申报	系列：教师
	专业：计算机科学与技术
	职称：副教授

业绩成果材料

（申报人的业绩成果材料包括论文、科研项目、获奖以及其他成果等）

单　位（二级单位） 数学与信息学院

姓　名 张广煜

材料核对人：　　单位盖章：

核对时间：

华南农业大学制

目　录

【佐证材料切记与目录页所列页码对应，不要用图片格式的材料进行打印。】

ART AND DESIGN PRESS INC
(United States)

2025年14期
ISSN(O): 3066-8557
ISSN(P): 3066-8549

职业发展与教育

VOCATIONAL DEVELOPMENT AND EDUCATION

职业发展与教育

Vocational Development and Education

半月刊

第1卷 第14期　2025年9月刊第二周

主　管　ART AND DESIGN PRESS INC.

主　办　ART AND DESIGN PRESS INC.

编　辑　《职业发展与教育》编辑部

ISSN(O): 3066-8557

ISSN(P): 3066-8549

地　址：119 S Atlantic Blvd, Suite 300D Monterey Park, CA 91754

网　址：https://www.artdesignp.com

本刊说明：

凡向本刊所投稿件，全体作者需签署论文著作权转让声明书和论文发表承诺书，声明、承诺及相关事项如下：

1. 作者将论文的复制权、发行权、网络传播权、翻译权、汇编权、信息网络传播权、改编权等著作权在世界范围内免费转让给本刊。
2. 论文不侵犯他人著作权和其他权利，否则作者将承担由此产生的全部责任，并赔偿由此给出版单位造成的全部损失。
3. 论文署名作者享有该作品的完全著作权，署名作者的身份真实。
4. 论文未曾以任何形式公开发表过。
5. 作者所投本刊稿件，本刊编辑部拥有修改权。

目录 CONTENTS

创新教育 | INNOVATIVE EDUCATION

职教前沿 | FRONTIERS OF VOCATIONAL EDUCATION

专业建设 | PROFESSIONAL DEVELOPMENT

基于多模态教学模式的《数据库应用》教学改革实践

张广煜
华南农业大学 数学与信息学院，广东 广州 510642
DOI: 10.61369/VDE.2025140037

摘　　要： 鉴于《数据库应用》课程理论抽象与实践性强的双重特征，本研究融合多模态教学理念与传统教学方法，提出“五维协同”教学模式。通过系统剖析多模态教学模式在课程中的价值意蕴，论证其在数据库教学中的必要性以及实践意义。经过连续两个学期的教学实证，该模式能显著提升课程教学效果。其以多模态认知映射为核心的改革范式，对同类技术应用型课程具有可迁移的实践价值。

关 键 词： 多模态教学模式；数据库应用；教学改革

Practice of Teaching Reform in Database Application under Multimodal Teaching Mode

Zhang GuangYu
College of Mathematics and Informatics, South China Agricultural University, Guangzhou, Guangdong 510642

Abstract: Based on the dual characteristics of theoretical abstraction and strong practicality of the course "Database Applications", this study integrates multimodal teaching concepts with traditional teaching methods and proposes "Five Dimensional Collaborative" teaching model. By systematically analyzing the value and implications of multimodal teaching mode in the curriculum, this paper demonstrates its necessity and practical significance in database teaching. After two consecutive semesters of teaching experience, this model can significantly improve the effectiveness of course teaching. Its reform paradigm centered on multimodal cognitive mapping has transferable practical value for similar technology application courses.

Keywords: multimodal teaching mode; database application; teaching reform

《数据库应用》课程是专门针对非计算机类本科生而开设的通识必修课，主要培养学生在数据库方面的理论、设计和应用能力，提升学生综合素质，强化信息管理和信息应用能力，为能够适应从事复杂数据库系统研究、设计、开发与应用工作的需求打下扎实的基础。

传统的《数据库应用》课程教学模式上课以 PPT 为主，课程内容呈现方式单一。教学过程中使用传统的教学方法，缺少师生互动，基本停留在老师教、学生学的阶段。课下师生交流时间少，难以及时解决学生在学习过程中遇到的疑难问题。在教学质量管控方面，教师无法及时掌握教学过程中动态的相关数据，难以及时发现问题。因此，有必要对《数据库应用》课程进行改革。

一、多模态教学模式

Stein在2000年发表的论文《反思教育中的权力：一种多模态教学法》中，提出了多模态教学模式的概念和框架[1]。多模态教学模式是一种基于多模态理论的教学方法，它强调在教学中整合多种感官通道和符号资源（模态）来呈现知识、促进互动和深化理解。其核心在于：利用人类多元化的感知和表达方式，创造更丰富、包容且有效的学习体验。

我国在多模式教学模式方面的研究主要应用于英语、汉语等语言教学领域[2-6]，并取得了较好的教学效果。在工程教育领域，多模态教学模式也涉及了部分相关课程[7-11]，但是在数据库应用课程的教学研究较少。因此，在数据库应用教学中运用多模态教学模式仍有很大的研究空间。将多模态融入数据库应用教学，不仅为数据库应用老师提供了理论框架和教学模式，还能更新现有的教学理念，丰富国内多模态教学模式在数据库应用教学方面的研究。同时也为一线教师开展数据库应用课程提供参考，为工程类课程的教学与研究开辟了新的可能。

二、多模态教学模式在《数据库应用》课程中的实践意义

多模态教学模式是整合多元感知通道的综合性教育范式。该模式通过协同激活触觉、听觉、视觉等感知系统，构建多维度的学习场景与信息输入方式。借助跨通道的认知体验，显著强化知

识理解与掌握效率，支持个性化学习路径，提升实践应用能力，激发协作互动效能。在《数据库应用》课程中实施多模态教学，可增强知识内化深度与学习驱动力，显著提升教学效能，系统化塑造学生综合素养，培育跨学科迁移能力。

在《数据库应用》课程教学中运用多模态的教学模式，具有以下实践意义。

1.构建多通道认知路径

传统数据库教学依赖单维度的口头讲授与 PPT 载体，导致学生对抽象概念没有深度理解、融会贯通。多模态教学通过整合视觉化演示（如图解 ER 模型）、动态交互（SQL 操作模拟）、听觉辅助（流程讲解音频）等媒介，激活多元感知通路。例如：使用 ACCESS 软件动态呈现数据表创建、查询优化的执行过程；通过 VBA 实时展示变量值的变化以及程序执行的结果。这种具象化认知体验将抽象理论转化为可观测现象，显著提升概念理解度与记忆留存率。

2.构建阶梯式实践体系

多模态实践是《数据库应用》课程能力培养的核心引擎。需重构课堂形态，通过分层任务驱动与逆向教学设计激发学生主体性，促成程序性知识的内化。制作微视频分解窗体、宏命令、SQL 优化等不同关键操作（如将课程中较难理解和掌握的宏操作与窗体结合的范例，创建制作流程录屏），支持学生的碎片化学习，推动理论认知向实践能力转化。构建 ACCESS 认知 – 操作融合训练矩阵，通过结构化异常注入与可视化逆向工程强化理论映射，构建认知 – 操作 – 创新的三维训练机制，使学生从被动接受知识转变为课程内容学习的主动建构者。

3.构建认知适应性学习系统

不同学生对数据库应用课程的认知有所差异，传统的课堂统一教学方式不能满足不同学生的学习需求。多模态的教学模式则提供了文字、音频、视频等多种学习资源，学生可以自主选择多模态中的一种或者几种，进行个性化的学习。多模态教学模式通过提供丰富的线上资源，随时满足学生对知识查缺补漏的需求，从而让学生迅速掌握相关知识。通过构建认知适应性学习系统，多模态教学模式能够更好提升学生的学习效果。

4.强化协作与互动机制

在《数据库应用》课程体系中，协作学习与师生或生生互动是提升教学效果的关键维度。传统授课模式往往难以提供充足的协作空间，制约了学生团队协作潜能的发挥。相比之下，融入多模态元素的教学策略，能够有效激活课程的协作与互动氛围。例如，当今 qq 群作为一款功能强大的社交工具被广泛应用于学生课程讨论、作业提交、资料分享等场景，借助协作讨论活动、课堂演示等多样化形式，引导学生完成共同观察、问题研讨及解决方案制定，从而系统性地培养其沟通能力和团队协作精神。更进一步，多模态教学模式还能利用 qq 群、微信群等数字化教学平台和工具，大幅提升师生互动的便利性与频率，实现问题的实时答疑与学习反馈的及时传递，有力激发学生更积极主动地投入学习过程。

5.提升知识记忆的长效性

《数据库应用》课程涵盖关系运算、数据库优化等理论体系与抽象概念。在传统教学模式下，学生常面临概念混淆或某些知识点难以理解的挑战。而引入多模态教学模式，则能通过课堂上的雨课堂练习，以及课下的雨课堂作业和上机作业等多元化的学习体验与信息呈现方式，在学习进程中同步激活学习者的多重感知通道。这种多感官协同刺激显著增强了知识的记忆强度。学生在多模态环境中形成的深刻认知印记与实践体验，极大地提高了知识在长期记忆中的留存率，并有效促进其在复杂应用场景中的高效检索与迁移能力。

综上所述，将多模态教学模式融入《数据库应用》课程，能够构建跨模态感知学习路径、提升技术实践转化效能、提供精准化学习支持框架、激活协作型学习生态。通过系统性整合多模态教学的核心要素与创新方法，可同步提升学生的知识内化率与课堂参与深度，全面培育其应对复杂工程问题的综合素养，从而为未来数据密集型工程实践提供坚实的能力支撑。

三、多模态教学模式在《数据库应用》课程中的实践架构

立足应用型高校人才培养导向，结合多模态教学的核心价值与独特优势，本文构建了适合农林院校的"五维协同"教学模式，如图 1 所示。该模式通过"听觉输入 – 视觉解析 – 言语交互 – 文本建构 – 操作实践"的跨通道认知循环，致力于建构全景式学习场景，深度激活学生对数据库知识的体系化理解与工程化应用能力。

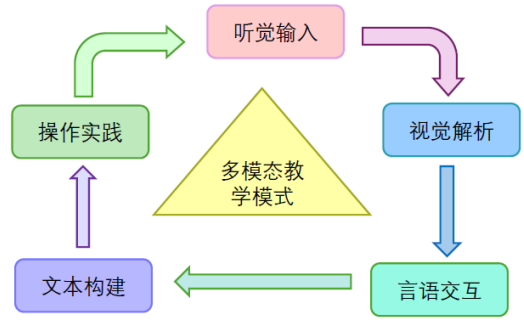


图 1 “五维协同”教学模式

1.听觉输入

听讲是课程教学的基石环节，主要通过教师课堂讲授的形式展开。其目标是系统传递《数据库应用》的核心知识体系，点燃学生的求知欲，并协助其初步搭建学科的理论框架。

在讲授过程中，主动融入贴近现实的案例或应用情境，以提升学生的参与度与好奇心。对于抽象概念与原理，采用具象化的比喻与类比进行阐释，促进学生形成直观的理解模式，从而更有效地吸收与记忆知识点。同时，对《数据库应用》课程较难理解的例题，录制讲解音频，便于学生课前预习和课后复习。引导学生洞察理论知识与实践应用的内在关联，深化其对知识的理解与解决实际问题的能力。

2.视觉解析

该环节需融合课堂导学与自主研学。教师依托数字媒体资源

（如投影、动态演示、实操录像等），将抽象理论与操作过程转化为可视化内容，提升认知具象度。针对《数据库应用》核心原理及实验步骤，可借助 ACCESS 平台开展实时交互演示，引导学生通过直接观察，进一步加深理解课程内容。课下教师布置课程的相关问题，学生通过查阅相关课程有关资料或重复观看上课录制的视频才能解决。该环节有效强化了学生的信息检索与创新探索能力。

3. 言语交互

转变教学范式，将学生置于知识建构的中心，通过主动阐释与表达深化其对课程内容的理解。打破传统单向传授，构建双向互动的课堂生态。积极引导学生提出疑问、参与深度研讨。教师针对学生生成性问题进行即时解析与拓展阐释，在动态问答中促进知识的深度内化。实施“学生讲师”机制，将特定课堂时段交由学生主导。学生需承担教师角色，对选定内容进行结构化讲解与演示。教师角色转换为引导者与促进者，提供过程性支持。通过融合同伴互评与教师专业点评的多维度反思性评估，培养了学生的知识整合能力、领导力、沟通协调能力及批判性思维，显著提升其知识内化程度与元认知能力。

4. 文本构建

教师依据课程目标设计相关习题，让学生将问题解析过程与解决方案进行文本构建，写出文字解答，以此促进知识内化，并提高学生的逻辑推理和系统表达素养。学生通过访问图书馆的文献索引、数据库及在线知识库，对习题内容做深度解构与延展探究。教师基于学生认知差异与兴趣导向，实施分层指导。培养学生解决问题的能力，并提高学生的学术写作的水平。

5. 操作实践

通过学生进行系统化练习，设计帮助学生掌握数据库应用知识的综合方案。对于理论知识方面，要求学生整理课程中的重要理论知识，通过堂上练习以及课后雨课堂的练习，助力学生厘清知识的思维路径，建构整个课程的系统性知识模型。对于实践方面，提供相关实验题目，让学生运用理论工具进行问题解构与方案验证，达成知识迁移与高阶工程思维的双重赋能。依托雨课堂的学习行为分析数据，教师可实施学情动态监测，从而精准调控教学内容结构与把控教学进度。

四、多模态教学模式下的《数据库应用》课程教改效果分析

1. 学生考核结果成效

为对多模态教学模式下的《数据库应用》课程进行教学效果评价，笔者将2022年春秋季、2023年春季的四个教学班做为研究对象，分别以传统的教学方式、多模态教学模式进行教学。在保持课程考核方式及难易程度一致的情况下，实施多模态教学模式的教改班级平均分比使用传统模式教学的班级平均分高3.35分。

2. 课程分析

从学生考核的结果来看，在《数据库应用》课程中运用多模态教学模式，取得了较好的教学效果，学生能更全面地掌握课程知识，成绩的整体水平也得到提升。评教时学生反馈的信息，也认同了多模态教学模式。

3. 存在问题

多模态教学模式提供了丰富的教学资源，有利于大部分学生采用灵活机动的方法快速掌握所学知识。但是也有个别学生不能统筹各种学习资源，学习能力不足，自控力较差，无法及时跟上课程的学习进度，从而产生畏难和懈怠情绪，最终导致成绩不合格。因此，多模态教学模式教改过程中，下一步应当注重全体学生的学习状态，及时调整课程内容，因材施教，做个性化的指导。

五、结束语

将多模态教学模式应用在《数据库应用》课程中，取得了较好的教学效果，学生能够在较短的时间、花较少的精力全面掌握课程内容，学生的成绩得到了进一步的提高。学生参与课堂教学的积极性进一步增强，对课程的满意度也得到了提升。未来，将结合人工智能学习系统，指导学生统筹兼顾使用各种教学资源的方式方法，进一步提高学习效率等做进一步的研究。同时，将多模态教学模式进一步拓展到 Python 程序设计、人工智能等其他课程教学中。

参考文献

[1]Stein，P.Rethinking Resources：Multimodal Pedagogies in the ESL Classroom[J].TESOL Quarterly，2000，34(2)：333-336.

[2] 张征 . 多模态 PPT 演示教学与学生学习绩效的相关性研究 [J]. 中国外语 ，2010，7(3)：54-58.

[3] 肖好章 ，杨颖滢 . 我国外语教育多模态效应生态分析 [J]. 北京科技大学学报 (社会科学版)，2018，34(6)：1-9.

[4]. 贾启行 . 多模态视角下英语专业课程教学体系重构研究 [J]. 北京城市学院学报 ，2019，(5)：32-35.

[5] 陈新 . 基于多模态理论框架的汉语视听说教学模式设计与研究 [J]. 云南大学学报 (自然科学版)，2020，42(S1)：116-122.

[6] 楚佳 . 应用型本科院校商务英语翻译多模态教学模式创新研究 [J]. 高教学刊 ，2020，(23)：29-32.

[7] 宋虹 ，段桂华 ，张士庚 .” 操作系统原理” 课程思政与多模态教学融合应用研究 [J]. 大学 ，2021，(52)：109-111.

[8] 张辰艳 ，王前进 ，张妍妮 .” 蛋白质与酶工程” 混合式教学的多模态教学改革思考 [J]. 科教导刊 ，2023，(35)：127-129.

[9] 张鹃 ，疏淑丽 ，魏树国 ，周妍 ，吴问天 . 多模态理论视域下机械控制工程课程教学模式探讨 [J]. 黑龙江工业学院学报 ，2024，24(8)：36-40.

[10] 顾力豪 ，陈丹 . 基于在线开放课程的 ESP 多模态教学模式 [J]. 江苏工程职业技术学院学报 ，2021，21(1)：86-89.

[11] 潘竞霞 ，吴丽娜 ，叶小露 . 多模态教学模式在中医院” 超声诊断学” 中的应用效果 [J]. 教育教学论坛 ，2022，(42)：157-160.

https://www.artdesignp.com
119 SAtlantic Blvd, Suite 300D Monterey Park, CA91754
(626) 810-4480
editorial@artdesignp.com

ART AND DESIGN PRESS INC
(United States)

2025年16期
ISSN(O): 3066-8557
ISSN(P): 3066-8549

职业发展与教育

VOCATIONAL DEVELOPMENT AND EDUCATION

目录 CONTENTS

职业发展与教育

Vocational Development and Education

第1卷 第16期　2025年9月刊第四周

主　管　ART AND DESIGN PRESS INC.

主　办　ART AND DESIGN PRESS INC.

编　辑　《职业发展与教育》编辑部

ISSN(O): 3066-8557

ISSN(P): 3066-8549

地　址：119 S Atlantic Blvd, Suite 300D Monterey Park, CA 91754

网　址：https://www.artdesignp.com

创新教育 | INNOVATIVE EDUCATION

技能培养 | SKILLS TRAINING

职教前沿 | FRONTIERS OF VOCATIONAL EDUCATION

专业建设 | PROFESSIONAL DEVELOPMENT

基于教育生态学的高校班主任工作实践

张广煜
华南农业大学 数学与信息学院，广东 广州 510642
DOI: 10.61369/VDE.2025160010

摘　　要： 立足教育生态学理论框架，将高校班级视为微型教育生态系统。结合作者班主任工作实践，提出高校班主任工作应遵循三项原则：人本性原则、系统性原则和动态平衡性原则。据此，班主任应在育人教育、系统教育和科学管理层面着力，运用生态思维重构工作范式，把班级变成一个能自我管理、适应变化、不断进步、持续成长的集体，为高等教育高质量发展提供微观治理支撑。

关 键 词： 教育生态学；高校班主任；学生管理

Practices for Class Advisor of University Based on Educational Ecology

Zhang GuangYu
College of Mathematics and Informatics, South China Agricultural University, Guangzhou, Guangdong 510642

Abstract: Based on the theoretical framework of educational ecology,university classes are regarded as micro educational ecosystems. Based on the author's experience as a class advisor of of university, it is proposed that the work of university class advisor should follow three principles: the principle of human nature, the principle of systematicity, and the principle of dynamic balance. Based on this, the university class advisor should focus on educating students, providing systematic education, and implementing scientific management. They should use ecological thinking to reconstruct the work paradigm, turning the class into a collective that can self manage, adapt to changes, continuously improve, and grow, providing micro governance support for the high-quality development of higher education.

Keywords: educational ecology; class adviser of university; student development

高校班主任作为校方委派到班级的专职教育管理者，是高等学校从事德育工作，开展大学生思想政治教育的骨干力量，是大学生健康成长的指导者和引路人[1]。其核心使命在于通过系统性组织、教育引导及科学管理，培育学风优良、凝聚力强的先进班集体，使学生成为具备德智体美劳五育融合素养的可持续发展型人才，并成为支撑社会进步的中坚力量。高校班主任岗位的特殊价值体现于两个维度：宏观层面，班主任作为国家教育政策的基层执行主体，承载着落实党的教育方针的关键职能；微观层面，其工作效能直接关系大学生发展质量，肩负着立德树人的重大责任[2]。高校班级作为微型教育生态系统，其管理效能的优化亟需范式重构。该机制作为高等教育体系的关键子系统，其发展与创新须植根于教育生态学的理论体系。

一、教育生态学

1966年，英国学者阿什比（Ashby）开创性地提出“高等教育生态学”概念，被视为运用生态学原理与方法探究高等教育的先行者。1976年，美国学者克雷明（Cremin）在其著作《公共教育》中系统确立了“教育生态学”这一学科术语[3]。在国内，吴鼎福是该领域早期研究的代表性学者，他发表了论文《教育生态学刍议》[4]，并推出了我国大陆首部以《教育生态学》为题的学术专著。此后，任凯与白燕合作编写了该学科的第二部同名专著。至2000年，范国睿撰著并发行了第三部《教育生态学》。

吴鼎福强调，教育生态学即应用生态学理论及方法解析教育现象，并主张将生态教育全面纳入各级学校教育体系，实现教育生态意识深度融入教学实践的各个环节。其研究视角侧重于环境问题引发的教育生态议题，呈现出较强的生态学理论特征[5]。任凯与白燕将教育生态系统界定为学科的核心研究对象，但在运用生态学原理进行演绎性分析方面相对有限。范国睿则提出，聚焦于教育生态系统本身，深入考察其与外部社会子系统（及其构成要素）之间复杂的物质、能量交换关系，是教育生态学研究的核心命题[6]。这一观点将教育生态系统定位为社会生态系统中一个具备相对独立性的子系统，标志着该领域认知的一次关键性突破[7]。

二、基于教育生态学视角的高校班主任工作核心原则

大学生作为教育生态学系统中的“生态主体”，其成长与发展受到班级、宿舍、院系、社会、家庭等多层次环境因子的综合影响，班级本身构成了一个动态开放的“微生态系统”。基于此，教育生态学为班主任工作提供了关键的理论指引，引导其超越孤立视角，系统剖析学生行为表现背后的环境诱因，并预判管理决

策对班级生态可能产生的连锁效应。其核心目标在于培育积极向上的班风，营造良好的学风，协调多元育人力量形成合力，最终构建和谐共生的班级育人环境。因此，将教育生态学原理深度融入高校班主任工作，已成为提升科学性、系统性与有效性的关键路径。

作为高等教育生态系统中的关键构成要素，高校班主任的工作原则本质上要对大学教育规律及教育生态学理论深刻把握，并在实践探索中将其贯穿育人全程。具体而言，高校班主任工作根本原则涵盖：以人为本的价值导向（人本性原则）、整体关联的系统思维（系统性原则）、适应演化的动态调节（动态平衡性原则）。

（一）人本性原则

教育生态学作为教育学的一个分支，核心思想就是“教育要以人为本”。在高校班主任工作中，班主任与学生共同构成了生态互动的核心主体，其他所有工作都要围绕这个核心关系展开。落实以人为本的原则，需要从两方面同时入手：

首先牢固确立“学生中心”的教育理念，深刻认识学生在教育过程中的主体性地位。充分保障学生的人格尊严，着力培育其个性化发展潜能。以服务学生全面发展为根本宗旨，覆盖全体学生需求，致力于学生终身成长。

另外要充分认识班主任作为思想政治教育工作者在落实立德树人根本任务中的关键角色。班主任作为国家教育政策的基层传导者，尤其承载着传导党的教育方针、巩固意识形态阵地、引领学生思想成长的关键政治职能；其工作效能直接关乎大学生培养质量，关系到每位学生的健康成长。

（二）系统性原则

系统思维是教育生态学的本质属性，系统性原则坚决摒弃任何割裂整体、孤立审视局部或个体的分析范式。强调所有教育现象都存在于复杂的关联网络之中，与环境要素间存在密不可分的相互作用。

高校班主任工作的成效，受到校内校外多种因素的共同影响。因此，用系统性思维看问题，这不仅是教育生态学的基本思路，更是班主任做好工作的必备能力，也是整个工作机制顺畅运行的关键。具体而言：班主任工作机制内部各要素（如班主任、学生、教育内容与方法、学校环境等）环环相扣、互相依赖，班主任工作持续与外部环境因子（家长、社会等）进行着动态的物质交换、能量传导与信息传递。正是这种内外部要素的协同作用，让整个高校班主任工作体系能健康、稳定、长久地运转。

（三）动态平衡性原则

从教育生态学视角审视，高校班主任工作机制本质上是一种具有开放性的动态网络结构，而非静态封闭系统。其运行状态始终处于动态演化之中，持续与内外部环境进行着复杂的交互作用。

因此，落实动态平衡原则需把握两个辩证统一的维度：深刻认识高校班主任工作机制运行过程固有的、持续的变化特性。在动态演化框架下，理解系统平衡的相对性本质。当机制内要素发生变动（如政策调整、学生群体变化、新引入等），系统必然经历短暂的失衡状态。然而，得益于系统内在的自组织调节功能，这种失衡将作为演化动力，驱使系统朝向新的、相对稳定的有序状态（即新的动态平衡点）演进。

三、高校班主任工作实践

（一）基于人本性原则的高校班主任工作实践

教育生态学视学生为具有主观能动性的生态主体，人本性原则强调班主任工作必须以学生的全面发展和个性化成长为核心，尊重其主体地位和发展需求。摒弃一刀切模式，通过深度谈心、学业分析、兴趣调查、生涯规划指导等方式，精准识别每位学生的优势潜能（“生态优势”）、发展短板（“生态限制因子”）及独特需求（“生态位”）。建立动态更新的学生成长档案[8]。

根据“生态位”识别结果，提供定制化支持。如针对学业发展受困学生群体，建立精准支持系统，实施个性化成长帮扶方案。班主任需以真诚、尊重、共情的态度与学生沟通，关注学生内在成长动力。通过成长性评价（关注个人进步）、成功体验强化（搭建展示平台）、挫折教育引导，帮助学生认识自我价值，明确发展目标，提升面对挑战的“生态适应力”与心理韧性。关注因挫折而产生的行为或心理上的异常现象的学生，及时开导和积极鼓励这些学生，有效预防自暴自弃等危险行为。对处于即将毕业关键期的学生，应协同疏导其面临的学术压力与职业焦虑，给予正确的指引和心理开导，实现毕业论文写作与就业的双赢[9]。倡导学生参加竞赛或申报大学生创新创业项目，为有科研潜质的学生搭建实验室或项目参与平台。尊重学生的兴趣特长，鼓励在适宜的“生态位”上绽放光彩。

打破传统单向管理，营造平等、民主、互助的班级氛围。鼓励学生参与班规制定、活动策划、评奖评优等班级事务，赋予其决策权与责任感。推广学习小组、朋辈导师制[10]等，促进知识共享与能力互助。构建班级的“学习共同体”。引导学生进行自我管理与生涯规划，成为自身成长的“第一责任人”，激发学生的内在动力与韧性。关注学生情感需求，及时察觉情绪波动，提供情感支持。在班级中大力倡导尊重差异、包容多元的价值观。反对任何形式的歧视，关注弱势群体需求（如经济困难生、特殊群体学生），营造让每个个体都感到安全、被接纳、有归属感的“班级心理微气候”，这是个体潜能释放的基础生态条件。

（二）基于系统性原则的高校班主任工作实践

系统性原则要求班主任将班级视为一个开放、复杂的微生态系统，并置于更大的学校、社会生态网络中，强调整体关联与资源整合，协同育人。

构建班级内部协同网络，优化班级内部结构，明确班委分工与协作机制，发挥班委会的引领与协调作用。通过班级口号、班服等文化符号建设，持续传递正向价值观，培育积极向上的班风，形成具有凝聚力的健康“群落”。建立有效的沟通反馈渠道（如线上平台、定期班会等），确保信息流畅通。设计多样化的集体活动，例如：社会实践、志愿者服务等，创造深度互动机会，打破小团体壁垒，促进不同特质学生间的理解、互助与优势互补，形

成积极互赖的“共生”关系。促进生生互动共生。着力培育严谨求实的学风、团结互助的班风、开放包容的人际氛围。营造浸润式“班级文化气候”。

班主任要做“生态位”拓展与资源链接，主动识别并链接外部优质资源。与专业课教师深度沟通，反馈学情，协同解决学业问题；与辅导员紧密配合，信息共享，共同应对思想心理危机；联动就业指导中心、创新创业学院等部门，为学生提供专业和个性化服务。积极与学生家长联系，形成家校育人合力；邀请优秀校友、行业导师开展讲座，对接社会实践基地、企业资源，拓展学生视野与实践平台。将家庭、学校、社会资源编织成支持学生成长的立体“生态网”。构建多元协同育人“生态链”。

（三）基于动态平衡性原则的高校班主任工作实践

班级生态系统处于永恒的变化中，动态平衡性原则要求班主任具备敏锐的洞察力和灵活的调控能力，及时应对内外部扰动，维护系统健康与活力。

运用信息化工具（如学工系统、学习平台数据）、深度访谈、班委观察等，持续跟踪学生学业进展、心理状态、人际关系、活动参与度等关键指标的变化趋势。建立“生态位”动态监测。设定关键阈值（如学业预警线、心理危机信号），建立预警模型。一旦监测到个体偏离“生态位”（如突然成绩下滑、社交退缩）或群体出现失衡苗头（如小团体对立、消极情绪蔓延），立即启动干预预案，如个别谈话、调整活动安排等，防止问题升级。根据问题的性质、严重程度及系统反馈，灵活选择干预策略与力度，因时因势施策。如对学业困难生，可调整帮扶计划或联系教师微调教学；面对外部冲击（如政策调整、突发事件），及时解读信息，疏导情绪，调整班级工作重心[11]。

定期评估班级内部资源（如干部作用发挥、活动效果）和外部链接的有效性。对表现不佳的“生态位”进行调整或重组；对低效或冗余的外部链接进行优化或替换，确保资源流向最需要、最能产生效益的地方。不仅关注学业成绩等显性指标，更要重视学生满意度、班级凝聚力、个体成长体验等隐性指标。通过个人谈话、座谈会等方式进行多维度效果评估。将评估结果及时反馈给相关学生、班委及相关教师，进行反馈驱动优化。反思工作策略的有效性，总结经验教训，持续优化班级管理制度、活动设计、资源链接方式和支持策略，形成“监测－干预－评估－反馈－优化”的动态闭环管理，推动班级生态系统螺旋式上升。

四、结语

高校班主任工作是一项充满挑战与创造的事业。以教育生态学的人本性、系统性、动态平衡性原则为指导，推动班主任工作从经验型、事务型向研究型、生态型转变，是提升新时代高校育人质量的必由之路。学生来自天南海北，个性与禀赋差异不同，唯有尊重个体生命的独特性与成长规律，整合优化育人生态系统的结构与功能，敏锐感知并有效调节系统的动态变化，方能营造出和谐稳定、生机勃勃、有利于每一位学生全面发展和终身成长的优质教育生态，真正落实立德树人的根本任务，培养担当民族复兴大任的时代新人。

参考文献

[1] 教育部 . 关于加强高等学校辅导员、班主任队伍建设的意见：教社政〔2005〕2 号 [A/OL].(2005-01-13)[2025-08-10]. http://www.moe.gov.cn/s78/A12/szs_lef/moe_1407/moe_1409/s3016/s3017/201006/t20100608_88984.html.

[2] 苏艳芳，林卉，林志彬 . 新形势下高校班主任工作实践与思考 [J]. 高等农业教育 ,2004,(5):75-78

[3] 贺祖斌 . 高等教育生态研究述评 [J]. 广西师范大学学报（哲学社会科学版）,2005,41(1):123-127.

[4] 吴鼎福 . 教育生态学刍议 [J]. 南京师大学报（社会科学版）,1988，(2):33-36+7.

[5] 吴鼎福，诸文蔚 . 教育生态学 [M]. 南京：江苏教育出版杜 ,1990:19.

[6] 范国睿 . 教育生态学 [M]. 北京：人民教育出版社 ,2000:31.

[7] 刘洋洋 . 教育生态学视角下高校班主任育人机制研究 [D]. 北京：华北电力大学 ,2020:7-8.

[8] 李景春 . 生态位理论视域中的教育生态系统及其发展 [J]. 教育科学 . 2006，22(3):26-29.

[9] 杨振刚 . 人本管理在高校班主任工作中的体现 [J]. 中山大学学报论丛 ,2006,26(12):113-115.

[10] 贺瀚，邓人心，张菁 . 大学生朋辈教育在学风建设中的作用探究 [J]. 黑龙江教育（高教研究与评估）,2023,(1):85-88.

[11] 郭鑫 . 论高校班主任工作的 " 四个有 "[J]. 思想政治教育研究 ,2009，25(2):121-123.

https://www.artdesignp.com
119 SAtlantic Blvd, Suite 300D Monterey Park, CA91754
(626) 810-4480
editorial@artdesignp.com

国家自然科学基金资助项目批准通知

（包干制项目）

张广煜　先生/女士：

根据《国家自然科学基金条例》、相关项目管理办法规定和专家评审意见，国家自然科学基金委员会（以下简称自然科学基金委）决定资助您申请的项目。项目批准号：62206099，项目名称：面向不完备多视图数据的高阶子空间聚类方法及其应用，资助经费：30.00万元，项目起止年月：2023年01月至 2025年 12月，有关项目的评审意见及修改意见附后。

请您尽快登录科学基金网络信息系统（https://isisn.nsfc.gov.cn），**认真阅读《国家自然科学基金资助项目计划书填报说明》并按要求填写《国家自然科学基金资助项目计划书》（以下简称计划书）**。对于有修改意见的项目，请您按修改意见及时调整计划书相关内容；如您对修改意见有异议，须在电子版计划书报送截止日期前向相关科学处提出。

请您将电子版计划书通过科学基金网络信息系统（https://isisn.nsfc.gov.cn）提交，由依托单位审核后提交至自然科学基金委。自然科学基金委审核未通过者，将退回的电子版计划书修改后再行提交；审核通过者，打印纸质版计划书（一式两份，双面打印）并在项目负责人承诺栏签字，由依托单位在承诺栏加盖依托单位公章，且将申请书纸质签字盖章页订在其中一份计划书之后，一并报送至自然科学基金委项目材料接收工作组。纸质版计划书应当保证与审核通过的电子版计划书内容一致。**自然科学基金委将对申请书纸质签字盖章页进行审核，对存在问题的，允许依托单位进行一次修改或补齐。**

向自然科学基金委提交电子版计划书、报送纸质版计划书并补交申请书纸质签字盖章页截止时间节点如下：

1. **2022年10月8日16点：**提交电子版计划书的截止时间；

2. **2022年10月14日16点：**提交修改后电子版计划书的截止时间；

3. **2022年10月19日：**报送纸质版计划书（一式两份，其中一份包含申请书纸质签字盖章页）的截止时间。

4. **2022年10月28日：**报送修改后的申请书纸质签字盖章页的截止时间。

请按照以上规定及时提交电子版计划书，并报送纸质版计划书和申请书纸质签字盖章页，逾期不报计划书或申请书纸质签字盖章页且未说明理由的，视为自动放弃接受资助；未按要求修改或逾期提交申请书纸质签字盖章页者，将视情况给予暂缓拨付经费等处理。

附件：项目评审意见及修改意见表

国家自然科学基金委员会

2022年9月7日

附件：项目评审意见及修改意见表

项目批准号	62206099	项目负责人	张广煜	申请代码1	F0605
项目名称	面向不完备多视图数据的高阶子空间聚类方法及其应用				
资助类别	青年科学基金项目		亚类说明		
附注说明					
依托单位	华南农业大学				
直接费用	30.00 万元		起止年月	2023年01月 至 2025年12月	

通讯评审意见:

<1>具体评价意见:

一、该申请项目的研究思想或方案是否具有新颖性和独特性？请详细阐述判断理由。

该项目研究不完备多视图数据的高阶子空间聚类方法及其应用，高阶子空间聚类是该领域重点研究问题。该项目提出针对视图数据缺失问题的研究具有一定创新性，从申请书描述的不同图像的差异，文档数据差异，信息缺失是不偏存在的问题。

二、请评述申请项目所关注问题的科学价值以及对相关前沿领域的潜在贡献。

该项目研究问题包括：多视图数据缺失值，高阶多视图表征学习，不完备多视图子空间聚类。缺失值问题属于传统数据挖掘问题，多视图缺失信息分析和子空间聚类的提出解决了该领域研究忽视的问题，对该方向的研究有一定科学价值和贡献。

三、请评述申请人的创新潜力与研究方案的可行性。

申请人在该相关领域发表过一些高水平论文，参加过国家项目，具备在该领域开展研究的创新潜力。研究方案从缺失值补全，提出基于数据自表达补全，基于深度学习补全，提出加权张量高阶多视图表征学习，子空间聚类，并给出医学，遥感影像应用思路。

四、其他建议

申请书排段比较混乱。

<2>具体评价意见:

一、该申请项目的研究思想或方案是否具有新颖性和独特性？请详细阐述判断理由。

该项目拟从理论、建模、应用三个层次出发，结合数据自表达、深度学习、多视图关联分析、多源信息融合与张量低秩近似等理论和技术，研究建立面向不完备多视图数据的高阶子空间聚类新方法，并验证这些新方法在生物信息学与地理信息科学等交叉领域的实际应用。研究思想具有独特性。

二、请评述申请项目所关注问题的科学价值以及对相关前沿领域的潜在贡献。

申请项目在多视图子空间聚类研究方面具有重要的科学价值，对相关学科前沿具有一定的贡献。

三、请评述申请人的创新潜力与研究方案的可行性。

申请人具有较好的研究基础，显示了较强的创新能力。研究方案详细，具有可行性。

四、其他建议

<3>具体评价意见:

一、该申请项目的研究思想或方案是否具有新颖性和独特性？请详细阐述判断理由。

该申请项目针对不完备多视图子空间聚类问题中，如何构建统一模型同时恢复各视图下的缺失数据以及挖掘跨视图间的高阶相关性等难点，研究内容包括数据潜在结构的多视图缺失值补全、高阶信息研究的多视图表征学习、多视图缺失值补全与高阶表征学习的多视图子空间聚类，研究内容自成体系，并探讨高阶不完备多视图子空间聚类的应用研究，在生物信息学、地理信

息科学等交叉学科具有重要实际应用价值，同时进一步丰富了数据挖掘与机器学习的理论方法。

二、请评述申请项目所关注问题的科学价值以及对相关前沿领域的潜在贡献。
该申请项目拟研究不完备多视图子空间聚类方法及应用，以多视图缺失值补全与高阶结果信息融合为主要的切入点，建立适用于不完备数据的多视图子空间聚类新框架，对问题分析全面、目标明确，预期成果将解决多视图数据存在任意缺失值、高阶多视图表征学习和高效优化建模等多个问题。

三、请评述申请人的创新潜力与研究方案的可行性。
项目申请人多年专注多视图子空间聚类研究，并取得较好的科研成果，研究基础可行性较高。提出了一系列基于高阶信息的不完备多视图子空间聚类方法，研究方案详实，项目的方案可行性较高。

四、其他建议
无

修改意见：

信息科学部

2022年9月7日

129

项目批准号	62206099
申请代码	F0605
归口管理部门	
依托单位代码	51064208A0499-0932

622060991001701

国家自然科学基金

资助项目计划书

（包干制项目）

资助类别：青年科学基金项目

亚类说明：

附注说明：

项目名称：面向不完备多视图数据的高阶子空间聚类方法及其应用

资助经费：30万元　　执行年限：2023.01-2025.12

负 责 人：张广煜

通讯地址：广东省广州市五山路483号大院华南农业大学数学与信息（软件）学院512办公室

邮政编码：　　电　　话：机密

电子邮件：guangyuzhg@foxmail.com

依托单位：华南农业大学

联 系 人：唐家林　　电　　话：020-85280070

填表日期：2022年09月21日

国家自然科学基金委员会制

Version：1.001.701

国家自然科学基金资助项目计划书填报说明

（包干制项目）

一、项目负责人收到《国家自然科学基金资助项目批准通知》（以下简称《批准通知》）后，请认真阅读本填报说明，参照国家自然科学基金相关项目管理办法和新修订的《国家自然科学基金资助项目资金管理办法》（以下简称《资金管理办法》，请查阅国家自然科学基金委员会官方网站首页“政策法规”栏目），按《批准通知》的要求认真填写和提交《国家自然科学基金资助项目计划书》（以下简称《计划书》）。

二、填写《计划书》时要科学严谨、实事求是、表述清晰、准确。《计划书》经国家自然科学基金委员会相关项目管理部门审核批准后，将作为项目研究计划执行、检查和验收的依据。

三、《计划书》各部分填写要求如下：

（一）简表：由系统自动生成。

（二）摘要及关键词：各类获资助项目都应当填写中、英文摘要及关键词。

（三）正文：

1. 青年科学基金项目：如果《批准通知》所附“项目评审意见及修改意见表”中“修改意见”栏目没有修改要求的，只需选择“研究内容和研究目标按照申请书执行”即可；如果《批准通知》中上述栏目明确要求调整研究期限或研究内容等的，须选择“根据研究方案修改意见更改”并填报相关修改内容。

2. 国家杰出青年科学基金项目和优秀青年科学基金项目按下列提纲撰写：

（1）研究方向；

（2）结合国内外研究现状，说明研究工作的学术思想和科学意义（限两个页面）；

（3）研究内容、研究方案及预期目标（限两个页面）；

（4）年度研究计划；

3. 科技管理专项项目按下列提纲撰写：

（1）科技战略研究领域方向；

（2）结合该领域国内外研究现状及发展趋势，分析科学基金资助战略与政策（限两个页面）；

（3）研究内容、研究方案及预期目标（限两个页面）；

（4）年度研究与实践计划。

四、资助经费相关要求：

1. 资助经费批准时不再区分直接费用和间接费用。

2. 项目负责人在提交计划书时需签署承诺书，承诺尊重科研规律，弘扬科学家精神，遵守科研伦理道德和作风学风诚信要求，认真开展科学研究工作；承诺项目经费全部用于与本项目研究工作相关的支出，不得用于与本项目研究无关的支出。

3. 项目负责人提交计划书时，无需编制项目预算。项目资金由项目负责人自主决

定使用，按照《资金管理办法》第九条规定的开支范围列支。有关管理费用的补助支出，由依托单位根据实际管理需要，在充分征求项目负责人意见基础上合理确定。绩效支出由项目负责人根据实际科研需要和相关薪酬标准自主确定，依托单位按照工资制度进行管理。其余用途经费无额度限制，由项目负责人根据实际需要自主决定使用。

4. 项目结题时，项目负责人根据实际使用情况编制项目经费决算，经依托单位财务、科研管理部门审核后，报自然科学基金委。依托单位应当在单位内部公开非涉密项目立项、主要研究人员、资金使用（重点是间接费用、外拨资金、结余资金使用等）、决算、大型仪器设备购置以及项目研究成果等情况，接受内部监督。

5. 自然科学基金委结合项目管理，对经费使用情况和依托单位管理情况定期开展抽查。

简表

项目负责人信息	姓　　名	张广煜	性 别	男	出生年月	1991年05月	民 族	汉族
	学　　位	博士			职称	讲师		
	是否在站博士后	否			电子邮件	guangyuzhg@foxmail.com		
	电　　话	机密			个人网页			
	工 作 单 位	华南农业大学						
	所 在 院 系 所	数学与信息(软件)学院						
依托单位信息	名　　称	华南农业大学					代码	51064208A0499
	联 系 人	唐家林			电子邮件	kycjhk@scau.edu.cn		
	电　　话	020-85280070			网站地址	http://kjc.scau.edu.cn/		
合作单位信息	单 位 名 称							
项目基本信息	项 目 名 称	面向不完备多视图数据的高阶子空间聚类方法及其应用						
	资 助 类 别	青年科学基金项目			亚 类 说 明			
	附 注 说 明							
	申 请 代 码	F0605:模式识别与数据挖掘			F0212:数据科学与大数据计算			
	基 地 类 别							
	执 行 年 限	2023.01-2025.12						
	资 助 经 费	30万元						

项目摘要

中文摘要：

在多视图学习研究领域，不完备多视图子空间聚类是当今最具挑战性的科学研究问题之一，同时也是学者们关注的一个焦点。其关键难题在于如何构建统一的模型，同时恢复各视图下的缺失数据以及挖掘跨视图间的高阶关联性。本项目拟从理论、建模、应用三个层次出发，结合数据自表达、深度学习、多视图关联分析、多源信息融合与张量低秩近似等理论和技术，重点解决这一难题，研究建立面向不完备多视图数据的高阶子空间聚类新方法，并验证这些新方法在生物信息学与地理信息科学等交叉领域的实际应用。项目的开展将进一步丰富数据挖掘与机器学习的理论与方法，特别是促进多视图子空间聚类研究的深入发展。

Abstract:

Incomplete multi-view subspace clustering is one of the most challenging tasks in multi-view learning. The key issue lies in designing a one-step framework, which can impute the incomplete data matrices in each view and discover the latent higher-order correlations across different views jointly. In this project, based on the theories of self-expressive representation learning, deep learning, multi-view correlation analysis, higher-order information fusion and low-rank tensor approximation, we aim to solve the above issue from the levels of theories, modeling and applications. Accordingly, we will propose two novel higher-order subspace clustering methods for incomplete multi-view data, as well as their applications in the fields of bioinformatics and geographic information science. The project would further enrich the theories and methods of data mining and machine learning. In particular, it would enhance the research development of heterogeneous multi-view subspace clustering.

关键词(用分号分开)： 数据挖掘；多源数据挖掘；聚类；多视图聚类

Keywords(用分号分开)： Data mining; Multi-view data mining; Clustering; Multi-view clustering

报告正文

研究内容和研究目标按照申请书执行。

国家自然科学基金项目负责人、依托单位承诺书

国家自然科学基金项目负责人承诺书

本人郑重承诺：我接受国家自然科学基金的资助，严格遵守中共中央办公厅、国务院办公厅《关于进一步加强科研诚信建设的若干意见》《关于进一步弘扬科学家精神加强作风和学风建设的意见》《关于加强科技伦理治理的意见》等规定，及国家自然科学基金委员会关于资助项目管理、项目资金管理等各项规章，在《计划书》填写及项目执行过程中：

（一）按照《批准通知》《国家自然科学基金资助项目计划书填报说明》的要求填写《计划书》，未自行降低、更改目标任务或约定要求，或缩减研究（研制）内容；

（二）树立“红线”意识，严格履行科研合同义务，按照《计划书》负责实施本项目（批准号：62206099），切实保证研究工作时间，按时报送有关材料，及时报告重大情况变动，不违规将科研任务转包、分包他人，不以项目实施周期外或不相关成果充抵交差；

（三）遵守科研诚信、科技伦理规范和学术道德，认真开展研究工作，对资助项目发表的论著和取得的研究成果按规定进行标注，不在非本项目资助的成果或其他无关成果上标注本项目批准号，反对无实质学术贡献者“挂名”，不在成果署名、知识产权归属等方面侵占他人合法权益，并如实报告本人及项目组成员发生的违背科研诚信要求的任何行为；

（四）尊重科研规律，弘扬科学家精神，严谨求实，追求卓越，反对浮夸浮躁、投机取巧，不人为夸大学术或技术价值，不传播未经科学验证的现象和观点；

（五）将项目资金全部用于与本项目研究工作相关的支出，并结合科研活动需要，科学合理安排项目资金支出进度；

（六）做好项目组成员的教育和管理，确保遵守以上相关要求。

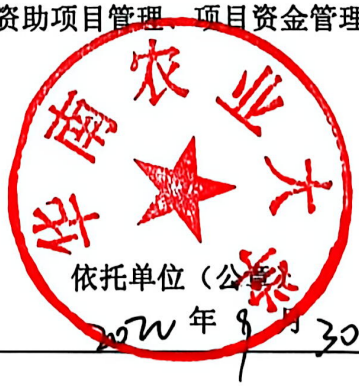
如违背上述承诺，本人愿接受国家自然科学基金委员会和相关部门做出的各项处理决定。

项目负责人（签字）：郑广煜

2022年 9月 29日

国家自然科学基金项目依托单位承诺书

我单位同意承担上述国家自然科学基金项目，将保证项目负责人及其研究队伍的稳定和研究项目实施所需的条件，严格遵守国家自然科学基金委员会有关资助项目管理、项目资金管理、科研诚信管理和科技伦理管理等各项规定，并督促实施。



依托单位（公章）

2022年 9月 30日

国家自然科学基金资助项目签批审核表

本栏目由自然科学基金委填写

科学处审查意见：

同意按照计划执行

吴国政

负责人（签章）：

年　月　日

2023-02-13

科学部审查意见：

同意

何杰

负责人（签章）：

年　月　日

2023-02-13

任务书编号：2024A04J4451

广州市科技计划项目

任务书

项目名称：	面向不完备数据的大规模多视图子空间聚类研究
承担单位：	华南农业大学
项目负责人：	张广煜
计划类别：	基础研究计划
专题名称：	2024年度基础与应用基础研究专题
支持方向：	青年博士“启航”项目
组织单位：	华南农业大学
起止时间：	2024-01-01 至 2025-12-31
主管处室：	基础研究处

广州市科学技术局制

二〇二四年

填写说明

1.任务书甲方为广州市科学技术局；乙方为项目承担单位；丙方为项目组织单位。

2.任务书基于项目申报书转换而成，请按照“广州科技大脑”提示在线填写核实，若存在不填写内容的栏目，请用“无”表示；任务书中的单位名称应为规范全称，并与单位公章一致。

3.乙方与合作单位的合作协议自动从项目申报书中读取，如需变化调整，须待任务书签订后，按要求及时办理重大变更。

4.乙方完成项目任务书在线填写，依次提交丙方和甲方审核确认后，按要求登录“穗好办”APP完成电子签章。不具备电子签章条件的单位，经与业务主管处室沟通对接后，可下载电子版项目任务书用A4纸双面打印装订签章；一式六份报甲方和丙方签章，其中甲方两份丙方两份，项目承担单位和项目负责人各一份。

5.涉密项目请在“广州科技大脑”下载项目任务书模板，按保密要求离线填写报送。

6.项目申报书是项目任务书填报的重要依据，未经甲方许可，乙方不得修改考核指标，调整主要研究内容。项目任务书将作为项目实施管理、验收结题和监督评估的重要依据。

7.项目任务书中的“备注”，包括重要的必须补充的内容。

8.“广州科技大脑”是项目管理过程中重要通知和文书的电子送达平台。为确保电子送达渠道畅通，乙方和项目负责人应及时更新维护“广州科技大脑”的单位和个人信息。

9.根据相关要求，项目涉及人体临床研究的，项目需经医学伦理委员会审查通过并在任务书附件栏上传相关佐证材料。

一、项目基本信息

项目基本信息	项目名称	面向不完备数据的大规模多视图子空间聚类研究
	申请市财政科技经费	5(万元)
	研究期限	2（年）
项目摘要	作为聚类研究的重点，多视图子空间聚类的目标在于融合多个视图间的互补信息以得到更加鲁棒的聚类结果。尽管国内外学者们提出了一系列多视图子空间聚类方法，但关于大规模不完备场景下的多视图子空间聚类研究仍十分匮乏。对于大规模且带有缺失值的多视图数据，当前多视图子空间聚类仍面临若干关键问题亟待解决，研究建立面向不完备数据的大规模多视图子空间聚类新框架。本项目的开展将进一步丰富数据挖掘与机器学习的理论与方法。	

二、项目单位情况

<table>
<tr><td rowspan="9">项目承担单位</td><td>单位名称</td><td colspan="2">华南农业大学</td><td>统一社会信用代码</td><td>12440000455416563 4</td></tr>
<tr><td>注册时间</td><td>1952-01-01</td><td colspan="2">单位类型</td><td>高等院校</td></tr>
<tr><td>注册地址</td><td colspan="4">广东省广州市天河区五山路483号</td></tr>
<tr><td>办公地址</td><td colspan="4">广东省广州市天河区五山路483号</td></tr>
<tr><td rowspan="3">联系人</td><td>姓名</td><td colspan="3">倪慧群</td></tr>
<tr><td>手机号码</td><td colspan="3">机密</td></tr>
<tr><td>电子邮箱</td><td colspan="3">kjcgxk@scau.edu.cn</td></tr>
<tr><td>开户银行</td><td colspan="4">广东广州工行五山支行</td></tr>
<tr><td>开户户名</td><td colspan="4">华南农业大学</td></tr>
<tr><td>银行账号</td><td colspan="4">3602002609000310520</td></tr>
</table>

三、项目负责人信息

姓名	张广煜	证件类型	身份证
证件号码	机密	性别	男
出生日期	1991-05-15	民族	汉族
国籍	中国	学历	博士研究生
学位	博士	学位授予国家（或地区）	中国
职务	教师	职称	中级
所学专业	计算机科学与技术	手机号码	机密
办公电话	020-85280070	电子邮箱	1538395488@qq.com

四、项目经费信息

本项目总投入：¥（5）万元，其中，市财政科技经费：¥（5）万元，自筹经费：¥（0）万元。

经费下达计划			
资金来源	小计	市财政科技经费	自筹经费
2024	5	5	0
总计	5	5	0

（单位：万元）

注：本专题纳入“包干制”，市财政科技经费按市科技计划项目经费“包干制”相关规定执行。

五、预期代表性成果

项目负责人在项目实施期内，以该项目作为资助项目获得以下5种情形之一且经费使用符合规定的，由组织单位审核后通过验收。

（一）项目实施期内，以第一作者/通讯作者发表论文1篇或以上（须标注资助项目编号）；

（二）项目实施期内，以第一完成人申请或授权专利、软件著作1项或以上；

（三）项目实施期内，获省级以上科技计划项目或人才项目支持1项或以上；

（四）项目实施期内，获省级以上科技奖励（含列入获奖团队成员名单）1项或以上；

（五）项目实施期内，获得职称晋升。

六、备注

专题补充约定条款：

甲方对未履行勤勉尽责义务的相关责任主体，自作出处理结论之日起，依照法律法规规定或任务书约定实施惩戒5年，取消相关责任主体申报市科技计划项目、申领市科技计划项目经费的资格。

预期代表性成果需在实施期内获得。

项目承担单位（乙方）及项目负责人承诺书

承诺书

本单位/本人作为广州市科技计划项目承担单位/项目负责人，将严格遵守广州市科技计划管理相关规定，严格履行自身责任，加强对项目组人员及合作单位的管理，在此郑重承诺：

（一）确保与本项目有关的全部材料真实、合法、有效，未侵犯其他方知识产权等权利，不存在多头申报、重复申报行为；

（二）严格遵守《广州市科技创新条例》《广州市科技计划项目管理办法》《广州市科技计划项目经费管理办法》《广州市科技计划科技报告管理办法》等相关规定，实施项目和经费管理；

（三）严格遵守国家、省、市关于科研诚信和科技伦理的有关法律、法规，相关政策以及各项规定，加强项目实施过程中的科研诚信及科技伦理管理，恪守科研道德准则。

如有违反，本单位/本人愿意接受相关部门做出的各项处理决定，包括但不限于终止项目、停拨经费、核减经费、追回经费，取消一定期限广州市科技计划项目申报资格，记入科研失信行为数据库，将不良行为向社会公开等。

项目承担单位：华南农业大学

日期：2023年12月19日

项目负责人：张广煜

日期：2023年12月19日

任务书签署

甲乙丙三方根据《广州市科技计划项目管理办法》《广州市科技计划项目经费管理办法》《广州市科技计划科技报告管理办法》等有关文件规定，以及有关法律、政策和管理要求，签署本任务书。

签订地点：广州市越秀区

广州市科学技术局（甲方）：广州市科学技术局
局项目经办人：蒋韬略　　联系电话：83124150
责任处室负责人：麦胜文

2024年01月17日

项目承担单位（乙方）：华南农业大学
二级部门：华南农业大学数学与信息学院
项目负责人：张广煜
项目经费汇入账号
账户名：华南农业大学　　账号：3602002609000310520
开户银行：广东广州工行五山支行
财务负责人：肖斐

2023年12月19日

组织单位（丙方）：华南农业大学
项目经办人：倪慧群

2023年12月20日

国家自然科学基金资助项目批准通知

（预算制项目）

王海燕　先生/女士：

根据《国家自然科学基金条例》、相关项目管理办法规定和专家评审意见，国家自然科学基金委员会（以下简称自然科学基金委）决定资助您申请的项目。项目批准号：62572201，项目名称：面向多源生物医学数据的高阶信息建模研究，直接费用：50.00万元，项目起止年月：2026年01月至2029年12月，有关项目的评审意见及修改意见附后。

请您尽快登录科学基金网络信息系统（https://grants.nsfc.gov.cn），**认真阅读《国家自然科学基金资助项目计划书填报说明》并按要求填写《国家自然科学基金资助项目计划书》（以下简称计划书）**。对于有修改意见的项目，请您按修改意见及时调整计划书相关内容；如您对修改意见有异议，须在电子版计划书报送截止日期前向相关科学处提出。

请您将电子版计划书通过科学基金网络信息系统（https://grants.nsfc.gov.cn）提交，由依托单位审核后提交至自然科学基金委。自然科学基金委审核未通过者，将退回的电子版计划书修改后再行提交；审核通过者，打印纸质版计划书（一式两份，双面打印）并在项目负责人承诺栏签字，由依托单位科研、财务管理等部门审核、签章并在承诺栏加盖依托单位公章，且将申请书纸质签字盖章页订在其中一份计划书之后，一并报送至自然科学基金委项目材料接收工作组。纸质版计划书应当保证与审核通过的电子版计划书内容一致。**自然科学基金委将对申请书纸质签字盖章页进行审核，对存在问题的，允许依托单位进行一次修改或补齐。**

向自然科学基金委提交电子版计划书、报送纸质版计划书并补交申请书纸质签字盖章页截止时间节点如下：

1. **2025年9月5日16点：**提交电子版计划书的截止时间；

2. **2025年9月12日16点：**提交修改后电子版计划书的截止时间；

3. **2025年9月23日：**报送纸质版计划书（一式两份，其中一份包含申请书纸质签字盖章页）的截止时间。

4. **2025年10月9日：**报送修改后的申请书纸质签字盖章页的截止时间。

请按照以上规定及时提交电子版计划书，并报送纸质版计划书和申请书纸质签字盖章页，逾期不报计划书或申请书纸质签字盖章页且未说明理由的，视为自动放弃接受资助；未按要求修改或逾期提交申请书纸质签字盖章页者，将视情况给予暂缓拨付经费等处理。

附件：项目评审意见及修改意见表

国家自然科学基金委员会

2025年8月27日

项目批准号	62572201
申请代码	F0213
归口管理部门	
依托单位代码	51064208A0499-0932

625722011004434

国家自然科学基金

资助项目计划书

（预算制项目）

资助类别：面上项目

亚类说明：

附注说明：

项目名称：面向多源生物医学数据的高阶信息建模研究

直接费用：50万元　　执行年限：2026.01-2029.12

负 责 人：王海燕　　BRID：08729.00.26952

通讯地址：广东省广州市天河区五山路483号华南农业大学

邮政编码：510642　　电　　话：机密

电子邮件：cshywang@scau.edu.cn

依托单位：华南农业大学

联 系 人：郑雪宜　　电　　话：020-85280070

填表日期：2025年09月03日

国家自然科学基金委员会制

Version：1.004.434

国家自然科学基金资助项目计划书填报说明

（预算制项目）

一、项目负责人收到《国家自然科学基金资助项目批准通知》（以下简称《批准通知》）后，请认真阅读本填报说明，参照国家自然科学基金相关项目管理办法和《国家自然科学基金资助项目资金管理办法》（以下简称《资金管理办法》，请查阅国家自然科学基金委员会门户网站首页“政策法规”栏目），按《批准通知》的要求认真填写和提交《国家自然科学基金资助项目计划书》（以下简称《计划书》）。

二、填写《计划书》时要科学严谨、实事求是、表述清晰、准确。《计划书》经国家自然科学基金委员会相关项目管理部门审核批准后，将作为项目研究计划执行、检查和验收的依据。

三、《计划书》各部分填写要求如下：

（一）简表：由系统自动生成。

（二）摘要及关键词：各类获资助项目都应当填写中、英文摘要及关键词。

（三）项目组主要成员：计划书中列出姓名的项目组主要成员由系统自动生成，与申请书原成员保持一致，不可随意调整。如果《批准通知》所附“项目评审意见及修改意见表”中“修改意见”栏目有调整项目组成员相关要求的，待项目开始执行后，按照项目成员变更程序另行办理。

（四）资金预算表：根据批准的项目资助额度，按规定调整项目预算，并按照《国家自然科学基金项目计划书预算表编制说明》填报资金预算表和预算说明书。

（五）正文：

1. 面上项目、地区科学基金项目：如果《批准通知》所附“项目评审意见及修改意见表”中“修改意见”栏目没有修改要求的，只需选择“研究内容和研究目标按照申请书执行”即可；如果《批准通知》中上述栏目明确要求调整研究期限或研究内容等的，须选择“根据研究方案修改意见更改”并填报相关修改内容。

2. 重点项目、重点国际（地区）合作研究项目、重大项目、重大研究计划重点支持项目、重大研究计划集成项目、国家重大科研仪器研制项目、联合基金项目、原创探索计划项目、重大科学基础设施国际合作研究计划专项：须选择“根据研究方案修改意见更改”，根据《批准通知》的要求填写研究（研制）内容，不得自行降低、更改研究目标（或仪器研制的技术性能与主要技术指标、验收技术指标等）或缩减研究（研制）内容。此外，还要突出以下几点：

（1）研究的难点和在实施过程中可能遇到的问题（或仪器研制风险），拟采用的研究（研制）方案和技术路线；

（2）项目主要参与者分工，合作研究单位（如有）之间的关系与分工，重大项目还需说明课题之间的关联；

（3）详细的年度研究（研制）计划。

3. 创新研究群体项目：须选择“根据研究方案修改意见更改”，按下列提纲撰写：

（1）研究方向；

（2）结合国内外研究现状，说明研究工作的学术思想和科学意义（限两个页面）；

（3）研究内容、研究方案及预期目标（限两个页面）；

（4）年度研究计划；

（5）研究队伍的组成情况。

4. 卓越研究群体项目：须选择“根据研究方案修改意见更改”，卓越研究群体项目根据《批准通知》的要求和现场考察专家组的意见和建议，进一步完善并细化研究计划。卓越研究群体项目（延续资助）根据《批准通知》的要求填写计划书即可。

卓越研究群体项目（含延续）按下列提纲撰写：

（1）五年拟开展的研究工作（包括主要研究方向、关键科学问题与研究内容）；

（2）研究方案（包括骨干成员之间的分工及合作方式、学科交叉融合研究计划等）；

（3）年度研究计划；

（4）五年预期目标和可能取得的重大突破等；

（5）研究队伍的组成情况。

5. 数学天元基金项目：天元前沿重点专项项目和数学与其他学科交叉联合资助项目，参照重点项目的方式进行选择和填写；其他类型项目，参照面上项目的方式进行选择和填写。

6. 对于其他类型项目，参照面上项目的方式进行选择和填写。

四、其他事项

（一）根据有关要求，国家自然科学基金资助项目研究形成的代表性论文中发表在我国科技期刊上的应占20%以上。

（二）国家自然科学基金资助项目研究形成的专利申请应按照《建立财政资助科研项目形成专利的声明制度实施方案》要求进行声明。

简表

项目负责人信息	姓　　名	王海燕	性 别	女	出生年月	1987年04月	民 族	汉族
	学　　位	博士			职称	副教授		
	是否在站博士后	否			电子邮件	cshywang@scau.edu.cn		
	电　　话	机密			个人网页			
	工 作 单 位	华南农业大学						
	所 在 院 系 所	数学与信息(软件)学院						
依托单位信息	名　　称	华南农业大学					代码	51064208A0499
	联 系 人	郑雪宜			电子邮件	kycjhk@scau.edu.cn		
	电　　话	020-85280070			网站地址	http://kjc.scau.edu.cn/		
合作单位信息	单 位 名 称							
	广东省智能科学与技术研究院							
项目基本信息	项 目 名 称	面向多源生物医学数据的高阶信息建模研究						
	资 助 类 别	面上项目			亚 类 说 明			
	附 注 说 明							
	申 请 代 码	F0213:生物信息计算与数字健康						
	执 行 年 限	2026.01-2029.12						
	直 接 费 用	50万元						

项目摘要

中文摘要：

多源异构的组学数据，为癌症提供了不同维度多方位的观测信息，并且是肿瘤研究的核心。然而、现有研究面临组学数据噪声高和共表达模式浅显的难题。为此，本项目拟构建以高阶信息建模理论为依托的多源生物医学数据融合分析，解决瓶颈背后的组学数据高阶关系提取难、度量难和高噪声痛点问题。具体地，项目拟建立基于图过滤器的自适应张量子空间学习模型，实现组学高阶关系提取和噪声过滤；建立基于张量二部图的Grassmann Manifold度量学习模型，提升高阶关系度量性能和大规模数据分析能力；建立基于超图网络的跨视图哈希表示学习模型，提升组学高阶关系挖掘和高维数据分析能力。该项目的成功实施对生物信息学分析、多源数据整合分析理论有着重大意义，为其提供理论基础和研究经验。同时为生物医学大数据分析提供直接分析工具，并有望在肿瘤的早期诊断和靶向治疗等领域得到广泛应用，为实现临床精准医疗奠定坚实基础。

Abstract:

Multi-source heterogeneous omics data provide multi-dimensional observation information for cancer and are the core of tumor research. However, existing research faces the problem of high noise in omics data and shallow co-expression patterns. To this end, this project intends to construct a multi-source biomedical data fusion analysis based on high-order information modeling theory to solve the bottleneck problems of high-order relationship extraction, measurement and high noise in omics data. Specifically, the project intends to establish an adaptive tensor subspace learning model based on graph filters to achieve omics high-order relationship extraction and noise filtering; establish a Grassmann Manifold metric learning model based on tensor bipartite graphs to improve high-order relationship measurement performance and large-scale data analysis capabilities; establish a cross-view hash representation learning model based on hypergraph networks to improve omics high-order relationship mining and high-dimensional data analysis capabilities. The successful implementation of this project is of great significance to bioinformatics analysis and multi-source data integration analysis theory, providing a theoretical basis and research experience for them. At the same time, it provides direct analysis tools for biomedical big data analysis, and is expected to be widely used in fields such as early diagnosis and targeted treatment of tumors, laying a solid foundation for the realization of clinical precision medicine.

关键词(用分号分开)：多模态数据融合；多源生物医学数据；模式挖掘；高阶信息建模；辅助诊断

Keywords(用分号分开): Multimodal data fusion; multi-source biomedical data; Pattern mining; Iligh-order Information Modeling; Assistant diagnosis

项目组主要成员

编号	姓名	出生年月	性别	职称	学位	单位名称	电话	证件号码	项目分工	每年工作时间（月）
1	王海燕	1987.04	女	副教授	博士	华南农业大学	机密		项目负责人	10
2	张广煜	1991.05	男	讲师	博士	华南农业大学	机密		算法检验	10
3	张慧玲	1986.10	女	讲师	博士	华南农业大学	机密		算法检验	10
4	张滨	1991.11	男	无	博士	广东省智能科学与技术研究院	机密		算法检验	10

总人数	高级	中级	初级	博士后	博士生	硕士生
10	1	2	0	1	0	6

国家自然科学基金预算制项目预算表

项目批准号：62572201　　项目负责人：王海燕　　金额单位：万元

行次	科目名称	金额
（1）	1、科学基金资助项目直接费用合计	50.0000
（2）	1.1 设备费	3.4000
（3）	其中：设备购置费	3.4000
（4）	1.2 业务费	25.3000
（5）	1.3 劳务费	21.3000
（6）	2. 直接费用中合作研究外拨资金	10.0000
（7）	3. 其他来源资金	0.0000

注：1.请按照项目研究实际需要合理填写各科目预算金额。
2.（1）=（2）+（4）+（5）。
3.如果不存在外拨资金的情形，（6）请填“0”。
4.如果无其他来源资金，（7）请填“0”。

预算说明书

1、科学基金资助项目直接费用

请按照《国家自然科学基金项目计划书预算编制说明》等有关要求，按照政策相符性、目标相关性和经济合理性原则，实事求是编制项目预算。填报时，每个科目应结合科研任务按支出用途进行基本测算说明。

1.1 设备费（是指在项目实施过程中购置或试制专用仪器设备，对现有仪器设备进行升级改造，以及租赁外单位仪器设备而发生的费用。计算类仪器设备和软件工具可在设备费科目列支。填报时，应对设备费支出的必要性和测算的合理性等内容进行说明，并按照设备购置费、试制改造费和租赁使用费的分类提供基本测算依据。单价大于50万元（含50万元）的设备需补充说明设备的主要性能指标、主要技术参数等内容。）

设备费：3.40万元

1.1.1 设备购置费（3.40万元）

本课题将为参与项目的研究生更新2台高性能图形工作站，用于构建组学数据分析模型。共计1.7万元/台×2台=3.40万元。

1.1.2 设备试制费（0.00万元）

实验室无需额外的设备试制费。

1.1.3 设备升级改造与租赁费（0.00万元）

实验室无设备升级改造与租赁费。

1.2 业务费（是指项目实施过程中消耗的各种材料、辅助材料等低值易耗品的采购、运输、装卸、整理等费用，发生的测试化验加工、燃料动力、出版/文献/信息传播/知识产权事务、会议/差旅/国际合作交流等费用，以及其他相关支出。填报时，应按照支出大类进行基本测算说明。）

业务费：25.30万元

1.2.1 差旅/会议/国际合作与交流费（15.70万元）

外埠差旅费：项目成员参加国内学术会议，年均4人次，共计16人次。会议注册费用按2000元/人，差旅往返费用按1800元/人，国内会议行程约3天，会议期间住宿费220元/人/天，会议期间伙食及交通补助180元/人/天，每人次差旅会议费为2000元/人+1800元/人+（220+180）元/人/天×3天=5000元/人。共计5000元/人×16人=8.00万元。

市内交通费：用于本项目在完成过程中小规模研讨会、校内或广州市内其他学校实验室或平台实验、测试等的市内交通费等，共计0.20万元×4年=0.80万元。

国际合作与交流费：项目成员参加国际学术会议，共计4人次。国际会议注册费用4000元/人，签证及往返国际机票费用按1.50万元/人，国际会议行程约5天，会议期间住宿费500元/人/天，会议期间伙食及交通补助300元/人/天，每人次差旅会议费为4000元/人+1.50万元/人+（600+300）元/人/天×5天=2.30万元/人。共计2.30万元/人×3人=6.90万元。

1.2.2 出版/文献/信息传播/知识产权事务费（9.60万元）

论文出版费用：预计发表SCI论文8篇，论文出版费用平均按1.00万元/篇，共计1.00万元/篇×8篇=8.00万元。

学术刊物订阅、专业书籍购买、文献查新等费用：每年预计0.4万元，共计0.40万元×4年=1.60万元。

1.3 劳务费（是指在项目实施过程中支付给参与项目研究的研究生、博士后、访问学者以及项目聘用的研究人员、科研辅助人员等的劳务性费用，以及支付给临时聘请的咨询专家的费用等。填报时，应综合考量劳务费支出对象所承担研究任务的必要性、投入本项目的工作时长、费用标准的合理性等因素，按照人员类别进行基本测算说明。专家咨询费应按照国家有关规定执行。）

劳务费：21.30万元

1.3.1 研究生、博士后劳务费（16.80万元）

参与课题的硕士研究生6人，按500元/月/人支付劳务费用，项目期间每人每年投入的科研时间为10个月，4年共计6人×500元/月/人×（10月/年×4年）=12.00万元。

参与课题的博士后1人，按1200元/月/人支付劳务费用，项目期间每人每年投入的科研时间为10个月，4年共计1人×1200元/月/人×（10月/年×4年）=4.80万元。

1.3.2 专家咨询费（4.50万元）

会议咨询：计划每年通过会议咨询专家并邀请专家做报告5人次。支付给临时聘请的咨询专家的费用按照均2250元/人次，共计2250元/人×5人/年×4年=4.50万元。

2. 直接费用中合作研究外拨资金

需对合作研究单位承担研究任务及协商一致后确定的直接费用资金外拨情况等进行必要说明。如存在多个合作研究单位，需逐一说明。

根据协商一致意见，项目获批直接费用的20%拨付给合作研究单位广东省智能科学与技术研究院；间接经费亦按同等比例（即各20%）同步分配给合作单位。用于对实验数据进行深入分析、优化研究模型等合作研究。

3. 其他来源资金

对其他来源资金的经费来源、资金具体开支用途做简要说明。

无。

报告正文

研究内容和研究目标按照申请书执行。

本项目研究形成的代表性论文中发表在我国科技期刊上的将占 20%以上。

本项目研究形成的专利申请将按照《建立财政资助科研项目形成专利的声明制度实施方案》要求进行声明。

国家自然科学基金项目负责人、依托单位承诺书

国家自然科学基金项目负责人承诺书

本人郑重承诺：我接受国家自然科学基金的资助，严格遵守中共中央办公厅、国务院办公厅《关于进一步加强科研诚信建设的若干意见》《关于进一步弘扬科学家精神加强作风和学风建设的意见》《关于加强科技伦理治理的意见》《科技伦理审查办法（试行）》等规定，和国家自然科学基金委员会关于资助项目管理、项目资金管理等各项规章，在《计划书》填写及项目执行过程中：

（一）按照《批准通知》《国家自然科学基金资助项目计划书填报说明》的要求填写《计划书》，未自行降低、更改目标任务或约定要求，或缩减研究（研制）内容；

（二）树立“红线”意识，严格履行科研合同义务，按照《计划书》负责实施本项目（批准号：62572201），切实保证研究工作时间，按时报送有关材料，及时报告重大情况变动，不违规将科研任务转包、分包他人，不以项目实施周期外或不相关成果充抵交差；

（三）遵守科研诚信、科技伦理规范和学术道德，认真开展研究工作，对资助项目发表的论著和取得的研究成果按规定进行标注，不在非本项目资助的成果或其他无关成果上标注本项目批准号，反对无实质学术贡献者“挂名”，不在成果署名、知识产权归属等方面侵占他人合法权益，并如实报告本人及项目组成员发生的违背科研诚信要求的任何行为；

（四）尊重科研规律，弘扬科学家精神，严谨求实，追求卓越，反对浮夸浮躁、投机取巧，不人为夸大学术或技术价值，不传播未经科学验证的现象和观点；

（五）将项目资金全部用于与本项目研究工作相关的支出，并结合科研活动需要，科学合理安排项目资金支出进度；

（六）做好项目组成员的教育和管理，确保遵守以上相关要求。

如违背上述承诺，本人愿接受国家自然科学基金委员会和相关部门做出的各项处理决定。

项目负责人（签字）：王渔燕

2025 年 9 月 16 日

依托单位科研管理部门：	依托单位财务管理部门：彭莹珊
负责人（签章）：	负责人（签章）：
2025 年 9 月 23 日	2025 年 9 月 23 日

国家自然科学基金项目依托单位承诺书

我单位同意承担上述国家自然科学基金项目，将保证项目负责人及其研究队伍的稳定和研究项目实施所需的条件，严格遵守中共中央办公厅、国务院办公厅《关于进一步加强科研诚信建设的若干意见》《关于进一步弘扬科学家精神加强作风和学风建设的意见》《关于加强科技伦理治理的意见》《科技伦理审查办法（试行）》等规定，和国家自然科学基金委员会有关资助项目管理、项目资金管理、科研诚信管理和科技伦理管理等各项规定，并督促实施。

依托单位（公章）

2025 年 9 月 23 日

国家自然科学基金资助项目签批审核表

本栏目由自然科学基金委填写

科学处审查意见：

同意按照计划执行

负责人（签章）：吴国政

2025-10-28日

科学部审查意见：

同意

负责人（签章）：

年 月 日

2025-10-28

SCAULIB202625903

检索证明

根据委托人提供的论文材料，委托人**华南农业大学数学与信息学院 张广煜(学科类型:自然科学)** 5 篇论文收录情况如下表。

序号	论文名称	发表刊物及发表的年月卷期/页码等	作者排名	论文等级	作者文中单位	收录情况	影响因子	中科院大类分区
1	Facilitated low-rank multi-view subspace clustering	KNOWLEDGE-BASED SYSTEMS 出版年：2023 出版日期：JAN 25 卷期：260 页码：- 文献号：110141 文献类型：Article	第一作者	T2 类	华南农业大学数学与信息学院	SCI	IP2-year=7.2 IP5-year=7.4 （2023）	计算机科学 1 区 Top 期刊：是 OA 期刊：否 （2023）
2	Scalable tri-factorization guided multi-view subspace clustering	KNOWLEDGE-BASED SYSTEMS 出版年：2025 出版日期：MAR 15 卷期：312 页码：- 文献号：113119 文献类型：Article	第一作者	T2 类	华南农业大学数学与信息学院	SCI	IP2-year=7.6 IP5-year=7.6 （2024）	计算机科学 1 区 Top 期刊：是 OA 期刊：否 （2025）
3	Unified and Tensorized Incomplete Multi-View Kernel Subspace Clustering	IEEE TRANSACTIONS ON EMERGING TOPICS IN COMPUTATIONAL INTELLIGENCE 出版年：2024 出版日期：APR	第一作者	A 类	华南农业大学数学与信息学院	SCI	IP2-year=6.5 IP5-year=6.6 （2024）	计算机科学 2 区 Top 期刊：否 OA 期刊：否 （2025）

		卷期：8 2 页码：1550-1566 文献类型：Article						
4	Tensorized Incomplete Multi-view Kernel Subspace Clustering	NEURAL NETWORKS 出版年：2024 出版日期：NOV 卷期：179 页码：- 文献号：106529 文献类型：Article	第一作者	A 类	华南农业大学 数学与信息学院	SCI	IP2-year=6.3 IP5-year=7.5 （2024）	计算机科学 2 区 Top 期刊：是 OA 期刊：否 （2025）
5	Large-Scale Tensorized Multi-View Kernel Subspace Clustering	ACM TRANSACTIONS ON INTELLIGENT SYSTEMS AND TECHNOLOGY 出版年：2025 出版日期：AUG 卷期：16 4 页码：- 文献号：85 文献类型：Review	第一作者	B 类	华南农业大学 数学与信息学院	SCI	IP2-year=6.6 IP5-year=8.4 （2024）	计算机科学 3 区 Top 期刊：否 OA 期刊：否 （2025）

说明：论文等级和中科院大类分区按《华南农业大学学术论文评价方案（试行）》划分。

报告免责声明:如未盖章,报告无效

检索员：王银林

华南农业大学图书馆

2026-03-11

SCAULIB202625686

检索证明

根据委托人提供的论文材料，委托人华南农业大学数学与信息学院 张广煜(学科类型:自然科学) 1 篇论文收录情况如下表。

序号	论文名称	发表刊物及发表的年月卷期/页码等	作者排名	论文等级	作者文中单位	收录情况	影响因子	中科院大类分区
1	Dual-level Facilitated Multi-view Contrastive Graph Clustering	The 31st IEEE International Conference on Parallel and Distributed Systems (ICPADS 2025) 出版年：2025 卷期： 页码：1 - 文献号： 文献类型：会议论文	第一作者	无	华南农业大学数学与信息学院	在线发表	无	无

说明：论文等级和中科院大类分区按《华南农业大学学术论文评价方案（试行）》划分。

报告免责声明:如未盖章,报告无效

检索员：田成

华南农业大学图书馆

2026-03-10

信息检索专用章

SCAULIB202625514

检索证明

根据委托人提供的论文材料，委托人华南农业大学数学与信息学院 张广煜(学科类型:自然科学) 1 篇论文收录情况如下表。

序号	论文名称	发表刊物及发表的年月卷期/页码等	作者排名	论文等级	作者文中单位	收录情况	影响因子	中科院大类分区
1	Multi-view Clustering via Flexible Dual-level Fusion	2025 IEEE 31th International Conference on Parallel and Distributed Systems (ICPADS) 出版年：2025 出版日期：14-18 December 卷期： 页码：- 文献类型：会议论文	第一作者	无	华南农业大学数学与信息学院	在线发表	无	无

说明：论文等级和中科院大类分区按《华南农业大学学术论文评价方案（试行）》划分。

报告免责声明:如未盖章,报告无效

检索员 田成

华南农业大学图书馆

2026-03-09

信息检索专用章

SCAULIB202625756

检索证明

根据委托人提供的论文材料，委托人华南农业大学数学与信息学院 张广煜(学科类型:自然科学) 1 篇论文收录情况如下表。

序号	论文名称	发表刊物及发表的年月卷期/页码等	作者排名	论文等级	作者文中单位	收录情况	影响因子	中科院大类分区
1	Confidence-oriented Contrastive Graph Clustering	2024 International Joint Conference on Neural Networks (IJCNN) 出版年：2024 卷期： 页码：- 文献号： 文献类型：会议论文	通讯作者	无	华南农业大学数学与信息学院	IEEE	无	无

说明：论文等级和中科院大类分区按《华南农业大学学术论文评价方案（试行）》划分。

报告免责声明:如未盖章,报告无效

检索员：张进

华南农业大学图书馆

2026-03-11

SCAULIB202523751

检索证明

根据委托人提供的论文材料，委托人**华南农业大学数学与信息学院 张广煜(学科类型:自然科学)** 1 篇论文收录情况如下表。

序号	论文名称	发表刊物及发表的年月卷期/页码等	作者排名	论文等级	作者文中单位	收录情况	影响因子	中科院大类分区
1	Multi-scale Multi-order Attributed Graph	International Conference on Intelligent Computing 2025 出版年：2025 出版日期：26 July 卷期： 页码：482-494 文献号： 文献类型：会议论文	通讯作者	无	College of Mathematics and Informatics, South China Agricultural University	在线发表	无	无

说明：论文等级和中科院大类分区按《华南农业大学学术论文评价方案（试行）》划分。

报告免责声明:如未盖章,报告无效

检索员：张进

华南农业大学图书馆

2025-11-05

Knowledge-Based Systems 260 (2023) 110141

Contents lists available at ScienceDirect

Knowledge-Based Systems

journal homepage: www.elsevier.com/locate/knosys

Facilitated low-rank multi-view subspace clustering

Guang-Yu Zhang [a], Dong Huang [a,*], Chang-Dong Wang [b]

[a] College of Mathematics and Informatics, South China Agricultural University, Guangzhou, China
[b] School of Computer Science and Engineering, Sun Yat-sen University, Guangzhou, China

ARTICLE INFO

Article history:
Received 7 July 2022
Received in revised form 14 November 2022
Accepted 16 November 2022
Available online 29 November 2022

Keywords:
Subspace clustering
Multi-view clustering
Low-rank representation
Facilitated strategy
Accelerated algorithm

ABSTRACT

Low-rank multi-view subspace clustering has recently attracted increasing attention in the multi-view learning research. Despite significant progress, most existing approaches still suffer from two issues. First, they mostly focus on exploiting the low-rank consistency across multiple views, but often ignore the low-rank structure within each view. Second, they often encounter the expensive time overhead, typically owing to the costs of matrix inversion and singular value decomposition (SVD) at each iteration. In light of this, we propose an efficient and effective approach termed Facilitated Low-rank Multi-view Subspace Clustering (FLMSC) in this paper. In terms of efficiency, our approach factorizes the view-specific representation matrix into two small factor matrices, i.e., an orthogonal dictionary and a latent representation, which mitigates the computation cost of solving SVD problems. In terms of effectiveness, our approach preserves the latent low-rank structure within each view, while at the same time encourages the structural consistency across different views by imposing an agreement term. Based on this, our approach can fully explore the underlying subspace structure of multiple views, so as to better serve for the following spectral clustering. Besides, a facilitated iterative algorithm is developed for the resultant optimization problem, upon which the matrix inversion operation is substantially accelerated during iterations. Experimental results on a variety of multi-view data sets demonstrate the effectiveness and efficiency of our approach. The source code and data sets are available at: https://www.researchgate.net/publication/365349930_FLMSC_v1.

1. Introduction

With the rapid development of sensor acquisition technology, multi-view data have been more frequently employed in different scientific studies. Take face recognition as an example, a human face can be observed at multiple light conditions. In web page analysis, the same web page can be represented by three types of features, e.g., text context, described pictures and the hyperlinks pointing to it. In document analysis, a news report may be available in multiple languages, and each language can be considered as a particular view. As the collection of multi-view data comes from diverse sources, how to efficiently utilize the rich and heterogeneous information from different views has become an interesting yet challenging problem. In light of this, some researchers raise great interest in the paradigm of multi-view learning [1–3]. As a hot research topic, multi-view learning has attracted considerable interest in different research tasks, such as multi-view deep learning [4,5], multi-view feature selection [6,7], multi-view dimensionality reduction [8,9], multi-view outlier detection [10,11] and multi-view clustering [12,13]. In this paper, we concentrate on the problem of multi-view clustering, whose goal is to enhance the performance by leveraging the complementary information across various views.

The multi-view clustering problem has been widely-studied from different perspectives. In the early stage, many studies focus on the k-means based approaches [14,15] and the matrix factorization based approaches [16,17]. The k-means based approaches, tend to learn the consistent cluster label across multiple views, while the matrix factorization based approaches, target at learning latent low-dimensional representation with specifically designed regularizers. Apart from these approaches, the kernel-based approaches have also drawn considerable attention, whose goal is to capture the nonlinear structure hidden in multi-view data via some kernel tricks. The representative works in the kernel-based approaches include [18,19]. In addition, benefiting from the theoretical guarantees on graph spectral theory, a vast number of spectral-based approaches have been also designed in multi-view clustering, such as [20,21]. In the spectral-based approaches, the main goal is to discover the latent relationship among data samples, so as to build a robust affinity matrix for the final spectral clustering.

Beyond the aforementioned approaches, the subspace-based approaches have been shown to be a powerful tool in computer

* Corresponding author.
E-mail addresses: guangyuzhg@foxmail.com (G.-Y. Zhang), huangdonghere@gmail.com (D. Huang), changdongwang@hotmail.com (C.-D. Wang).

https://doi.org/10.1016/j.knosys.2022.110141
0950-7051/

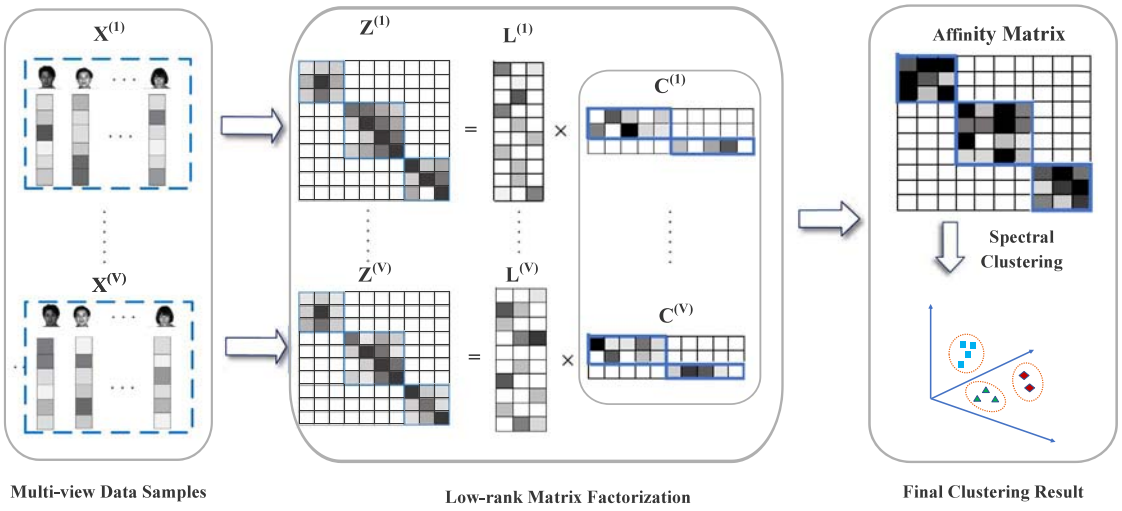


Fig. 1. The framework of our approach.

vision and data mining. The key idea of subspace-based approaches is to uncover intrinsic subspace structure across multiple views. Starting from different perspectives, a family of subspace-based approaches [22–25] have been developed and implemented over the last few years. Among these approaches, the low-rank subspace-based approaches [26,27] have become a promising topic with wide applicability and elegant theoretical guarantees. The typical works in low-rank subspace-based approaches include Latent Multi-view Subspace Clustering (LMSC) [28], Multimodal Sparse and Low-rank Subspace Clustering (MLRSSC) [29], Flexible Multi-view Representation Learning for Subspace Clustering (FMR) [30] and Dual Shared-Specific Multi-view Subspace Clustering (DSS-MSC) [31]. Based on the well-defined framework of low-rank representation, these recently developed approaches have reached competitive clustering performance in empirical study. However, it is observed that they still have two common drawbacks. On the one hand, the above approaches only concentrate on the low-rank structure across different views, but often neglect the latent low-rank structure within each view. On the other hand, the existing approaches still suffer from high computational costs at each iteration, which severely limit their applicability to solve large-scale or even medium-scale problems.

To address the aforementioned drawbacks, this paper develops a Facilitated Low-rank Multi-view Subspace Clustering (FLMSC) approach (as shown in Fig. 1), which is featured by its low complexity and competitive performance. In terms of the efficiency, the proposed approach factorizes the view-specific representation into an orthogonal dictionary and a latent representation. By this means, the original singular value decomposition (SVD) problem can be converted into two smaller-scale optimization problems that efficiently reduces the computational cost. In terms of effectiveness, our approach preserves the low-rank structure within individual view, and simultaneously discovers the structural consistency of multiple views in a unified framework. This is beneficial to capture the underlying subspace across different views, leading to the promising clustering performances. The main contributions and novelty of this work are given below.

- This paper presents an efficient and effective low-rank based multi-view subspace clustering approach termed FLMSC.
- This paper develops a facilitated optimization algorithm for low-rank multi-view subspace clustering.
- Comprehensive experiments are conducted on six benchmark data sets, which have shown the advantage of our approach in both efficiency and effectiveness.

The rest of this paper is organized as follows. Section 2 briefly reviews the relevant literature on multi-view clustering as well as multi-view subspace clustering. After that, Section 3 introduces our approach FLMSC and its optimization scheme. In Section 4, a variety of experiments are conducted on practical multi-view data sets to demonstrate the superior of FLMSC approach. Finally, our conclusion is summarized in Section 5.

2. Related work

This section first reviews the representative studies on multi-view clustering, and then introduces the recent development of multi-view subspace clustering.

2.1. Multi-view clustering

Multi-view clustering plays a critical role in various scientific problems. Over the past decade, substantial efforts have been made towards the development of multi-view clustering. Naturally, owing to the limitation of knowledge and technology, the earliest studies were proposed to handle the two-view cases. In [32], Bickel et al. presented the first attempt to cope with two-view cases, which aims at clustering document data by considering them as independent subsets. Besides, De et al. [33] proposed a spectral-based approach for two-view cases. The basic idea is to extract the relationship of web page data with bipartition graph. Apart from these two pioneering approaches, the subsequent studies in multi-view clustering focused on investigating the cases with arbitrary number of views. Along this line, Kumar et al. [34] proposed a novel multi-view spectral clustering approach under the co-training framework. By utilizing the complementary information from diverse views, this approach can iteratively learn the robust spectral embedding within individual view. In [35], Wang et al.presented a graph-based multi-view clustering approach based on two-step graph fusion strategy. Furthermore, Zhu et al. [36] introduced a multi-view spectral clustering approach via one-step strategy, which seamlessly leans the view-graph matrices and one unified affinity matrix in a mutual reinforcement manner. Based on the multiple graph fusion strategy, Zhan et al. [37] derived a graph fusion structure (GSF) to address the multi-view clustering problem. In GSF, both the graph

structure within individual view as well as the consistency across heterogeneous views are considered in the graph fusion process. Based on the bipartite graph strategy, Li et al. [38] designed a novel multi-view spectral clustering approach for large-scale applications. In [17], Liu et al. introduced the Joint Nonnegative Matrix Factorization (NMF) framework to discover the common clustering factor from heterogeneous views. Based on this work, Zhao et al. [39] developed a more effective clustering approach by means of Deep Nonnegative Matrix Factorization (DNMF) [40] solution. The specific approach seeks the deep structure in multi-view data layer by layer, which in turn facilitates clustering performance. Inspired by this, Huang et al. [41] designed an auto-weighted multi-view clustering approach, whose general idea relies to explore the hierarchical information hidden in multi-view structure. To efficiently tackle large-scale data problems, Cai et al. [15] presented a clustering approach that intends to capture the consistent cluster structure by matrix factorization. Particularly, this approach introduces a simple objective function with $l_{2,1}$ constraint, which is able to eliminate the negative influence of the outliers and noise in the data. Besides, Chen et al. [14] developed a classical multi-view k-means clustering approach, which can automatically compute the weights of views and features w.r.t. their importance. Following this approach, Zhang et al. [42] proposed another multi-view k-means clustering approach named Two-level Weighted Collaborative k-means (TW-Co-k-means). It learns the final cluster label by considering the mutual relationships between multiple views. Moreover, Jiang et al. [43] developed a groundbreaking work in multi-view fuzzy k-means clustering. The approach learns the fuzzy cluster assignment matrix in a collaborative manner. Inspired by the success of [12], Huang et al. [44] developed a multi-view clustering approach termed Multi-view intact space clustering (MVIC). This approach first constructs the intact space of all views, then captures the common clustering structure based on matrix factorization. The aforementioned approaches address the multi-view clustering problem with respect to different strategies. To acquire more detailed knowledge about multi-view clustering, please refer to the survey in [45] [46].

2.2. Multi-view subspace clustering

During the last few years, there emergences a trend towards the study of multi-view subspace clustering. As an early attempt, Gao et al. [47] extended the classical subspace clustering to multi-view case, which not only performs subspace clustering on each specific view, but also guarantees the indicator consistency across different views. Along this work, Cao et al. [48] further developed a novel subspace-based approach for multi-view clustering. The approach uses HSIC to exploit the complementarity across diverse views. To handle the nonlinearities hidden in multi-view data, Zhang et al. [24] designed a kernelized approach for multi-view subspace clustering. This approach can cluster the high-dimensional data effectively via the one-step framework. Recently, Chen et al. [49] presented a novel approach to partition the multi-view features in the shared embedding space. Specially, this approach unifies latent embedding space learning and the structured subspace clustering into one framework. Meanwhile, Kang et al. [50] proposed an efficient subspace-based approach that conducts large-scale multi-view clustering in linear time. Following this work, Sun et al. [51] developed a large-scale subspace-based approach via unified anchors selection strategy. In this approach, two important tasks, i.e., the subspace clustering and unified anchors selection are integrated into a joint framework. Recently, Chen et al. [52] proposed a large-scale subspace-based approach with efficient orthogonal constraint. Apart from these approaches, low-rank subspace-based approaches have been attracted more and more attention because of the well-defined framework. For instance, Abavisani et al. [29] presented a classical low-rank subspace-based approach, whose idea relies to learn the shared subspace structure across different views by low-rank and sparse constraints. Based on the assumption that all views are generated from the same latent low-dimensional space, Zhang et al. [28] designed an advanced low-rank subspace-based approach named LMSC. Different from the most existing approaches, LMSC jointly learns latent representation from multiple sources, meanwhile conducts low-rank subspace clustering in the learned latent representation. Starting from both consistent and individual perspectives, Luo et al. [23] developed an effective low-rank approach for multi-view subspace clustering. In this approach, the structural consistency across all views is modeled by low-rank constraint. Borrowing the idea from [53], Pan et al. [54] recently proposed a simple but promising low-rank subspace-based approach. This approach utilizes the low-rank constraint to characterize the consistent subspace structure of multiple views. Very recently, Cai et al. [55] developed an effective approach to seek both commonness and inconsistencies across different views. The aforementioned low-rank subspace-based approaches have achieved prominent performances in applications, however, they still suffer from two common drawbacks. First, many of these approaches only consider the structural consistency across multiple views, neglecting the latent low-rank structure within individual view. Second, the existing approaches require to solve many SVD and matrix inversion problems during optimization, which inhibits their ability to solve larger-scale problems in practice.

3. Proposed approach

3.1. Basic notations

In the following, uppercase letters are used to define the matrices, such as M. Besides, $\mathrm{m}_{\cdot j}$, $\mathrm{m}_{i\cdot}$ and m_{ij} to denote as the jth column, ith row, and ijth element of the matrix M, respectively. Tr(M) is defined as the trace definition of the matrix M. M^T and M^{-1} represent the transpose as well as inverse of the matrix M, respectively. The identity matrix is represented by the letter I.

3.2. Preliminary knowledge

Before this subsection illustrates the proposed target function, a brief review is provided for the general framework of single-view subspace clustering, as well as a representative approach, namely Low-Rank Representation (LRR).

Let $\mathrm{X} = \{\mathrm{x}_{:1}, \ldots, \mathrm{x}_{:n}\} \in \mathbb{R}^{d\times n}$ be a collection of multi-view data samples, where d is the feature dimension and n is the number of data samples. For traditional subspace clustering, the task is to investigate the linear relationship among data samples with self-expression property. To be more specific, the self-expression property means that input data samples embed in a union of latent subspaces, where each sample can be written by a linear combination of other samples from the same subspace. Mathematically, we can express this as follows:

$$\mathrm{x}_{:i} = \sum_{j\neq i} \mathrm{x}_{:j}\mathrm{z}_{ji}, \ \forall i, \tag{1}$$

If we further consider the noise and outlier in data, the self-expression property can be written in the following form:

$$\mathrm{X} = \mathrm{XZ} + \mathrm{E}, \tag{2}$$

where $\mathrm{Z} \in \mathbb{R}^{n\times n}$ denotes the subspace representation and $\mathrm{E} \in \mathbb{R}^{d\times n}$ denotes the error matrix. Based on this self-expressiveness

property, we can formulate the optimization problem of general subspace clustering as follows:

$$\min_{\mathrm{Z,E}} \|\mathrm{E}\|_p + \alpha\|\mathrm{Z}\|_q \quad \text{s.t. } \mathrm{X} = \mathrm{XZ} + \mathrm{E}, \tag{3}$$

in which $\|\cdot\|_p$ and $\|\cdot\|_q$ are two matrix norm constraints, and $\alpha > 0$ is a balanced parameter. Note that the obtained subspace representation Z cannot describe the symmetrical relationship among data samples, since element Z_{ij} may be not equal to element Z_{ji}. Hence, we employ a necessary step to compute the affinity matrix W, which is shown as follows:

$$\mathrm{W} = (|\mathrm{Z}| + |\mathrm{Z}^T|)/2. \tag{4}$$

After computing the affinity matrix W, we could obtain the final clustering results by conducting spectral clustering algorithm.

3.2.1. Low-rank representation (LRR)

LRR [56] is one of the most popular approaches in subspace clustering. In LRR, the key is to recover the low-dimensional subspace structure from original data matrix with low-rank constraint. Concretely, the approach attempts to find a low-rank subspace representation that reflects the latent correlation among data samples by a given dictionary. In practice, the data matrix itself is always selected as the given dictionary. Therefore, given a collection of data samples $\mathrm{X} = \{\mathrm{x}_{:1}, \ldots, \mathrm{x}_{:n}\}$, the target function w.r.t. LRR could be written as follows:

$$\min_{\mathrm{Z,E}} \|\mathrm{E}\|_F^2 + \alpha\|\mathrm{Z}\|_* \quad \text{s.t. } \mathrm{X} = \mathrm{XZ} + \mathrm{E}, \tag{5}$$

in which $\|\mathrm{E}\|_F^2 = \sum_{i,j=1}^{n} \mathrm{e}_{ij}{}^2$ is the F-norm on error matrix E, as well as α is a positive balanced parameter. The aforementioned target function can be rewritten in the form as:

$$\min_{\mathrm{Z}} \|\mathrm{X} - \mathrm{XZ}\|_F^2 + \alpha\|\mathrm{Z}\|_* \tag{6}$$

where $\|\mathrm{Z}\|_*$ denotes the nuclear norm on subspace representation matrix Z. Specifically, the LRR approach has been successfully applied in image segmentation, face recognition and saliency detection.

3.3. The proposed target function

This subsection introduces the proposed target function in a gradual manner. Let us consider a collection of multi-view data samples $\mathrm{X} = [\mathrm{X}^{(1)^T}, \ldots, \mathrm{X}^{(V)^T}]^T$, in which $\mathrm{X}^{(v)} = [\mathrm{x}_{:1}^{(v)}, \ldots, \mathrm{x}_{:n}^{(v)}] \in \mathbb{R}^{d_v \times n}$ represents the data matrix in the vth view. If we only consider the consistency of all views, then the subspace representations across different views should share the common patterns. If we only consider the low-rank structure individually within individual view, then the self-expression formulation for multi-view data can be written by:

$$\mathrm{X}^{(v)} = \mathrm{X}^{(v)}\mathrm{Z}^{(v)} + \mathrm{E}^{(v)}, \forall v, \tag{7}$$

in which $\mathrm{Z}^{(v)}$ denotes the vth view subspace representation, while at the same time $\mathrm{E}^{(v)}$ denotes the vth view error matrix. Based on the above equation, we can further extend LRR to multi-view setting, which comes to the following optimization problem:

$$\min_{\{\mathrm{Z}^{(v)}\}_{v=1}^{V},\{\mathrm{E}^{(v)}\}_{v=1}^{V}} \sum_{v=1}^{V} \|\mathrm{E}^{(v)}\|_F^2 + \lambda \sum_{v=1}^{V} \|\mathrm{Z}^{(v)}\|_* \quad \text{s.t. } \mathrm{X}^{(v)} = \mathrm{X}^{(v)}\mathrm{Z}^{(v)} + \mathrm{E}^{(v)}, \forall v, \tag{8}$$

and achieve the following equivalent form:

$$\min_{\{\mathrm{Z}^{(v)}\}_{v=1}^{V}} \sum_{v=1}^{V} \|\mathrm{X}^{(v)} - \mathrm{X}^{(v)}\mathrm{Z}^{(v)}\|_F^2 + \lambda \sum_{v=1}^{V} \|\mathrm{Z}^{(v)}\|_* \tag{9}$$

It is observed that the above formulation reveals the desired low-rank structure within individual view. However, this multi-view LRR formulation suffers from very high per-iteration cost, i.e. the cubic of the data samples. In order to improve its scalability when solving larger-scale practical problems, we further incorporate the idea of low-rank matrix factorization. Consequently, we have the following lemma:

Lemma 1. *For any matrix* $\mathrm{X} \in \mathbb{R}^{m\times n}$ *with rank*$(\mathrm{X}) = r \le d$*, we can factorize it into two smaller matrices* $\mathrm{U} \in \mathbb{R}^{m\times d}$ *and* $\mathrm{A} \in \mathbb{R}^{d\times n}$ *such that* $\mathrm{X} = \mathrm{UA}$*. If matrix* U *has orthogonal columns, i.e.,* $\mathrm{U}^T\mathrm{U} = \mathrm{I}$*, then we have the following property:*

$$\|\mathrm{X}\|_* = \|\mathrm{A}\|_* \tag{10}$$

where $\|\cdot\|_*$ *denotes the nuclear matrix.*

Proof. Since $\mathrm{X} = \mathrm{UA}$ and $\mathrm{U}^T\mathrm{U} = \mathrm{I}$, then we have $\|\mathrm{X}\|_* = \mathrm{Tr}(\sqrt{\mathrm{X}^T\mathrm{X}}) = \mathrm{Tr}(\sqrt{\mathrm{A}^T\mathrm{U}^T\mathrm{UA}}) = \mathrm{Tr}(\sqrt{\mathrm{A}^T\mathrm{A}}) = \|\mathrm{A}\|_*$ □

According to this lemma, we can naturally factorize the view-specific representation into two small matrices, like:

$$\mathrm{Z}^{(v)} = \mathrm{L}^{(v)}\mathrm{C}^{(v)}, \forall v \tag{11}$$

where $\mathrm{L}^{(v)} \in \mathbb{R}^{n\times\gamma}$ represents the orthogonal dictionary in the vth view ($\gamma \ll n$), which has orthogonal columns. Besides, $\mathrm{C}^{(v)} \in \mathbb{R}^{\gamma\times n}$ represents the latent representation in the vth view, which is able to efficiently encode the low-rank property from matrix $\mathrm{Z}^{(v)}$. By considering the formulations in Eqs. (11) and (9), the new multi-view LRR formulation can be expressed as follows:

$$\min_{\{\mathrm{Z}^{(v)},\mathrm{L}^{(v)},\mathrm{C}^{(v)}\}_{v=1}^{V}} \sum_{v=1}^{V} \|\mathrm{X}^{(v)} - \mathrm{X}^{(v)}\mathrm{Z}^{(v)}\|_F^2 + \lambda \sum_{v=1}^{V} \|\mathrm{C}^{(v)}\|_* \quad \text{s.t. } \mathrm{Z}^{(v)} = \mathrm{L}^{(v)}\mathrm{C}^{(v)}, \mathrm{L}^{(v)^T}\mathrm{L}^{(v)} = \mathrm{I}, \forall v \tag{12}$$

Similar to the formulation in Eq. (9), the aforementioned formulation still fails to discover the consistent structure across different views, which may further improve the clustering performance. Before introducing the solution to solve this shortcoming, we obtain the following lemma:

Lemma 2. *Let* $\mathrm{X} \in \mathbb{R}^{m\times n}$ *and* $\mathrm{Y} \in \mathbb{R}^{m\times n}$ *be two matrices with the same dimensions, we can factorize them into much smaller matrices respectively, i.e.,* $\mathrm{X} = \mathrm{UA}$ *and* $\mathrm{Y} = \mathrm{WB}$*, where* $\mathrm{U}, \mathrm{W} \in \mathbb{R}^{m\times d}$ *and* $\mathrm{A}, \mathrm{B} \in \mathbb{R}^{d\times n}$*. If matrices* U, W *have orthogonal columns and* $\|\mathrm{A}\|_* = \|\mathrm{B}\|_*$ *holds, then we have the following property:*

$$\|\mathrm{X}\|_* = \|\mathrm{Y}\|_* \tag{13}$$

This lemma reveals that the structural consistency across different views can be measured by latent representations in diverse views. Inspired by this, we naturally introduce the following agreement term:

$$\min_{\{\mathrm{C}^{(v)}\}_{v=1}^{V}} \frac{1}{V-1} \sum_{\substack{v=1,\\ v\neq w}} \|\mathrm{C}^{(v)} - \mathrm{C}^{(w)}\|_F^2 \tag{14}$$

Derived from the idea of pairwise similarity, the above agreement term measures the average consistence of different latent representations effectively. By further incorporating this agree-

ment term into Eq. (12), we can fulfill the formulation of FLMSC as follows:

$$\begin{aligned}\min_{\{Z^{(v)},L^{(v)},C^{(v)}\}_{v=1}^{V}} & \sum_{v=1}^{V}\|X^{(v)}-X^{(v)}Z^{(v)}\|_F^2+\lambda_1\sum_{v=1}^{V}\|C^{(v)}\|_* \\ & +\frac{\lambda_2}{V-1}\sum_{\substack{v=1,\\ v\neq w}}^{V}\|C^{(v)}-C^{(w)}\|_F^2 \\ & \text{s.t. } Z^{(v)}=L^{(v)}C^{(v)},\ L^{(v)^T}L^{(v)}=I,\forall v\end{aligned} \tag{15}$$

Here λ_1 and λ_2 are two positive balancing parameters. As can be seen, the proposed formulation preserves the low-rank structure within each view, and simultaneously discovers the structural consistency across different views jointly. By this means, both the complementary principle and the consistent principle can be obeyed, which facilitates the performance of multi-view clustering in an effective manner. For clarity, Fig. 1 gives the flowchart of the proposed approach.

3.4. Optimization

Based on the framework of ADMM [57], we propose an efficient optimization algorithm to solve the minimization problem in Eq. (15) efficiently. First, the corresponding augmented Lagrangian function is formulated as follows:

$$\begin{aligned}&\mathcal{L}_\mu(\{Z^{(v)}\}_{v=1}^V,\{L^{(v)}\}_{v=1}^V,\{C^{(v)}\}_{v=1}^V) \\ &=\sum_{v=1}^{V}\|X^{(v)}-X^{(v)}Z^{(v)}\|_F^2+\lambda_1\sum_{v=1}^{V}\|C^{(v)}\|_* \\ &\quad+\frac{\lambda_2}{V-1}\sum_{\substack{v=1,\\ v\neq w}}^{V}\|C^{(v)}-C^{(w)}\|_F^2 \\ &\quad+\sum_{v=1}^{V}<Y^{(v)},Z^{(v)}-L^{(v)}C^{(v)}>+\frac{\mu}{2}\sum_{v=1}^{V}\|Z^{(v)}-L^{(v)}C^{(v)}\|_F^2, \\ &\quad\text{s.t. } L^{(v)^T}L^{(v)}=I,\forall v,\end{aligned} \tag{16}$$

where $\{Y^{(v)}\}_{v=1}^V$ represent the Lagrange multipliers and $\mu>0$ represents the penalty parameter. Apparently, it is not easy to optimize all the variables in Eq. (16) at the same time. Therefore, we adopt an iterative optimization scheme to update the variables one by one. The corresponding procedure of updating steps as shown in what follows.

3.4.1. Update rule for the variables $\{Z^{(v)}\}_{v=1}^V$

When fixing the other variables, we can solve the following minimization subproblem w.r.t. variable $Z^{(v)}$:

$$\min_{Z^{(v)}}\ \|X^{(v)}-X^{(v)}Z^{(v)}\|_F^2+<Y^{(v)},Z^{(v)}-L^{(v)}C^{(v)}>+\frac{\mu}{2}\|Z^{(v)}-L^{(v)}C^{(v)}\|_F^2 \tag{17}$$

By setting the derivative of above subproblem to zero, we can get the following optimal solution:

$$Z^{(v)}=(X^{(v)^T}X^{(v)}+\mu I)^{-1}(X^{(v)^T}X^{(v)}-Y^{(v)}+\mu L^{(v)}C^{(v)}). \tag{18}$$

The above solution requires to compute the matrix inversion of $(X^{(v)^T}X^{(v)}+\mu I)$ during optimization. This approach may suffer from expensive computational costs in handling large-scale multi-view data sets. To remedy this issue, a feasible way is to reduce the size of matrix inversion. First of all, a useful lemma is given as follows:

Lemma 3. *For any three matrices $E\in\mathbb{R}^{n\times n}$, $F\in\mathbb{R}^{n\times d}$ and $G\in\mathbb{R}^{n\times d}$, then the following equation holds:*

$$(E+FG^T)^{-1}=E^{-1}-E^{-1}F(I+G^TE^{-1}F)^{-1}G^TE^{-1}. \tag{19}$$

The above lemma is the so-called Sherman–Morrison–Woodbury equation.

Proof.

$$\begin{aligned}&(E^{-1}-E^{-1}F(I+G^TE^{-1}F)^{-1}G^TE^{-1})(E+FG^T)\\ &=I+E^{-1}FG^T-E^{-1}F(I+G^TE^{-1}F)^{-1}G^T\\ &\quad-E^{-1}F(I+G^TE^{-1}F)^{-1}G^TE^{-1}FG^T\\ &=I+E^{-1}FG^T-E^{-1}F(I+G^TE^{-1}F)^{-1}(I+G^TE^{-1}F)G^T\\ &=I\end{aligned}$$ □

Based on this, if the condition $n>d_v$ is satisfied, then the new solution for variable $Z^{(v)}$ can be shown by:

$$Z^{(v)}=(\frac{1}{\mu}I-\frac{1}{\mu}X^{(v)^T}(I+\frac{1}{\mu}X^{(v)}X^{(v)^T})^{-1}X^{(v)})(X^{(v)^T}X^{(v)}-Y^{(v)}+\mu L^{(v)}C^{(v)}). \tag{20}$$

Notice that we only need to pre-compute $X^{(v)}X^{(v)^T}$ once before the whole algorithm begins.

3.4.2. Update rule for the variables $\{L^{(v)}\}_{v=1}^V$

When fixing the other variables, we can solve the following minimization subproblem w.r.t. variable $L^{(v)}$:

$$\begin{aligned}\min_{L^{(v)}}\ & <Y^{(v)},Z^{(v)}-L^{(v)}C^{(v)}>+\frac{\mu}{2}\|Z^{(v)}-L^{(v)}C^{(v)}\|_F^2 \\ & \text{s.t. } L^{(v)^T}L^{(v)}=I\end{aligned} \tag{21}$$

This constrained problem could be further reduced into the form as follows:

$$\max_{L^{(v)}}\ \mathrm{Tr}(L^{(v)^T}R^{(v)})\quad\text{s.t. } L^{(v)^T}L^{(v)}=I \tag{22}$$

where $R^{(v)}=(Z^{(v)}+\frac{1}{\mu}Y^{(v)})C^{(v)^T}$. Before giving the solution of this problem, we first give the following lemma:

Lemma 4. *For any matrices $A\in\mathbb{R}^{m\times n}$, suppose the singular value decomposition (SVD) of matrix A is $U\Lambda V^T$, then we consider the following constrained problem:*

$$\max_{Y}\ \mathrm{Tr}(Y^TA)\quad\text{s.t. } Y^TY=I \tag{23}$$

has closed form as follows:

$$Y=UV^T. \tag{24}$$

Proof. Since $A=U\Lambda V^T$, $U^TU=I$ and $V^TV=I$, thus we have:

$$\begin{aligned}&\mathrm{Tr}(Y^TA)\\ &=\mathrm{Tr}(Y^TU\Lambda V^T)\\ &=\mathrm{Tr}(\Lambda V^TY^TU)\\ &\le\mathrm{Tr}(\Lambda I)\end{aligned}$$

Note that $\mathrm{Tr}(\Lambda I)=\sum_{i=1}^{k}\delta_i$, where δ_i is the ith singular value of matrix A. Apparently, the above equality holds when $\Lambda P=\Lambda I$, that is $V^TY^TU=I$. Therefore, the closed solution is $Y=UV^T$. □

Based on the lemma 4, by performing the SVD decomposition of matrix $R^{(v)}$ as $R^{(v)}=U_R^{(v)}\Lambda_R^{(v)}W_R^{(v)^T}$, the solution for problem in Eq. (22) can be achieved as follows:

$$L^{(v)}=U_R^{(v)}W_R^{(v)^T} \tag{25}$$

3.4.3. Update rule for the variables $\{C^{(v)}\}_{v=1}^{V}$

When fixing the other variables, we can solve the following minimization subproblem w.r.t. variable $C^{(v)}$:

$$\min_{C^{(v)}} \lambda_1\|C^{(v)}\|_* + \frac{\lambda_2}{V-1}\sum_{\substack{v=1,\\ v\neq w}}\|C^{(v)} - C^{(w)}\|_F^2 \tag{26}$$
$$+ < Y^{(v)}, Z^{(v)} - L^{(v)}C^{(v)} > + \frac{\mu}{2}\|Z^{(v)} - L^{(v)}C^{(v)}\|_F^2$$

which can be further reduced into the following form:

$$\min_{C^{(v)}} \frac{\lambda_1}{2(\mu+\lambda_2)}\|C^{(v)}\|_* + \frac{1}{2}\|C^{(v)} - \frac{1}{\mu+\lambda_2}H^{(v)}\|_F^2, \tag{27}$$

where

$$H^{(v)} = \mu L^{(v)^T}(Z^{(v)} + \frac{1}{\mu}Y^{(v)}) + \frac{\lambda_2}{V-1}\sum_{w\neq v} C^{(w)}$$

According to the lemma in [58], the aforementioned problem can be solved efficiently by Singular Value Threshold (SVT) approach. The concrete lemma for SVT approach is introduced as follows:

Lemma 5. *For a given matrix* F *and a positive parameter* $\tau > 0$*, the optimal solution to the following problem*

$$\min_{D} \tau\|D\|_* + \frac{1}{2}\|D - F\|_F^2 \tag{28}$$

is given by

$$D = U_F\Theta_\tau(\Sigma_F)W_F{}^T. \tag{29}$$

where $U_F\Sigma_F W_F{}^T$ *is the SVD decomposition of matrix* F*. Meanwhile,* $\Theta_\tau(\cdot)$ *represents the SVT operator. The SVT operator is defined as follows:*

$$\Theta_\tau(\Sigma_F) = \max(0, \Sigma_F - \tau) + \min(0, \Sigma_F + \tau). \tag{30}$$

Based on this, by setting $\gamma = \frac{\lambda_1}{2(\mu+\lambda_2)}$, the closed-formed solution for variable $C^{(v)}$ is shown as follows:

$$C^{(v)} = U_H^{(v)}\Theta_\gamma(\Sigma_H^{(v)})W_H^{(v)^T}. \tag{31}$$

Specially, $U_H^{(v)}\Sigma_H^{(v)}W_H^{(v)^T}$ represents the SVD decomposition of $\frac{1}{\mu+\lambda_2}H^{(v)}$.

In this optimization algorithm, before the stopping criterions are met, the above three steps are updated in an iterative way. Subsequently, the final affinity matrix S could be obtained as follows:

$$\widetilde{S} = \frac{1}{V}\sum_{v=1}^{V} L^{(v)}C^{(v)} \tag{32}$$

and

$$S = \frac{|\widetilde{S}| + |\widetilde{S}|^T}{2}$$

In a nutshell, the detailed optimization process for FLMSC is summarized in Algorithm 1.

3.5. Computational complexity analysis

This subsection analyzes the computational complexity of our proposed approach. As shown in Algorithm 1, our approach mainly consists of three sub-processes in each iteration. For the first sub-process, the computation of the variables $\{Z^{(v)}\}_{v=1}^{V}$ takes $O(\sum_{v=1}^{V}\min(n^3, n^2d_v))$ time, where d_v denotes the dimension in the vth view, and n denotes the number of data samples. For the second sub-process, the computation of the variables $\{L^{(v)}\}_{v=1}^{V}$ takes $O(Vn^2)$ time, where V denotes the number of views. For the third sub-process, the computation of the variables $\{C^{(v)}\}_{v=1}^{V}$ takes $O(Vn)$ time. Thus, the overall time complexity of our approach is $O(T(\sum_{v=1}^{V}\min(n^3, n^2d_v) + Vn^2 + Vn))$, where T is the number of iterations. With $T \ll n, \gamma \ll n$, the overall time complexity of Algorithm 1 can be written as $O(\sum_{v=1}^{V}\min(n^3, n^2d_v))$. In general scenarios with $n > d_v$, the computational complexity of Algorithm 1 is $O(n^2\sum_{v=1}^{V}d_v)$. In the scenarios of very high-dimensional data with $n < d_v$, the computational complexity of Algorithm 1 is $O(n^3)$. Therefore, our FLMSC approach is generally more efficient than the conventional low-rank subspace clustering algorithms, especially for the large-scale data sets whose data sizes are much larger than their dimensions.

Algorithm 1: FLMSC.

Input: A collection of multi-view data samples $X = \{X^{(1)}, \cdots, X^{(V)}\}$, the number of clusters and balancing parameters λ_1, λ_2.

Parameter Setup: Set $\mu = 10^{-3}$, $\rho = 1.2$, $\mu_{max} = 10^6$ and γ = 100.

Initialization: Initialize $L^{(v)} = \mathbf{0}$, $C^{(v)} = \mathbf{0}$ and $Y^{(v)} = \mathbf{0}, \forall v$. Compute the $\{X^{(v)^T}X^{(v)}\}$ ($d_v \geqslant n$) or $\{X^{(v)}X^{(v)^T}\}$ ($n > d_v$) , $\forall v$. Set $t = 1$.
repeat
  Obtain $\{Z^{(v)}\}_{v=1}^{V}$ via Eq. (18) or Eq. (20).
  Obtain $\{L^{(v)}\}_{v=1}^{V}$ via Eq. (25).
  Obtain $\{C^{(v)}\}_{v=1}^{V}$ via Eq. (31).
  Obtain the multipliers $\{Y^{(v)}\}_{v=1}^{V}$ via: $Y^{(v)} = Y^{(v)} + \mu(Z^{(v)} - L^{(v)}C^{(v)})$.
  Obtain μ via $\mu = \min(\rho\mu, \mu_{\max})$.
  $t = t + 1$.
until The following stopping criterion is met or $t > t_{max}$.

$$\sum_{v=1}^{V}\|Z^{(v)} - L^{(v)}C^{(v)}\|_\infty < \varepsilon$$

Construct the graph similarity matrix via Eq. (32).
Perform Ncut algorithm on the graph similarity matrix S and generate the final labels for all data samples.
Output: The final labels for all data samples.

4. Experiments

This section carries out substantial experiments to validate the clustering property of FLMSC in terms of three perspectives, namely, comparison with state-of-arts, parameter sensitivity analysis and running time analysis. In this paper, all the experiments are conducted in Matlab 2019b on a PC with an Intel i5-6600 CPU and 64 GB of RAM.

4.1. Experimental materials and evaluation metrics

To verify both the efficiency and effectiveness of our approach, six practical data sets are adopted for the experimental purpose. The experimental data sets are selected with different characteristics, whose data size range from six hundred to more than ten thousand. Among these data sets, we use ALOI-100, Caltech101, UCI-Digit and VIS/NIR for image clustering, while adopting Reuters and BBC for document clustering. The detailed statistics of them are shown in Table 1. Meanwhile, we give the brief descriptions of them as follows:

Table 1
The statistics of six data sets used in the experiments.

Data set	#Sample	#View	#Class	Dimension
ALOI-100	10,800	4	100	Similarity (77), Haralick (13), HSV (64), RGB (125)
Caltech101	9,144	6	102	Gabor (48), WM (40), CENTRIST (254), HOG (1984), GIST (512), LBP (928)
UCI-Digit	2,000	3	10	FOU (76), PIX (240), MOR (6)
Reuters	1,200	5	6	French (2,000), German (2,000), English (2,000), Spanish (2,000), Italian(2,000)
VIS/NIR	1,056	2	22	Visible Light (10,000), Near-IR illumination (10,000)
BBC	685	4	6	View 1 (4,659), View 2 (4,633), View 3 (4,665), View 4 (4,684)

- **AlOI-100**[1] is a multi-view data set for object detection/clustering problems. Following the previous work, we use the first 100 categories from the Amsterdam Library Of Images (ALOI) to build ALOI-100. In ALOI-100 data set, it contains totally 10 800 object images, where each image is described by four different feature sets.
- **Caltech101**[2] is a widely-used image data set in many computer vision tasks. The data set is composed of 9144 object images covering 102 subjects (101 object categories and a background class), with about 40 to 800 images per subject. In the data set, each image is extracted from six visual feature sets (views).
- **UCI-digit**[3] is a classical image data set for handwritten digit recognition (from '0' to '9'). This data set contains 2000 images covering 10 classes. Similar to the previous work, we extract three different feature sets (views) for each image, i.e., Fourier coefficients (Fou), pixels averages (PIX) and morphological features (MOR).
- **Reuters**[4] is a multi-view data sets with documents available in five distinct languages. For our experiments, we randomly extract a sampling subset consisting of 1200 documents, with each category having 200 documents. In the experiments, five different languages are used as views.
- **VIS/NIR** [59] is a collection of face images for multi-modal face clustering problem. It is generated from two diverse sources (views), where each sources corresponding to one specific light condition. A subset containing 1056 images is constructed and used in the following experiments.
- **BBC** [44] is a collection of news reports derived from the BBC website. In total, there are 2225 documents over 5 annotated topics in this data set. In the experiments, we use a sampled subset of original BBC consisting of 685 documents and four different views, with 4659, 4633, 4665 and 4684 in each view, respectively.

For the evaluation metrics, Normalized Mutual Information (NMI) [60] as well as Accuracy (ACC) [61], are used to report clustering results in our experiments.

4.2. Comparison with state-of-arts

This subsection analyzes the performance of FLMSC and ten baselines by comparison experiments. To ensure the reliability and comprehensiveness of comparisons, three types of multi-view clustering approaches are selected as the baselines. As far as the first type of baselines is concerned, two classical graph-based approaches, e.g., SwMC [62] and GMC [63], are used for comparisons. For the second type of baselines, four representative subspace-based approaches (include two low-rank subspace-based approaches),[5] namely, DiMSC [48], LMSC [28,64], CSMSC [23] and COMVSC [65], are chosen in our experiments. Apart from the above two types of baselines, three large-scale subspace-based approaches are chosen as the third type of baselines, i.e., LMVSC [50], SMVSC [51] as well as FPMSC [66]. The detailed description for these baselines is shown in the original papers.

At the same time, we comply with the following experimental setting for fair comparisons:

- The source code of all baseline approaches is directly downloaded from the author's websites. Meanwhile, the optimal parameters of these baseline approaches are tuned in the range of $[1e^{-4}, 1e^{-3}, \ldots, 1e^{3}, 1e^{4}]$.
- For the proposed approach, the parameter γ is set as static value (i.e., 100) in the experiments. Besides, the optimal parameters (i.e., λ_1 and λ_2) of our approach are tuned in the range of $[1e^{-4}, 1e^{-3}, \ldots, 1e^{3}, 1e^{4}]$. The corresponding experimental settings of our method are listed in Table 2.
- For each experiment, we run all the clustering approaches 20 times under the same computing environment, and record the average results with standard deviations.
- Besides the above described parameters, the parameter K in LMSC and SMVSC will be tuned in the range of [10, 100, 300, 500], as suggested by the corresponding papers.

In terms of NMI and ACC, the comparison results are shown in Table 3 and Table 4 respectively. The best results in comparisons are highlighted in bold while the second best competitors are marked with underlines. By observing the results reported in these tables, four important observations could be obtained.

- Most of the baseline approaches (including the graph-based approaches as well as traditional subspace-based approaches) are not computationally suitable for ten-million-level data sets (i.e., AloI-101 and Caltech101). Out of the eleven comparison approaches, only the proposed approach and three large-scale subspace approaches are able to handle all of the experimental data sets.
- In the most of cases, traditional subspace-based approaches (i.e., DiMSC, LMSC, CSMSC, MCLES and COMVSC) obtain better clustering results over the large-scale subspace-based approaches (i.e., LMVSC, SMVSC and FPMVSC). However, the expensive time cost limit the applications of traditional subspace-based approaches when facing relatively large-scale clustering problem.
- For the document clustering tasks (i.e., Reuters and BBC), the performance of traditional subspace-based approaches are more reliable and consistent compared to that of the graph-based approaches. This finding indicates that the subspace-based approaches are more feasible for dealing with multi-view document clustering tasks.
- Finally but most important, results indicate that our approach overall outperforms the baseline approaches w.r.t. NMI and Accuracy metrics. Particularly, our approach obtains the best clustering performance on ALOI-100, UCI-Digit and VIS/NIR data sets, while it also shows very competitive performance on the rest of data sets.

[1] https://elki-project.github.io/datasets/multi_view
[2] http://www.vision.caltech.edu/Image_Datasets/Caltech101/
[3] http://archive.ics.uci.edu/ml/datasets/Multiple+Features
[4] http://lig-membres.imag.fr/grimal/data.html
[5] Note that LMSC, CSMSC are two widely used low-rank subspace-based approaches.

Table 2
Parameter setting of FLMSC on the six multi-view data sets.

	ALOI-100	Caltech101	UCI-Digit	Reuters	VIS/NIR	BBC
λ_1	$1e^{-3}$	$1e^{-3}$	$1e^{-3}$	$1e^{0}$	$1e^{-2}$	$1e^{-2}$
λ_2	$1e^{2}$	$1e^{2}$	$1e^{1}$	$1e^{2}$	$1e^{-2}$	$1e^{-1}$

* Note that the parameter γ is set as 100 in the experiments.

Table 3
Comparison results in terms of NMI (%) on six benchmark data sets.

	Approach	ALOI-100	Caltech101	UCI-Digit	Reuters	VIS/NIR	BBC	Avg.Rank
	SwMC	O/M	$15.9_{\pm 0.0}$	$69.8_{\pm 0.0}$	$4.1_{\pm 0.0}$	$96.1_{\pm 0.6}$	$55.8_{\pm 0.0}$	7.2
	GMC	O/M	O/M	$81.4_{\pm 0.0}$	$8.2_{\pm 0.0}$	$98.5_{\pm 0.0}$	$56.3_{\pm 0.0}$	5.5
	DiMSC	O/M	O/M	$79.2_{\pm 0.2}$	$27.3_{\pm 0.1}$	$99.3_{\pm 1.0}$	**81.5**$_{\pm 0.0}$	3.8
	LMSC	–	–	$73.6_{\pm 3.2}$	$31.8_{\pm 1.2}$	$97.8_{\pm 1.9}$	$69.8_{\pm 3.5}$	5.2
	CSMSC	O/M	O/M	$79.8_{\pm 0.4}$	$25.9_{\pm 0.0}$	$98.5_{\pm 0.9}$	$77.6_{\pm 0.0}$	4.2
	COMVSC	O/M	O/M	$75.0_{\pm 1.3}$	$13.0_{\pm 1.0}$	**99.4**$_{\pm 0.1}$	$37.2_{\pm 8.2}$	6.0
	LMVSC	$74.9_{\pm 0.0}$	$27.2_{\pm 0.0}$	$75.6_{\pm 0.0}$	**36.0**$_{\pm 0.0}$	$72.1_{\pm 0.0}$	$56.7_{\pm 0.0}$	4.7
	SMVSC	$57.4_{\pm 0.0}$	$37.0_{\pm 0.0}$	$80.4_{\pm 0.0}$	$18.9_{\pm 0.0}$	$94.1_{\pm 0.0}$	$69.0_{\pm 0.0}$	4.5
	FPMVS	$55.5_{\pm 0.0}$	$35.9_{\pm 0.0}$	$67.2_{\pm 0.0}$	$20.6_{\pm 0.0}$	$73.1_{\pm 0.0}$	$7.1_{\pm 0.0}$	6.8
	FLMSC	**79.6**$_{\pm 0.5}$	**50.5**$_{\pm 0.4}$	**83.4**$_{\pm 0.0}$	$32.0_{\pm 0.1}$	**99.4**$_{\pm 0.1}$	$76.2_{\pm 0.0}$	**1.5**

* Note that "O/M" means the out-of-memory error; And "-" means the algorithm exceeds more than 24 h.

Table 4
Comparison results in terms of ACC (%) on six benchmark data sets.

	Approach	ALOI-100	Caltech101	UCI-Digit	Reuters	VIS/NIR	BBC	Avg.Rank
	SwMC	O/M	$16.1_{\pm 0.0}$	$65.5_{\pm 0.0}$	$18.4_{\pm 0.0}$	$91.6_{\pm 1.0}$	$69.2_{\pm 0.0}$	7.5
	GMC	O/M	O/M	$78.0_{\pm 0.0}$	$24.2_{\pm 0.0}$	$95.5_{\pm 0.0}$	$69.3_{\pm 0.0}$	6.2
	DiMSC	O/M	O/M	$86.5_{\pm 0.4}$	$46.9_{\pm 0.1}$	$97.7_{\pm 1.0}$	**93.6**$_{\pm 0.0}$	3.5
	LMSC	–	–	$81.7_{\pm 5.6}$	$52.8_{\pm 1.5}$	$94.3_{\pm 1.6}$	$88.2_{\pm 4.7}$	4.7
	CSMSC	O/M	O/M	$87.4_{\pm 0.5}$	$41.3_{\pm 0.0}$	$95.2_{\pm 2.8}$	$92.0_{\pm 0.0}$	4.2
	COMVSC	O/M	O/M	$77.4_{\pm 2.4}$	$32.1_{\pm 1.8}$	**99.6**$_{\pm 0.1}$	$54.8_{\pm 8.3}$	6.2
	LMVSC	$56.6_{\pm 0.0}$	$29.1_{\pm 0.0}$	$80.3_{\pm 0.0}$	$53.3_{\pm 0.0}$	$57.4_{\pm 0.0}$	$76.4_{\pm 0.0}$	4.5
	SMVSC	$34.3_{\pm 0.0}$	$26.7_{\pm 0.0}$	$83.4_{\pm 0.0}$	$38.1_{\pm 0.0}$	$93.2_{\pm 0.0}$	$87.5_{\pm 0.0}$	4.8
	FPMVS	$31.5_{\pm 0.0}$	**29.5**$_{\pm 0.0}$	$72.2_{\pm 0.0}$	$44.3_{\pm 0.0}$	$58.8_{\pm 0.0}$	$35.2_{\pm 0.0}$	6.2
	FLMSC	**67.9**$_{\pm 1.5}$	$26.7_{\pm 0.8}$	**91.4**$_{\pm 0.0}$	**56.6**$_{\pm 0.2}$	**99.6**$_{\pm 0.1}$	$90.8_{\pm 0.0}$	**1.7**

* Note that "O/M" means the out-of-memory error; And "-" means the algorithm exceeds more than 24 h.

The reason for the remarkable performance maybe that, our approach benefits from utilizing the inherent characteristics of multi-view data, (i.e., incorporating the view-specific low-rank structure and cross-view structural consistency for subspace representation learning). Thus, the final affinity graph well exploits the view-specificity as well as complementary information of multiple views, which enables the underlying cluster structure to be easily detected by the following spectral clustering. In summary, the extensive results reflect the effectiveness of FLMSC approach in image and document clustering applications.

4.3. Parameter sensitivity analysis

This subsection investigates the parameter sensitivity of parameters λ_1 and λ_2 in the proposed approach. Based on the grid search strategy, we fix one parameter as a static value, and allow the other parameter to vary in the relatively large range of $\{1e^{-4}, 1e^{-3}, \ldots, 1e^{3}, 1e^{4}\}$. At the same time, the NMI as well as Accuracy are selected as the clustering metrics in the experiments.

The effects of different balancing parameters on clustering results are given in Figs. 2 to 3, respectively. As can be seen, when the parameter λ_2 is tuned, our approach maintains relatively stable performance on the ALOI101, Caltech101 and BBC data sets. In contrast, the clustering performance of our approach is sensitive to the parameter perturbation of λ_1 on six tested data sets. As a result, balancing parameter λ_1 is crucial in our approach.

4.4. Running time analysis and convergence analysis

In this subsection, we analyze the computational efficiency of FLMSC approach and the baseline approaches on the tested benchmark data sets. Specially, comparison results reported in Table 5 indicate the following three observations.

- Large-scale subspace-based approaches perform much faster than the traditional subspace-based approaches, which demonstrates the high efficiency of anchor graph learning strategy.
- In the most cases, our approach shows comparable efficiency over two large-scale subspace approaches, e.g., SMVSC and FPMVSC. On some data sets, the running efficiency of our approach is even higher than these two approaches. The reason is that the proposed approach requires few iterations to converge.
- Notably, our approach needs much shorter time over the traditional subspace-based approaches. Especially, the proposed approach is significantly more efficient than two low-rank subspace-based approaches, e.g., LMSC and CSMSC. This is because the low-rank subspace-based approaches have to solve the whole SVD operations at each iteration.

Overall, in terms of the running speed, the proposed approach ranks at upper level among the eight multi-view subspace approaches, which proves the relatively high efficiency of our approach by experimental results. Besides, in the empirical experiments, the running time are determined by many factors, like the

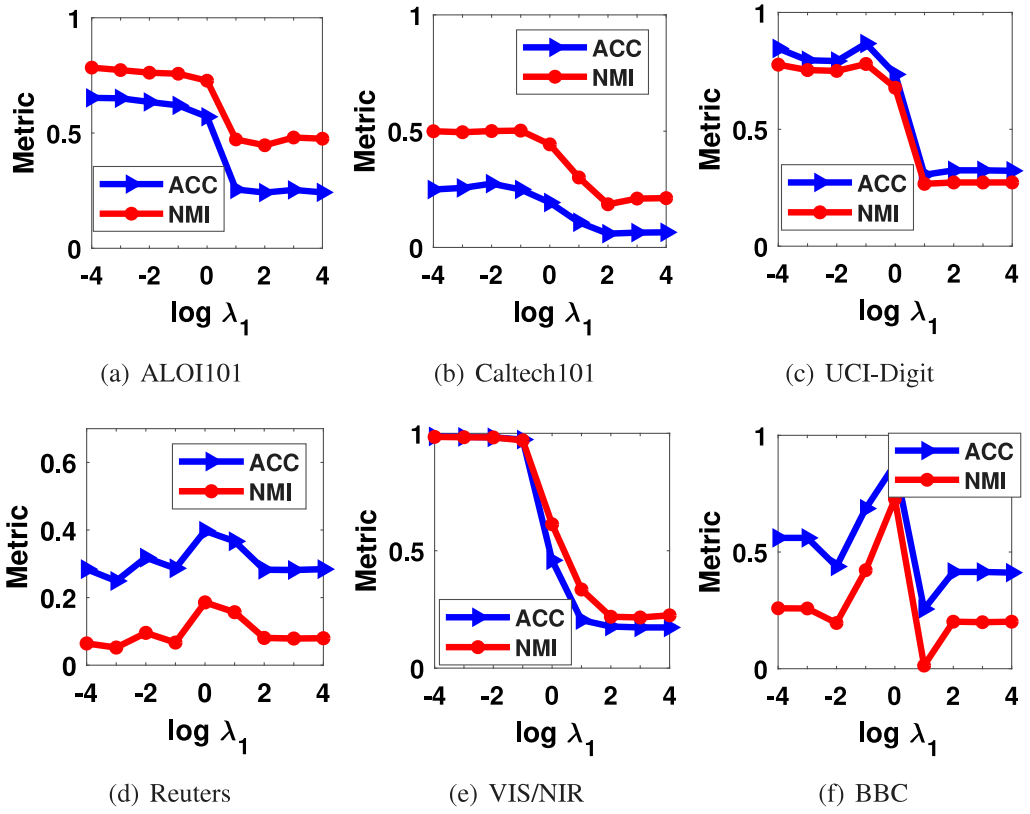


Fig. 2. Analysis on the effect of parameter λ_1 by fixing parameter λ_2 as 1.

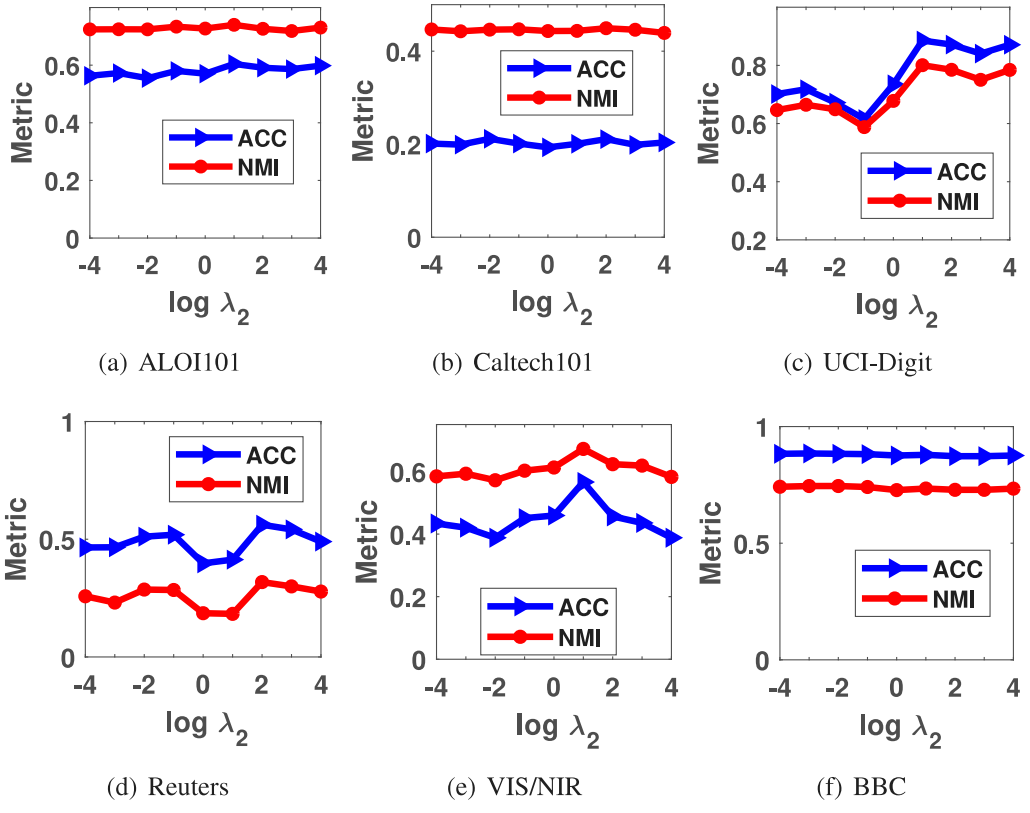


Fig. 3. Analysis on the effect of parameter λ_2 by fixing parameter λ_1 as 1.

running platform, the iteration times, the distribution structure of data sets and so on. As the future work, it would be interesting to investigate the potential relationship between complexity model and empirical time in the clustering research.

Table 5
Running time (in seconds) of all clustering approaches on six benchmark data sets.

Approach	ALOI-100	Caltech101	UCI-Digit	Reuters	VIS/NIR	BBC	Avg.Rank
SwMC	O/M	5012.0	521.2	4.4	42.2	26.2	5.3
GMC	O/M	O/M	10.2	6.4	3.4	2.9	3.2
DiMSC	O/M	O/M	116.9	121.7	12.7	21.1	6.0
LMSC	–	–	223.7	93.6	134.5	121.4	7.2
CSMSC	O/M	O/M	157.5	160.7	192.2	124.3	7.8
COMVSC	O/M	O/M	225.7	160.0	70.6	63.6	7.2
LMVSC	25.5	48.2	4.4	2.9	9.0	1.7	1.5
SMVSC	124.9	682.3	14.0	9.0	44.2	18.7	3.7
FPMVC	241.0	400.5	22.0	16.9	136.6	36.2	5.2
FLMSC	532.4	1092.7	20.1	21.2	5.9	6.1	3.8

* Note that "O/M" means the out-of-memory error; And "-" means the algorithm exceeds more than 24 h.

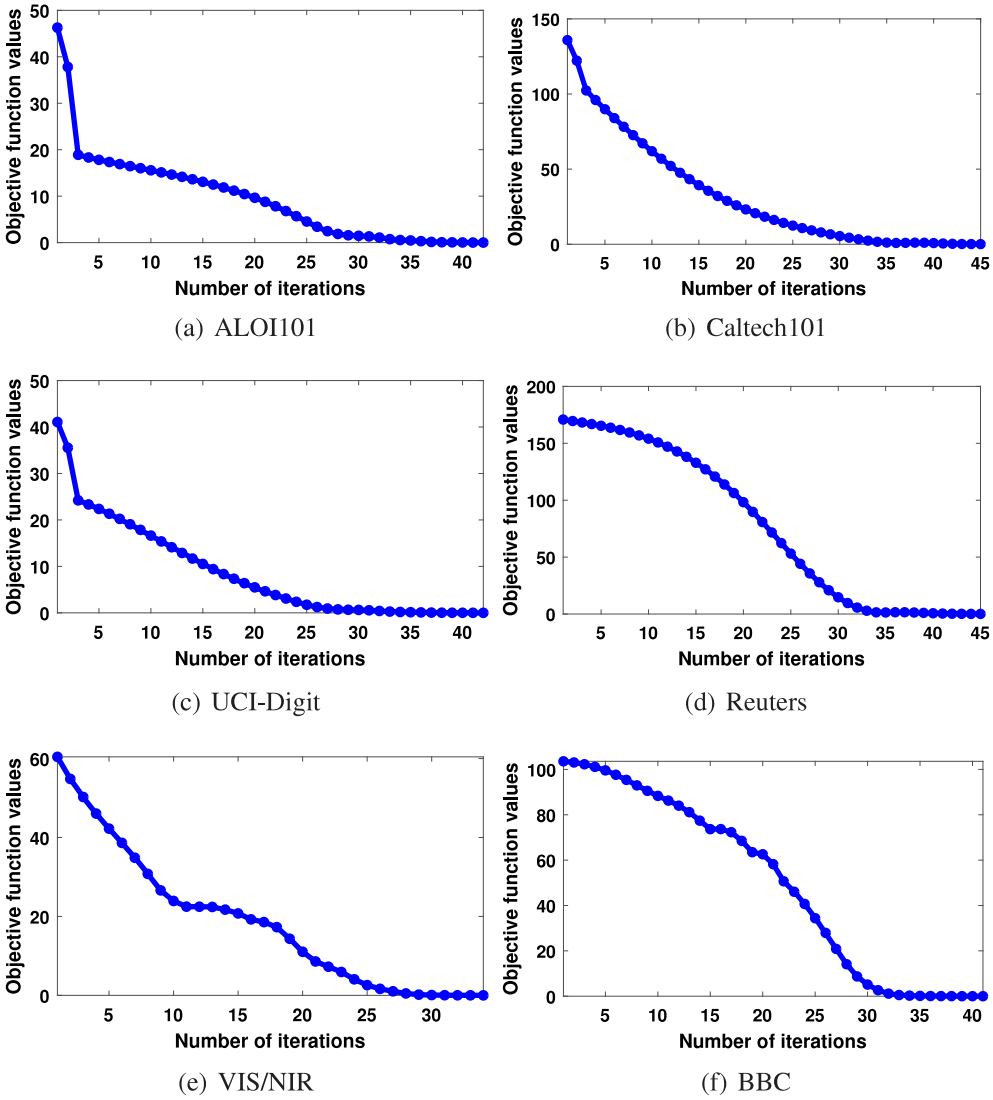


Fig. 4. The convergence curves of our method on six data sets.

Additionally, we analyze the convergence property of our approach empirically. The convergence curves with increasing iterations are plotted in Fig. 4. As observed, the loss of our approach (see the stopping criterion in Algorithm 1) monotonically decreases at each iteration and the convergence condition is generally met less than 45 iterations.

4.5. Statistical test

In this subsection, we further investigate the experimental results in Table 3 to Table 5 from statistical perspective. Following the previous works in [24,67,68], we use the t-test with $p < 0.05$ and the statistical-test results are recorded in Table 6. As can be

Table 6
The t-test comparison results ($p < 0.05$) of the proposed approach over baseline approaches in terms of NMI, ACC and Running Time on six benchmarks. The three integers inside the brackets respectively correspond to the number of times that the performance of our approach is (significantly better than, comparable to, and significantly worse than) a baseline approach.

	Caltech101	Caltech101	UCI-Digit	Reuters	VIS/NIR	BBC
NMI	(**3** / 0 / 0)	(**4** / 0 / 0)	(**9** / 0 / 0)	(**7** / 1 / 1)	(**7** / 2 / 0)	(**7** / 0 / 2)
ACC	(**3** / 0 / 0)	(**1** / 1 / 2)	(**9** / 0 / 0)	(**9** / 0 / 0)	(**8** / 1 / 0)	(**6** / 1 / 2)
Running time	(**0** / 0 / 3)	(**1** / 0 / 3)	(**5** / 0 / 4)	(**4** / 0 / 5)	(**8** / 0 / 1)	(**7** / 0 / 2)
Overall	(**6** / 0 / 3)	(**6** / 1 / 5)	(**23** / 0 / 4)	(**20** / 1 / 6)	(**23** / 3 / 1)	(**20** / 1 / 6)

observed, our approach shows significant improvements over the baseline approaches in terms of ACC and NMI. Besides, although our approach has not shown the same significant improvement like the ACC and NMI metrics, our approach still significantly outperforms the competitors at least 4 times in terms of the running time. Overall, the above t-test results have validated the superiority of our method from the statistical perspective.

5. Conclusion

This paper proposes a novel approach for Low-rank multi-view subspace clustering, which aims to learn robust subspace representations from effective and efficient perspectives. Notably, our approach is able not only to preserve the low-rank structure within individual view, but also to capture the structural consistence across different views. This design ensures the proposed approach to flexibly suit the characteristics of multi-view data, in which the subsequent spectral clustering can easily detect the underlying cluster structure. Moreover, our approach enjoy high efficiency by converting the original SVD problem into smaller-scale optimization problem. Finally, the experimental results on six benchmark data sets confirm the effectiveness as well as efficiency of the proposed approach. In the future, we plan to extend our FLMSC approach with deep feature learning, so as to design more effective multi-view subspace clustering approaches with the facilitated strategy.

CRediT authorship contribution statement

Guang-Yu Zhang: Formal analysis, Investigation, Methodology, Resources, Software, Writing – original draft, Writing – review & editing. **Dong Huang:** Original draft, Data curation, Funding acquisition, Writing – review & editing, Supervision. **Chang-Dong Wang:** Formal analysis, Investigation, Methodology, Funding acquisition, Supervision.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgments

This work was supported by the NSFC, China (62206099 & 61976097 & 61876193).

References

[1] S. Sun, A survey of multi-view machine learning, Neural Comput. Appl. 23 (2013) 2031–2038.
[2] W. Wang, R. Arora, K. Livescu, J. Bilmes, On deep multi-view representation learning, in: Proceedings of the 32th International Conference on Machine Learning, 2015, pp. 1083–1092.
[3] B. Liu, X. Chen, Y. Xiao, W. Li, L. Liu, C. Liu, An efficient dictionary-based multi-view learning method, Inform. Sci. 576 (2021) 157–172.
[4] M. Kan, S. Shan, X. Chen, Multi-view deep network for cross-view classification, in: Proceedings of the 29th IEEE Conference on Computer Vision and Pattern Recognition, 2016, pp. 4847–4855.
[5] C. Wang, X. Chen, B. Chen, F. Nie, B. Wang, Z. Ming, Learning unsupervised node representation from multi-view network, Inform. Sci. 579 (2021) 700–716.
[6] C. Hou, F. Nie, H. Tao, D. Yi, Multi-view unsupervised feature selection with adaptive similarity and view weight, IEEE Trans. Knowl. Data Eng. 29 (2017) 1998–2011.
[7] C. Tang, J. Chen, X. Liu, M. Li, P. Wang, M. Wang, et al., Consensus learning guided multi-view unsupervised feature selection, Knowl.-Based Syst. 160 (2018) 49–60.
[8] C. Zhang, H. Fu, Q. Hu, P. Zhu, X. Cao, Flexible multi-view dimensionality co-reduction, IEEE Trans. Image Process. 26 (2017) 648–659.
[9] C. Zhang, Y. Liu, Y. Liu, Q. Hu, X. Liu, P. Zhu, FISH-MML: Fisher-HSIC multi-view metric learning, in: Proceedings of the 27th International Joint Conference on Artificial Intelligence, 2018, pp. 3054–3060.
[10] S. Li, M. Shao, Y. Fu, Multi-view low-rank analysis for outlier detection, in: Proceedings of the 15th SIAM International Conference on Data Mining, 2015, pp. 748–756.
[11] C.G. Li, R. Vidal, Structured sparse subspace clustering: A unified optimization framework, in: Proceedings of the 28th IEEE Conference on Computer Vision and Pattern Recognition, 2015, pp. 277–286.
[12] C. Xu, D. Tao, C. Xu, Multi-view self-paced learning for clustering, in: Proceedings of the 27th International Joint Conference on Artificial Intelligence, 2015, pp. 3974–3980.
[13] S. Huang, Z. Kang, Z. Xu, Self-weighted multi-view clustering with soft capped norm, Knowl.-Based Syst. 158 (2018) 1–8.
[14] X. Chen, X. Xu, J.Z. Huang, Y. Ye, TW-k-means: automated two-level variable weighting clustering algorithm for multiview data, IEEE Trans. Knowl. Data Eng. 25 (2013) 932–944.
[15] X. Cai, F. Nie, H. Huang, Multi-view k-means clustering on big data, in: Proceedings of the 23th International Joint Conference on Artificial Intelligence, 2013, pp. 2598–2604.
[16] D. Greene, P. Cunningham, A matrix factorization approach for integrating multiple data views, in: Proceedings of the 13th Joint European Conference on Machine Learning and Knowledge Discovery in Databases, 2009, pp. 423–438.
[17] J. Liu, C. Wang, J. Gao, J. Han, Multi-view clustering via joint nonnegative matrix factorization, in: Proceedings of the 13th SIAM International Conference on Data Mining, SIAM, 2013, pp. 252–260.
[18] G. Tzortzis, A. Likas, Kernel-based weighted multi-view clustering, in: Proceedings of the 12th IEEE International Conference on Data Mining, 2012, pp. 675–684.
[19] X. Liu, Y. Dou, J. Yin, L. Wang, E. Zhu, Multiple kernel k-means clustering with matrix-induced regularization, in: Proceedings of the 30th AAAI Conference on Artificial Intelligence, 2016, pp. 1888–1894.
[20] F. Nie, G. Cai, X. Li, Multi-view clustering and semi-supervised classification with adaptive neighbours, in: Proceedings of the 31st AAAI Conference on Artificial Intelligence, 2017, pp. 2408–2414.
[21] Y.W. Liang, D. Huang, C.D. Wang, Consistency meets inconsistency: A unified graph learning framework for multi-view clustering, in: Proceedings of the 19th IEEE International Conference on Data Mining, 2019, pp. 1204–1209.
[22] C. Tang, X. Zhu, X. Liu, M. Li, P. Wang, C. Zhang, et al., Learning a joint affinity graph for multiview subspace clustering, IEEE Trans. Multimed. 21 (2018) 1724–1736.
[23] S. Luo, C. Zhang, W. Zhang, X. Cao, Consistent and specific multi-view subspace clustering, in: Proceedings of the 32th AAAI Conference on Artificial Intelligence, 2018, pp. 3730–3737.
[24] G.Y. Zhang, Y.R. Zhou, X.Y. He, C.D. Wang, D. Huang, One-step kernel multi-view subspace clustering, Knowl.-Based Syst. (2019) 105126.
[25] C. Zhang, H. Fu, J. Wang, W. Li, X. Cao, Q. Hu, Tensorized multi-view subspace representation learning, Int. J. Comput. Vis. 128 (2020) 2344–2361.
[26] Z. Xue, J. Du, D. Du, S. Lyu, Deep low-rank subspace ensemble for multi-view clustering, Inform. Sci. 482 (2019) 210–227.

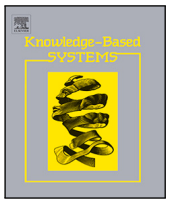
[27] X. Zhang, H. Sun, Z. Liu, Z. Ren, Q. Cui, Y. Li, Robust low-rank kernel multi-view subspace clustering based on the schatten p-norm and correntropy, Inform. Sci. 477 (2019) 430–447.
[28] C. Zhang, Q. Hu, H. Fu, P. Zhu, X. Cao, Latent multi-view subspace clustering, in: Proceedings of the 30th IEEE Conference on Computer Vision and Pattern Recognition, 2017, pp. 4279–4287.
[29] M. Abavisani, V.M. Patel, Multimodal sparse and low-rank subspace clustering, Inf. Fusion 39 (2018) 168–177.
[30] R. Li, C. Zhang, Q. Hu, P. Zhu, Z. Wang, Flexible multi-view representation learning for subspace clustering, in: Proceedings of the 28th International Joint Conference on Artificial Intelligence, 2019, pp. 2916–2922.
[31] T. Zhou, C. Zhang, X. Peng, H. Bhaskar, J. Yang, Dual shared-specific multiview subspace clustering, IEEE Trans. Cybern. 50 (2019) 3517–3530.
[32] S. Bickel, T. Scheffer, Multi-view clustering, in: Proceedings of the 4th IEEE International Conference on Data Mining, Vol. 4, 2004, pp. 19–26.
[33] V.R. De Sa, Spectral clustering with two views, in: Proceedings of the 22th International Conference on Machine Learning Workshop on Learning with Multiple Views, 2005, pp. 20–27.
[34] A. Kumar, H. Daumé, A co-training approach for multi-view spectral clustering, in: Proceedings of the 28th International Conference on Machine Learning, 2011, pp. 393–400.
[35] H. Wang, Y. Yang, B. Liu, H. Fujita, A study of graph-based system for multi-view clustering, Knowl.-Based Syst. 163 (2019) 1009–1019.
[36] X. Zhu, S. Zhang, R. Hu, W. He, C. Lei, P. Zhu, One-step multi-view spectral clustering, IEEE Trans. Knowl. Data Eng. 31 (2018) 2022–2034.
[37] K. Zhan, C. Niu, C. Chen, F. Nie, C. Zhang, Y. Yang, Graph structure fusion for multiview clustering, IEEE Trans. Knowl. Data Eng. 31 (2018) 1984–1993.
[38] Y. Li, F. Nie, H. Huang, J. Huang, Large-scale multi-view spectral clustering via bipartite graph, in: Proceedings of the 29th AAAI Conference on Artificial Intelligence, 2015, pp. 2750–2756.
[39] H. Zhao, Z. Ding, Y. Fu, Multi-view clustering via deep matrix factorization, in: Proceedings of the 31st AAAI Conference on Artificial Intelligence, 2017, pp. 2921–2927.
[40] G. Trigeorgis, K. Bousmalis, S. Zafeiriou, B.W. Schuller, A deep matrix factorization method for learning attribute representations, IEEE Trans. Pattern Anal. Mach. Intell. 39 (2016) 417–429.
[41] S. Huang, Z. Kang, I.W. Tsang, Z. Xu, Auto-weighted multi-view clustering via kernelized graph learning, Pattern Recognit. 88 (2019) 174–184.
[42] G.Y. Zhang, C.D. Wang, D. Huang, W.S. Zheng, Y.R. Zhou, TW-Co-k-means: Two-level weighted collaborative k-means for multi-view clustering, Knowl.-Based Syst. 150 (2018) 127–138.
[43] Y. Jiang, F.L. Chung, S. Wang, Z. Deng, J. Wang, P. Qian, Collaborative fuzzy clustering from multiple weighted views, IEEE Trans. Cybern. 45 (2014) 688–701.
[44] L. Huang, H.Y. Chao, C.D. Wang, Multi-view intact space clustering, Pattern Recognit. 86 (2019) 344–353.
[45] Y. Yang, H. Wang, Multi-view clustering: a survey, Big Data Min. Anal. 1 (2) (2018) 83–107.
[46] M.S. Chen, J.Q. Lin, X.L. Li, B.Y. Liu, C.D. Wang, D. Huang, et al., Representation learning in multi-view clustering: A literature review, Data Sci. Eng. (2022) 1–17.
[47] H. Gao, F. Nie, X. Li, H. Huang, Multi-view subspace clustering, in: Proceedings of the 28th IEEE International Conference on Computer Vision, 2015, pp. 4238–4246.
[48] X. Cao, C. Zhang, H. Fu, S. Liu, H. Zhang, Diversity-induced multi-view subspace clustering, in: Proceedings of the 28th IEEE Conference on Computer Vision and Pattern Recognition, 2015, pp. 586–594.
[49] M.S. Chen, L. Huang, C.D. Wang, D. Huang, Multi-view clustering in latent embedding space, in: Proceedings of the 34th AAAI Conference on Artificial Intelligence, 2020.
[50] Z. Kang, W. Zhou, Z. Zhao, J. Shao, M. Han, Z. Xu, Large-scale multi-view subspace clustering in linear time, in: Proceedings of the 34th AAAI Conference on Artificial Intelligence, Vol. 34, 2020, pp. 4412–4419.
[51] M. Sun, P. Zhang, S. Wang, S. Zhou, W. Tu, X. Liu, et al., Scalable multi-view subspace clustering with unified anchors, in: Proceedings of the 29th ACM International Conference on Multimedia, 2021, pp. 3528–3536.
[52] M.S. Chen, C.D. Wang, D. Huang, J.H. Lai, P.S. Yu, Efficient orthogonal multi-view subspace clustering, in: Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining, 2022, pp. 127–135.
[53] R. Xia, Y. Pan, L. Du, J. Yin, Robust multi-view spectral clustering via low-rank and sparse decomposition, in: Proceedings of the 28th AAAI Conference on Artificial Intelligence, 2014, pp. 2149–2155.
[54] Y. Pan, C.Q. Huang, D. Wang, Multiview spectral clustering via robust subspace segmentation, IEEE Trans. Cybern. 52 (2022) 2467–2476.
[55] X. Cai, D. Huang, G. Zhang, C.D. Wang, Seeking commonness and inconsistencies: A jointly smoothed approach to multi-view subspace clustering, Inf. Fusion 91 (2023) 364–375.
[56] G. Liu, Z. Lin, S. Yan, J. Sun, Y. Yu, Y. Ma, Robust recovery of subspace structures by low-rank representation, IEEE Trans. Pattern Anal. Mach. Intell. 35 (1) (2012) 171–184.
[57] S. Boyd, N. Parikh, E. Chu, B. Peleato, J. Eckstein, et al., Distributed optimization and statistical learning via the alternating direction method of multipliers, Found. Trends Mach. Learn. 3 (1) (2011) 1–122.
[58] J.F. Cai, E.J. Candès, Z. Shen, A singular value thresholding algorithm for matrix completion, SIAM J. Optim. 20 (4) (2010) 1956–1982.
[59] C.D. Wang, J.H. Lai, S.Y. Philip, Multi-view clustering based on belief propagation, IEEE Trans. Knowl. Data Eng. 28 (4) (2015) 1007–1021.
[60] D. Huang, C.D. Wang, J. Wu, J.H. Lai, C.K. Kwoh, Ultra-scalable spectral clustering and ensemble clustering, IEEE Trans. Knowl. Data Eng. 32 (2020) 1212–1226.
[61] Y.W. Liang, D. Huang, C.D. Wang, P.S. Yu, Multi-view graph learning by joint modeling of consistency and inconsistency, IEEE Trans. Neural Netw. Learn. Syst. (2022).
[62] F. Nie, J. Li, X. Li, Self-weighted multiview clustering with multiple graphs, in: Proceedings of the 26th International Joint Conference on Artificial Intelligence, 2017, pp. 2564–2570.
[63] H. Wang, Y. Yang, B. Liu, GMC: Graph-based multi-view clustering, IEEE Trans. Knowl. Data Eng. 32 (2020) 1116–1129.
[64] C. Zhang, H. Fu, Q. Hu, X. Cao, Y. Xie, D. Tao, et al., Generalized latent multi-view subspace clustering, IEEE Trans. Pattern Anal. Mach. Intell. 42 (1) (2018) 86–99.
[65] P. Zhang, X. Liu, J. Xiong, S. Zhou, W. Zhao, E. Zhu, et al., Consensus one-step multi-view subspace clustering, IEEE Trans. Knowl. Data Eng. 34 (10) (2022) 4676–4689.
[66] S. Wang, X. Liu, X. Zhu, P. Zhang, Y. Zhang, F. Gao, et al., Fast parameter-free multi-view subspace clustering with consensus anchor guidance, IEEE Trans. Image Process. 31 (2021) 556–568.
[67] D. Huang, C.D. Wang, J.H. Lai, Locally weighted ensemble clustering, IEEE Trans. Cybern. 48 (5) (2017) 1460–1473.
[68] D. Huang, C.D. Wang, H. Peng, J. Lai, C.K. Kwoh, Enhanced ensemble clustering via fast propagation of cluster-wise similarities, IEEE Trans. Syst. Man Cybern. Syst. 51 (1) (2021) 508–520.

Knowledge-Based Systems 312 (2025) 113119

Contents lists available at ScienceDirect

Knowledge-Based Systems

journal homepage: www.elsevier.com/locate/knosys



Scalable tri-factorization guided multi-view subspace clustering

Guang-Yu Zhang [a], Chang-Bin Guan [a,b], Dong Huang [a], Zihao Wen [a,*], Chang-Dong Wang [c,d], Lei Xiao [a]

[a] College of Mathematics and Informatics, South China Agricultural University, Guangzhou, China
[b] State Key Laboratory of Swine and Poultry Breeding Industry, Guangzhou, China
[c] School of Computer Science and Engineering, Sun Yat-sen University, Guangzhou, China
[d] Guangdong Key Laboratory of Big Data Analysis and Processing, Guangzhou, China

ARTICLE INFO

Keywords:
Data clustering
Multi-view clustering
Anchor-based multi-view subspace clustering
Tri-factorization
Low-rank tensor learning

ABSTRACT

Anchor-based Multi-view Subspace Clustering (AMSC) has exhibited its outstanding capability in large-scale multi-view clustering. Despite significant progress, previous AMSC approaches still suffer from two limitations. First, they mostly neglect the high-order correlation, which undermines their ability in discovering complex cluster structures. Second, they frequently overlook the potential connection between multi-view dimension reduction and anchor subspace clustering, which affects their robustness to low-quality views. In view of these issues, we present a Scalable Tri-factorization Guided Multi-view Subspace Clustering (ST-MSC) approach. Specifically, the proposed approach seeks to recover the latent sample-anchor relationships in multiple embedded spaces, where the multi-view anchor representations are stacked into a low-rank tensor to enhance their high-order correlations with tri-factorization guidance. Theoretical analysis indicates that the tri-factorization paradigm has inherent relevance with two mutually beneficial tasks, namely, the multi-view dimensionality reduction and the anchor-based multi-view subspace clustering. Furthermore, a simple yet fast algorithm is devised to minimize the objective model, where the latent embedding spaces and the anchor subspace structure can be iteratively updated in a unified manner. Experiments have been conducted to verify the effectiveness and efficiency of our ST-MSC approach in comparison with the advanced approaches.

1. Introduction

Data clustering, as an essential yet challenging research problem in machine learning and knowledge discovery, has been widely studied over the past decades. Conventional clustering approaches, such as k-means [1] or spectral clustering [2], are mostly designed for single-view datasets, whose features are typically collected from a single source or depicted by a single feature descriptor. To go beyond the single-view to multi-view scenarios, the multi-view clustering technique rapidly emerges in recent years, which targets at enhancing the clustering robustness by exploring the rich and complementary information from multiple sources (or views).

In the multi-view clustering research, the graph-based approaches have received significant popularity owing to their high interpretability and flexibility. With the aim of capturing the shared block-diagonal structure across the graphs of various views, many graph-based approaches [3–6] have been developed with different graph learning strategies. In these graph-based approaches, the core steps involve building a unified affinity matrix (or spectral embedding) from different view-specific graphs, followed by performing the spectral clustering to generate the discrete label. Thereby, the quality of the pre-defined graphs (including the corresponding parameter settings) may heavily affect the clustering quality. Besides the graph-based approaches, another research hotspot is the subspace-based approaches. Derived from the self-expressive model, the subspace approaches assume that data samples (within each individual view) are embedded in a set of latent subspaces, with the purpose of learning the latent subspace representations among multiple views. Over the past few years, a considerable number of subspace-based approaches [7–9] have been presented, which benefit from their theoretical guarantees and have achieved significant success in many applications.

Motivation. Despite the impressive efforts that have been made, many multi-view clustering approaches still face an intensive computational burden, which severely limits their practical capacity in large-scale or even medium-scale scenarios. Aiming to break through this bottleneck, the anchor-based multi-view subspace clustering approached

* Corresponding author.
E-mail addresses: guangyuzhg@foxmail.com (G.-Y. Zhang), kwansheung@163.com (C.-B. Guan), huangdonghere@gmail.com (D. Huang), zihao.wen@hotmail.com (Z. Wen), changdongwang@hotmail.com (C.-D. Wang), lein_xiao@scau.edu.cn (L. Xiao).

https://doi.org/10.1016/j.knosys.2025.113119
Received 8 August 2024; Received in revised form 7 January 2025; Accepted 31 January 2025
Available online 8 February 2025

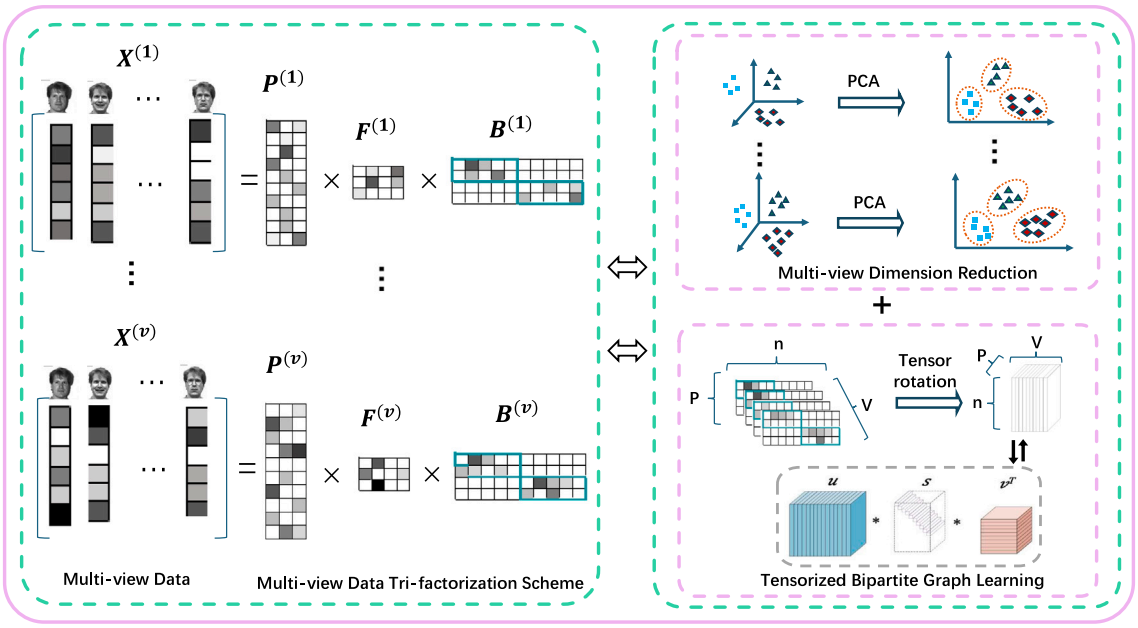


Fig. 1. The flowchart of the proposed approach.

(abbreviated as anchor subspace-based approaches in the following) has shown its promising ability in tackling large-scale problems. Instead of building a set of $n \times n$ subspace representations from multiple sources, the anchor subspace-based approaches seek to model the latent anchor-sample relationship by a union of $n \times p$ anchor representations, where n is the number of data samples and p is the number of anchors (also known as landmarks [10] or representatives [2]). While the anchor subspace-based approaches have exhibited their advantages in efficiency, which typically reduce the time complexity (of conventional multi-view subspace clustering) from $O(n^3)$ to $O(n)$ (with the condition $p \ll n$ holds), there are still two essential questions that remain to be addressed. First, how to characterize the high-order correlation among different anchor representations? Second, how to effectively leverage the multi-view dimension reduction in collaboration with anchor multi-view subspace clustering for more robust clustering performance?

Our work. In view of these questions, we present a **S**calable **T**ri-factorization guided **M**ulti-view **S**ubspace **C**lustering (ST-MSC) approach. Our approach aims to capture the deep anchor-sample relationship hidden in multiple embedded spaces. Typically, the t-SVD tensor decomposition technique is utilized to capture the high-order correlation across different anchor representations. As illustrated in Fig. 1, our tri-factorization mechanism can bridge the gap between two multi-view learning tasks, namely, multi-view dimension reduction and anchor-based multi-view subspace clustering. Particularly, a unified objective model with theoretical guarantee is formulated, for which a highly efficient optimization algorithm is designed. Comprehensive experiments have been conducted on nine benchmark datasets, which confirm the superiority of our ST-MSC approach over the state-of-the-art.

For clarity, we conclude the main contributions of this work in the following:

- This paper formulates two mutually beneficial tasks, e.g., multi-view dimensionality reduction and anchor multi-view subspace clustering, into a unified tri-factorization framework.
- A tri-factorization guided multi-view subspace clustering approach called ST-MSC is presented, which is featured by its highly competitive time efficiency and robust clustering performance.
- An efficient optimization algorithm is devised to solve our model, which enjoys low time complexity in terms of both sample size and dimension size. Experiments on several real-world benchmarks demonstrate the efficiency and effectiveness of our approach.

Organization. The subsequent sections will be described in what follows. First, Section 2 outlines the background and relevant works. Next, Section 3 details the proposed ST-MSC approach, followed by the introduction of the optimization algorithm. Following this line, Section 4 records the experimental results on nine benchmarks, and Section 5 give a conclusion of the proposed work.

2. Related work

In this section, we briefly review the related works on two topics, namely, multi-view subspace clustering and large-scale multi-view clustering.

2.1. Multi-view subspace clustering

During the past decade, subspace learning has been extensively applied in several applications, such as the anomaly detection [11–14] and clustering analysis scenarios. Inheriting from the subspace learning framework, multi-view subspace clustering (MSC) has provided a prominent tool for high-dimensional analysis of multi-view clustering [7–9,15].

In general, many of the existing works stem from the single-view subspace learning models. For example, Gao et al. [16] made an early effort to investigate the MSC problem, which extends the classical LSR model to enhance the consistency across diverse sources. Cao et al. [17] presented a diversity-guided approach for MSC, whose purpose is to exploit the statistical complementary information of various representations via the Hilbert–Schmidt Independence Criterion (HSIC) [18]. To address the issue of view-insufficiency, Zhang et al. [19] assumed that multiple sources share the same latent space. Furthermore, Wang et al. [20] designed a unified MSC approach, which models the complementarity and consistency of multi-view representations into a general scheme. Following this line, Huang et al. [21] proposed a kernelized subspace clustering approach, where two essential tasks, namely multi-view clustering and graph learning, are jointly formulated. Chen et al. [22] developed a multi-view clustering in latent space (MVCLS)

approach, which aims to recover the embedded space from different sources, while simultaneously discovering the global cluster structure.

In addition to these approaches, continual endeavors have been devoted to exploring the deeper (or higher-order) correlation of multi-view representations. Specifically, Xie et al. [23] devised a tensorized MSC approach, where a robust t-SVD tensor constraint is introduced on the rotated tensor. By preserving the locality within the individual view, Xie et al. [24] further leveraged the hyper-Laplacian graph learning technique. They suggested two novel approaches for clustering and semi-supervised learning tasks, namely, HLR-M^{2}VS and semi-HLR-M^{2}VS, respectively. Starting from the local and global viewpoints, Wang et al. [25] recently proposed a deep MSC approach termed DMSC-UDL. Gu et al. [26] sought to formulate the cross-view discrepancy and multi-view high-order correlations for facilitating the downstream clustering task. Although many subspace-based approaches have been designed, most of them still suffer from expensive computational costs, which restrict their capability to handle massive-scale or even medium-scale datasets.

2.2. Large-scale multi-view clustering

Over the past years, anchor-based multi-view clustering (AMC) has demonstrated remarkable advantages in handling complex large-scale problems, upon which a variety of AMC representatives are proposed using different strategies [27–29].

Specifically, Li et al. [30] designed a scalable AMC approach by leveraging the bipartite graph learning technique. After that, Li et al. [5] proposed a parameter-free and scalable approach that focuses on fusing the heterogeneous knowledge among multiple bipartite graphs. Using the anchor representation building strategy, Kang et al. [31] further devised a novel MSC approach termed LMVSC for scalable multi-view clustering. Following this work, Kang et al. [32] developed another scalable MSC approach by utilizing the connectivity of multiple bipartite graphs. In additional, Sun et al. [33] sought to incorporate the unified anchor learning with scalable multi-view subspace learning. Further, Wang et al. [34] proposed a novel large-scale subspace-based approach, in which the low-rank anchor graph is built without any hyper-parameter tuning. Guided by the tri-factorization based NMF framework, Yang et al. [35] designed a highly-efficient clustering approach to uniform the latent distributions among diverse views. In terms of view-independent perspective, Liu et al. [7] developed a fast subspace-based approach that extracts the common block-diagonal relationship among various independent anchor graphs. Moreover, Wan et al. [36] proposed a one-step framework for large-scale clustering, where multiple diverse representations are projected to achieve the consensus clustering labels. In order to simultaneously model the individuality and the commonality latent in multiple views, Gu et al. [37] derived a robust multi-view graph learning approach with sparse feature selection and discriminative subspace extraction techniques. More recently, Chen et al. [38] sought to capture the semantic correlations between anchor sets and cluster centers. To achieve this goal, the cluster centers can be formulated as the linear combinations of the anchor representatives, which provides a new insight for large-scale problem. Beyond the above subspace-based approaches, Liang et al. [39] developed a novel weighted sample-level graph learning approach to address the incomplete multi-view clustering problem.

Regardless of the remarkable progress, these previous works are still faced with two critical problems, namely, how to enhance the high-order complementarity of heterogeneous anchor graphs, and how to bridge the gap between multi-view dimension reduction and anchor multi-view subspace clustering to improve large-scale clustering robustness.

3. Our ST-MSC approach

This section details the proposed approach in what follows. The main notations and definitions are first provided in Section 3.1. After that, the target formulation of our proposed approach is given in Section 3.2. Then a fast optimization algorithm with three steps is presented in Section 3.3. Moving forward, the time and space complexity of the proposed algorithm is analyzed in Section 3.4. Finally, the convergence property of the optimization algorithm is analyzed in Section 3.5.

3.1. Main notations

Throughout the rest of this paper, we represent the 3-mode tensors in calligraphy letters (like $\mathcal{Q}$) and the matrices in upper-case letters (like P). Notice that, the jth column, lth row and jlth element of matrix Q are denoted by $\mathrm{Q}_{:j}$, $\mathrm{Q}_{l:}$ and Q_{jl}, respectively. Additionally, the jth lateral slice, lth horizontal slice and mth frontal slice of 3-mode tensor $\mathcal{Q} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$ are respectively represented as $\mathcal{Q}(:, j, :)$, $\mathcal{Q}(l, :, :)$ and $\mathcal{Q}(:, :, m)$. The transform operator $\overline{\mathcal{Q}} = \mathrm{fft}(\mathcal{Q}, [], 3)$ means the DFT along the mth dimension w.r.t. tensor $\mathcal{Q}$, meanwhile the inverse DFT w.r.t. tensor $\mathcal{Q}$ can be represented by $\mathcal{Q} = \mathrm{ifft}(\overline{\mathcal{Q}}, [], 3)$ here. For simplicity, matrix $\mathrm{Q}^{(m)}$ is utilized to represent the frontal slice $\mathcal{Q}(:, :, m)$.

Moving forward, we give several basic definitions wr.t. the tensor SVD operation. First of all, we have the following "circulant" matrix w.r.t. tensor $\mathcal{Q} \in \mathcal{R}^{n_1 n_3 \times n_2 n_3}$.

$$\mathrm{bcirc}(\mathcal{Q}) = \begin{bmatrix} \mathrm{Q}^{(1)} & \mathrm{Q}^{(n_3)} & \cdots & \mathrm{Q}^{(2)} \\ \mathrm{Q}^{(2)} & \mathrm{Q}^{(1)} & \cdots & \mathrm{Q}^{(3)} \\ \vdots & \ddots & \ddots & \vdots \\ \mathrm{Q}^{(n_3)} & \mathrm{Q}^{(n_3-1)} & \cdots & \mathrm{Q}^{(1)} \end{bmatrix}.$$

Further, the "unfold" and "fold" operations of tensor $\mathcal{Q}$ is given as follows.

$$\mathrm{unfold}(\mathcal{Q}) = \begin{bmatrix} \mathrm{Q}^{(1)} \\ \mathrm{Q}^{(2)} \\ \vdots \\ \mathrm{Q}^{(n_3)} \end{bmatrix}, \ \mathrm{fold}(\mathrm{unfold}(\mathcal{Q})) = \mathcal{Q}.$$

For a tensor $\mathcal{Q}$, its "bdiag" and the inverse operator "bdfold" is given in following.

$$\mathrm{bdiag}(\mathcal{Q}) = \begin{bmatrix} \mathrm{Q}^{(1)} & & & \\ & \mathrm{Q}^{(2)} & & \\ & & \ddots & \\ & & & \mathrm{Q}^{(n_3)} \end{bmatrix}, \ \mathrm{bdfold}(\mathrm{bdaig}(\mathcal{Q})) = \mathcal{Q}.$$

Definition 1 (*Tensor Product [40,41]*)**.** Given any two tensors $\mathcal{A} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$ and $\mathcal{B} \in \mathbb{R}^{n_2 \times n_4 \times n_3}$, the corresponding t-product $\mathcal{C} = \mathcal{A} * \mathcal{B}$ is a tensor of size $n_1 \times n_4 \times n_3$,

$$\mathcal{C} = \mathcal{A} * \mathcal{B} = \mathrm{fold}(\mathrm{bcirc}(\mathcal{A}) \cdot \mathrm{unfold}(\mathcal{B})). \tag{1}$$

Definition 2 (*Tensor SVD [40,41]*)**.** Suppose a tensor $\mathcal{P} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$ is given, we can denote its tensor SVD (t-SVD) in the following.

$$\mathcal{Q} = \mathcal{A} * \mathcal{O} * \mathcal{B}^T, \tag{2}$$

Here $\mathcal{A} \in \mathbb{R}^{n_1 \times n_1 \times n_3}$ and $\mathcal{B} \in \mathbb{R}^{n_2 \times n_2 \times n_3}$ represent two orthogonal tensors, and $\mathcal{O} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$ represents a f-diagonal tensor.

Definition 3 (*t-SVD Based Nuclear Norm [40,41]*)**.** Suppose a tensor $\mathcal{Q} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$ is given, we can denote its t-SVD based nuclear norm (i.e., $\|\mathcal{Q}\|_{\circledast}$) as follows:

$$\|\mathcal{Q}\|_{\circledast} = \sum_{k=1}^{n_3} \|\mathcal{Q}_f^{(k)}\|_* = \sum_{i=1}^{min(n_1,n_2)} \sum_{k=1}^{n_3} | \, O_f^{(k)}(i,i) \, | \tag{3}$$

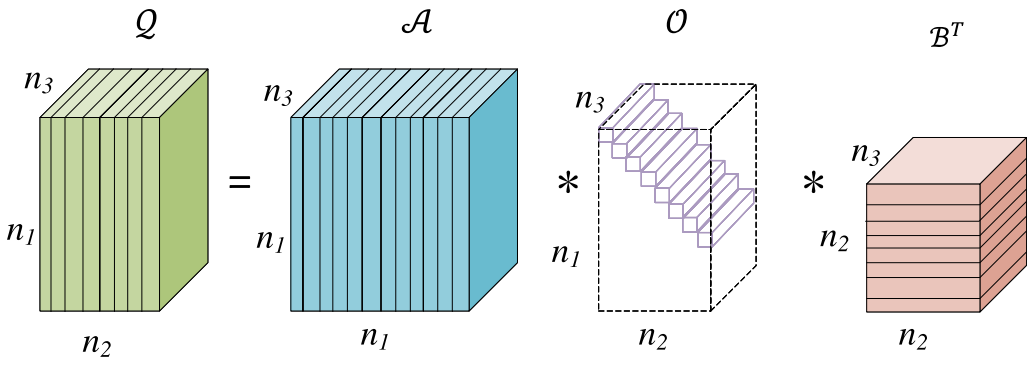


Fig. 2. The t-SVD decomposition for a tensor $\mathcal{Q} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$.

Specifically, Fig. 2 shows an example of the t-SVD decomposition for a tensor $\mathcal{Q} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$.

3.2. The target formulation

Before elaborating on our target formulation, we delve into some preliminary knowledge. Formally, a multi-view dataset consisting of V sources can be denoted as $\mathrm{X}_d = [\mathrm{X}^{(1)}; \ldots; \mathrm{X}^{(V)}]$, among which $\mathrm{X}^{(v)} \in \mathbb{R}^{d_v \times n}$ indicates the data sub-matrix of the vth view. Current MSC research is based on the self-expressive learning theory, which treats the whole set of data samples as the self-supervision dictionary. Starting from the self-expressive assumption, each individual data sample lies in a series of embedded subspaces within the original feature spaces. Therefore, we derive the vth view self-expressive formulation in what follows:

$$\mathrm{X}^{(v)} \approx \mathrm{X}^{(v)}\mathrm{Z}^{(v)}, \forall v, \tag{4}$$

in which $\mathrm{Z}^{(v)} \in \mathbb{R}^{n \times n}$ indicates the subspace representation of the vth view. Based on this, the basic model for multi-view subspace clustering can be formulated as follows:

$$\min_{\{\mathrm{Z}^{(v)}\}_{v=1}^{V}} \sum_{v=1}^{V} \left\| \mathrm{X}^{(v)} - \mathrm{X}^{(v)}\mathrm{Z}^{(v)} \right\|_F^2 + \sum_{v=1}^{V} \Omega(\mathrm{Z}^{(v)}) \tag{5}$$

in which $\Omega(.)$ indicates a specific regularization term for $\{\mathrm{Z}^{(v)}\}_{v=1}^{V}$. As shown in Eq. (5), the conventional MSC approaches aim to explore the global subspace relationship across different sources. However, the time complexity (over $\mathrm{O}(n^3)$) restricts these MSC approaches from handling massive-scale or even medium-scale datasets.

In light of this, some interests have been focused on the context of anchor multi-view subspace clustering (also abbreviation for anchor subspace-based approach).

Instead of pursuing the subspace representations (with a size of $n \times n$) from multiple sources, anchor subspace-based approaches concentrate on constructing several bipartite graphs (also known as anchor graphs) with a much smaller size (i.e., $p \times n$). The core of anchor subspace-based approaches is to exploit the local relationship between the data samples and a set of selected anchors. Specifically, each data sample can be linearly represented by a small set of anchors, thus leading to the formulation as follows:

$$\mathrm{X}^{(v)} \approx \mathrm{A}^{(v)}\mathrm{B}^{(v)}, \forall v, \tag{6}$$

Especially, $\mathrm{A}^{(v)}$ means the anchor set of the vth view, while $\mathrm{B}^{(v)}$ means the anchor representation in the same source. Within each individual source, the anchor set is usually generated by applying k-means algorithm or random sampling strategy. Thereby, the target formulation for anchor MSC can be set as:

$$\min_{\{\mathrm{A}^{(v)},\mathrm{B}^{(v)}\}_{v=1}^{V}} \sum_{v=1}^{V} \left\| \mathrm{X}^{(v)} - \mathrm{A}^{(v)}\mathrm{B}^{(v)} \right\|_F^2 + \sum_{v=1}^{V} \Theta(\mathrm{B}^{(v)}) \tag{7}$$

in which $\Theta(.)$ is a specific regularization term for $\{\mathrm{B}^{(v)}\}_{v=1}^{V}$. In particular, this function (i.e., Eq. (7)) can efficiently accelerate the computational speed by utilizing the lightweight structure of bipartite graph. Based on this, we expect to further enhance its scalability from deeper viewpoint.

Then we investigate the target formulation of the tri-factorization scheme. Aiming to capture the latent semantic embedded in original feature spaces, we can factorize the anchor matrix $\{\mathrm{A}^{(v)}\}_{v=1}^{V}$ into two components, that is

$$\mathrm{A}^{(v)} \approx \mathrm{P}^{(v)}\mathrm{F}^{(v)}, \text{ s.t. } (\mathrm{P}^{(v)^T})\mathrm{P}^{(v)} = \mathrm{I}, \forall v, \tag{8}$$

Here $\mathrm{P}^{(v)}$ indicates the vth view base matrix while $\mathrm{F}^{(v)}$ indicates the anchor set of the vth view in the embedded spaces. When integrating Eqs. (6) and (8) together, the tri-factorization formulation w.r.t. the vth view data matrix is given in the following:

$$\mathrm{X}^{(v)} \approx \mathrm{P}^{(v)}\mathrm{F}^{(v)}\mathrm{B}^{(v)}, \text{ s.t. } (\mathrm{P}^{(v)^T})\mathrm{P}^{(v)} = \mathrm{I}, \forall v, \tag{9}$$

Derived from the tri-factorization scheme, the target function of our ST-MSC approach is expressed in what follows:

$$\begin{aligned} \min_{\{\mathrm{P}^{(v)},\mathrm{A}^{(v)},\mathrm{B}^{(v)}\}_{v=1}^{V}} & \sum_{v=1}^{V} \left\| \mathrm{X}^{(v)} - \mathrm{P}^{(v)}\mathrm{F}^{(v)}\mathrm{B}^{(v)} \right\|_F^2 + \sum_{v=1}^{V} \Omega(\mathrm{B}^{(v)}) \\ \text{s.t. } & (\mathrm{P}^{(v)^T})\mathrm{P}^{(v)} = \mathrm{I}. \end{aligned} \tag{10}$$

The column orthogonal constraint on $\{\mathrm{P}^{(v)}\}_{v=1}^{V}$ is incorporated for preserving the independence among different bases in embedded spaces. Inspired by the previous tensor clustering approaches [42], we seek to capture the high-order correlation latent in different bipartite graphs (with each bipartite graph corresponding to one view). Mathematically, by imposing the t-SVD tensor norm on anchor representations, we can obtain the target formulation of our ST-MSC as follows:

$$\begin{aligned} \min_{\{\mathrm{P}^{(v)},\mathrm{F}^{(v)},\mathrm{B}^{(v)}\}_{v=1}^{V}} & \sum_{v=1}^{V} \left\| \mathrm{X}^{(v)} - \mathrm{P}^{(v)}\mathrm{F}^{(v)}\mathrm{B}^{(v)} \right\|_F^2 + \lambda \|\mathcal{B}\|_{\circledast} \\ \text{s.t. } & (\mathrm{P}^{(v)^T})\mathrm{P}^{(v)} = \mathrm{I}, \ \mathcal{B} = \Psi(\mathrm{B}^{(1)}, \mathrm{B}^{(2)}, \ldots, \mathrm{B}^{(V)}). \end{aligned} \tag{11}$$

As indicated in Fig. 3, operator $\Psi(\cdot)$ here heaps the anchor representations $\{\mathrm{B}^{(v)}\}_{v=1}^{V}$ into a tensor $\mathcal{B}$ with 3-mode, and then rotates its size to $n \times V \times p$. With the t-SVD-based tensor constraint $\|\mathcal{B}\|_{\circledast}$ given, our proposed model is expected to give a novel insight into multi-view anchor-sample relationship, which efficiently characterizes the low-rank spatial information embedded in heterogeneous views.

Specifically, we proceed to analyze some inherent properties of the proposed tri-factorization scheme. We first provide an important lemma in the following.

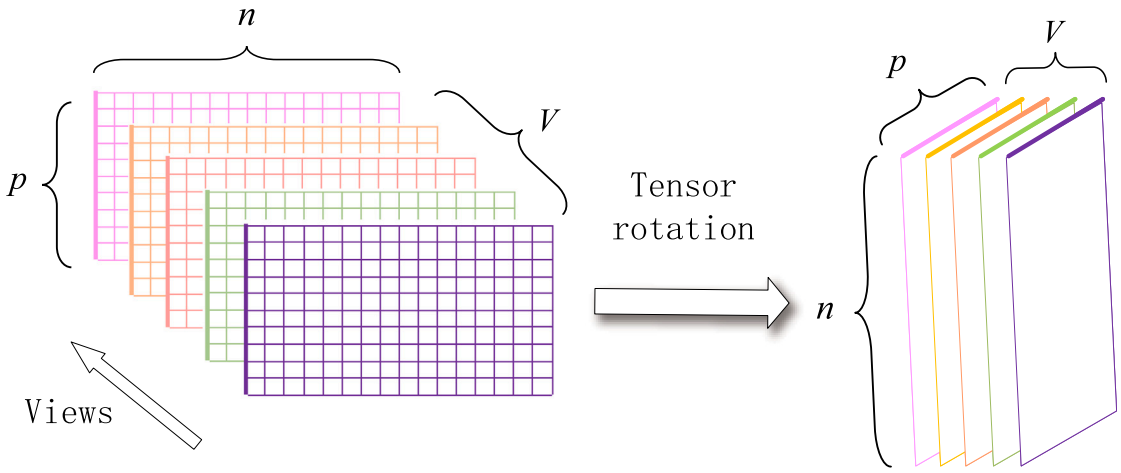


Fig. 3. Example of rotation operation.

Lemma 1. *With* $(\mathrm{P}^{(v)^T})\mathrm{P}^{(v)} = \mathrm{I}, \forall v$ *given, the reconstruction term in Eq. (11) can be factorized into two components, corresponding to multi-view dimension reduction and anchor multi-view subspace clustering, respectively. Then we have*

$$\begin{aligned}&\sum_{v=1}^{V}\left\|\mathrm{X}^{(v)}-\mathrm{P}^{(v)}(\mathrm{F}^{(v)})(\mathrm{B}^{(v)})\right\|_F^2\\&=\sum_{v=1}^{V}\left\|\mathrm{X}^{(v)}-\mathrm{P}^{(v)}(\mathrm{P}^{(v)})^T\mathrm{X}^{(v)}\right\|_F^2+\sum_{v=1}^{V}\left\|(\mathrm{P}^{(v)})^T\mathrm{X}^{(v)}-\mathrm{F}^{(v)}\mathrm{B}^{(v)}\right\|_F^2\end{aligned}\tag{12}$$

Proof. In the first place, the left-hand side of Eq. (12) can be rewritten as

$$\begin{aligned}&\sum_{v=1}^{V}\left\|\mathrm{X}^{(v)}-\mathrm{P}^{(v)}\mathrm{F}^{(v)}\mathrm{B}^{(v)}\right\|_F^2\\&=\sum_{v=1}^{V}\left\|\mathrm{X}^{(v)}\right\|_F^2+\sum_{v=1}^{V}\left\|\mathrm{P}^{(v)}\mathrm{F}^{(v)}\mathrm{B}^{(v)}\right\|_F^2-2\sum_{v=1}^{V}\mathrm{Tr}\{(\mathrm{X}^{(v)})^T\mathrm{P}^{(v)}\mathrm{F}^{(v)}\mathrm{B}^{(v)}\}\end{aligned}\tag{13}$$

With the condition $(\mathrm{P}^{(v)^T})\mathrm{P}^{(v)} = \mathrm{I}, \forall v$, the first term on the right-hand side of Eq. (12) can be rewritten as

$$\begin{aligned}&\sum_{v=1}^{V}\left\|\mathrm{X}^{(v)}-\mathrm{P}^{(v)}(\mathrm{P}^{(v)})^T\mathrm{X}^{(v)}\right\|_F^2\\&=\sum_{v=1}^{V}\left\|\mathrm{X}^{(v)}\right\|_F^2+\sum_{v=1}^{V}\left\|\mathrm{P}^{(v)}(\mathrm{P}^{(v)})^T\mathrm{X}^{(v)}\right\|_F^2-2\sum_{v=1}^{V}\mathrm{Tr}\{(\mathrm{X}^{(v)})^T\mathrm{P}^{(v)}(\mathrm{P}^{(v)})^T\mathrm{X}^{(v)}\}\\&=\sum_{v=1}^{V}\left\|\mathrm{X}^{(v)}\right\|_F^2-\sum_{v=1}^{V}\left\|(\mathrm{P}^{(v)})^T\mathrm{X}^{(v)}\right\|_F^2\end{aligned}\tag{14}$$

With $\sum_{v=1}^{V}\left\|\mathrm{F}^{(v)}\mathrm{B}^{(v)}\right\|_F^2=\sum_{v=1}^{V}\left\|\mathrm{P}^{(v)}\mathrm{F}^{(v)}\mathrm{B}^{(v)}\right\|_F^2$, we can rewrite the last term of Eq. (12) as

$$\begin{aligned}&\sum_{v=1}^{V}\left\|(\mathrm{P}^{(v)})^T\mathrm{X}^{(v)}-\mathrm{F}^{(v)}\mathrm{B}^{(v)}\right\|_F^2\\&=\sum_{v=1}^{V}\left\|(\mathrm{P}^{(v)})^T\mathrm{X}^{(v)}\right\|_F^2+\sum_{v=1}^{V}\left\|\mathrm{F}^{(v)}\mathrm{B}^{(v)}\right\|_F^2-2\sum_{v=1}^{V}\mathrm{Tr}\{(\mathrm{X}^{(v)})^T\mathrm{P}^{(v)}\mathrm{F}^{(v)}\mathrm{B}^{(v)}\}\\&=\sum_{v=1}^{V}\left\|(\mathrm{P}^{(v)})^T\mathrm{X}^{(v)}\right\|_F^2+\sum_{v=1}^{V}\left\|\mathrm{P}^{(v)}\mathrm{F}^{(v)}\mathrm{B}^{(v)}\right\|_F^2-2\sum_{v=1}^{V}\mathrm{Tr}\{(\mathrm{X}^{(v)})^T\mathrm{P}^{(v)}\mathrm{F}^{(v)}\mathrm{B}^{(v)}\}\end{aligned}\tag{15}$$

Obviously, the sum of the right-hand sides of Eqs. (14) and (15) equals the right-hand side of Eq. (13) □

Inheriting from Lemma 1, we can further rewrite the proposed target formulation into the following form:

$$\begin{aligned}\min_{\{\mathrm{P}^{(v)},\mathrm{F}^{(v)},\mathrm{B}^{(v)}\}_{v=1}^{V}}\ &\underbrace{\sum_{v=1}^{V}\left\|\mathrm{X}^{(v)}-\mathrm{P}^{(v)}(\mathrm{P}^{(v)})^T\mathrm{X}^{(v)}\right\|_F^2}_{\text{Multi-view dimension reduction}}\\&+\underbrace{\sum_{v=1}^{V}\left\|(\mathrm{P}^{(v)})^T\mathrm{X}^{(v)}-\mathrm{F}^{(v)}\mathrm{B}^{(v)}\right\|_F^2+\lambda\|\mathcal{B}\|_{\circledast}}_{\text{Tensorized anchor multi-view subspace clustering}}\\\text{s.t. }\ &(\mathrm{P}^{(v)^T})\mathrm{P}^{(v)}=\mathrm{I},\ \mathcal{B}=\Psi(\mathrm{B}^{(1)},\mathrm{B}^{(2)},\ldots,\mathrm{B}^{(V)}).\end{aligned}\tag{16}$$

in which λ is a positive hyper-parameter.

For further discussions and analysis, we give several remarks as follows.

Remark 1. The target formulation has two parts. The first part (i.e., *the multi-view dimension reduction part*) seeks to project the original features into multiple embedded spaces. The second part focuses on the *tensorized anchor multi-view subspace clustering*, which helps to exploit the deep relationship (i.e., high-order complementarity) between the samples and the selected anchors. Inheriting from the tri-factorization scheme, these two parts can benefit each other during iterations.

Remark 2. In the tri-factorization scheme, minimizing the first part (i.e., *the multi-view dimension reduction*) is equivalent to the multi-view principal component analysis (PCA) term.

Proof.

$$\begin{aligned}&\min_{(\mathrm{P}^{(v)^T})\mathrm{P}^{(v)}=\mathrm{I}}\sum_{v=1}^{V}\left\|\mathrm{X}^{(v)}-\mathrm{P}^{(v)}(\mathrm{P}^{(v)})^T\mathrm{X}^{(v)}\right\|_F^2\\\Leftrightarrow\ &\min_{(\mathrm{P}^{(v)^T})\mathrm{P}^{(v)}=\mathrm{I}}\{\sum_{v=1}^{V}\left\|\mathrm{P}^{(v)}(\mathrm{P}^{(v)})^T\mathrm{X}^{(v)}\right\|_F^2-2\sum_{v=1}^{V}\mathrm{Tr}\{(\mathrm{X}^{(v)})^T\mathrm{P}^{(v)}(\mathrm{P}^{(v)})^T\mathrm{X}^{(v)}\}\}\\\Leftrightarrow\ &\max_{(\mathrm{P}^{(v)^T})\mathrm{P}^{(v)}=\mathrm{I}}\sum_{v=1}^{V}\mathrm{Tr}\{(\mathrm{P}^{(v)})^T\mathrm{X}^{(v)}(\mathrm{X}^{(v)})^T\mathrm{P}^{(v)}\}\ \square\end{aligned}\tag{17}$$

With the development of clustering research, several fields like text mining, recommender systems and genetic data analysis have gradually turned their attention towards the tri-factorization framework, among which the representatives include the co-clustering technique [43] and concept factorization technique [44]. In the subsequent remarks, we will discuss the potential relationships among co-clustering, concept factorization, and the proposed approach.

Remark 3. As one of the tri-factorization representatives in the literature, the co-clustering (also known as biclustering) technique [43] seeks to simultaneously perform partitioning at two perspectives (both rows and columns of the data matrix). Inheriting from this, we can factorize the data matrix X into the following form:

$$\mathrm{X} \approx \mathrm{PHG}, \ \text{s.t.} \ \mathrm{P}, \mathrm{G} \in \Delta \tag{18}$$

where variables P and G denote the clustering partitions of features and samples of data matrix X. Notably, Δ means the hard clustering constraint on variables P and G (i.e., $\mathrm{P} \geq 0, \mathrm{P}^T\mathrm{P} = \mathrm{I}$ and $\mathrm{G} \geq 0, \mathrm{G}^T\mathrm{G} = \mathrm{I}$). Hence, the general model of co-clustering can be formulated in the subsequent equation:

$$\min_{\mathrm{P,H,G}} \|\mathrm{X} - \mathrm{PHG}\|_F^2, \ \text{s.t.} \ \mathrm{P}, \mathrm{G} \in \Delta \tag{19}$$

By extending the above general model from single-view to multiple views, we can obtain the following equation:

$$\min_{\mathrm{P}^{(v)},\mathrm{H}^{(v)},\mathrm{G}} \sum_{v=1}^{V} \|\mathrm{X}^{(v)} - \mathrm{P}^{(v)}\mathrm{H}^{(v)}\mathrm{G}\|_F^2, \ \text{s.t.} \ \mathrm{P}^{(v)}, \mathrm{G} \in \Delta, \forall v \tag{20}$$

where $\mathrm{P}^{(v)}$ indicates the clustering partition of the features of the vth view, which is actually a feature selection matrix of the vth view. Additionally, G means the consistent clustering partition across multiple views. According to the observation in Lemma 1, we can rewrite the above objective model as follows:

$$\begin{aligned}\min_{\{\mathrm{P}^{(v)},\mathrm{H}^{(v)},\mathrm{G}\}_{v=1}^{V}} & \underbrace{\sum_{v=1}^{V} \left\|\mathrm{X}^{(v)} - \mathrm{P}^{(v)}(\mathrm{P}^{(v)})^T\mathrm{X}^{(v)}\right\|_F^2}_{\text{Multi-view feature selection}} \\ & + \underbrace{\sum_{v=1}^{V} \left\|(\mathrm{P}^{(v)})^T\mathrm{X}^{(v)} - \mathrm{H}^{(v)}\mathrm{G}\right\|_F^2}_{\text{Projectived multi-view } k\text{-means clustering}} \\ \text{s.t.} \ & \mathrm{P}^{(v)}, \mathrm{G} \in \Delta. \end{aligned} \tag{21}$$

On the basis of the aforementioned analysis, the multi-view co-clustering aims to unify the multi-view k-means with multi-view feature selection. In conclusion, the proposed tri-factorization approach for scalable multi-view subspace clustering can be viewed as a special extension of the multi-view co-clustering scheme.

Remark 4. Besides co-clustering, another representative technique in the tri-factorization clustering research is the concept factorization [44]. In particular, the classical k-means method aims to factorize the data matrix X into two parts, which can be formulated into the following form:

$$\mathrm{X} \approx \mathrm{CG}, \ \text{s.t.} \ \mathrm{G} \in \Delta \tag{22}$$

where matrices C and G denote the cluster centers and clustering partition, respectively. Similarity, Δ means the hard clustering constraint on G (i.e., $\mathrm{G} \geq 0, \mathrm{G}^T\mathrm{G} = \mathrm{I}$). As a variant of k-means clustering, the concept factorization technique further models the cluster centers as a linear combination of data points (i.e., $\mathrm{C} = \mathrm{XW}$), which inspires the following formulation:

$$\mathrm{X} \approx \mathrm{XWG}, \ \text{s.t.} \ \mathrm{W} \geq 0, \ \mathrm{G} \in \Delta \tag{23}$$

where matrix W represents the coefficient matrix. Accordingly, the general model of concept factorization can be formulated in the subsequent equation:

$$\min_{\mathrm{W,G}} \|\mathrm{X} - \mathrm{XWG}\|_F^2, \ \text{s.t.} \ \mathrm{W} \geq 0, \ \mathrm{G} \in \Delta \tag{24}$$

Furthermore, if we focus on the multi-view scenario, then we can have the following equation:

$$\min_{\mathrm{W}^{(v)},\mathrm{G}} \sum_{v=1}^{V} \|\mathrm{X}^{(v)} - \mathrm{X}^{(v)}\mathrm{W}^{(v)}\mathrm{G}\|_F^2, \ \text{s.t.} \ \mathrm{W}^{(v)} \geq 0, \ \mathrm{G} \in \Delta, \forall v, \tag{25}$$

where $\mathrm{W}^{(v)}$ is the vth view coefficient matrix, and G is the consistent clustering partition of all views. In contrast to the proposed approach in Eq. (16), the concept factorization technique aims to capture the common cluster structure across diverse views, while at the same time discovering the semantic correlation across cluster centers and data samples.

Remark 5. Attributed graph clustering (AGC) [45] has gained considerable attention in recent years, which draws inspiration primary from deep learning and matrix factorization. In the subsequent descriptions, we will discuss the potential connections between the ST-MSC approach with graph conventional networks (GCN) and the AGC problem.

Let a multi-view attributed graph dataset with V views be denoted as $\mathcal{G} = (\mathcal{G}^{(1)}, \mathcal{G}^{(2)}, \ldots, \mathcal{G}^{(V)})$, where $\mathcal{G}^{(v)} = (\mathrm{X}^{(v)}, \mathrm{A}^{(v)})$ is the vth view of the attributed graph data, $\mathrm{X}^{(v)}$ is the feature matrix of the vth view, and $\mathrm{A}^{(v)}$ is the affinity matrix of the vth view. For a specific view, say, the vth view, with the modified affinity matrix $\widetilde{\mathrm{A}}^{(v)} = \mathrm{A}^{(v)} + \mathrm{I}$ given, the corresponding normalized affinity matrix can be represented as $\mathrm{W}^{(v)} = (\widetilde{\mathrm{D}}^{(v)})^{-1}\widetilde{\mathrm{A}}^{(v)}$. In order to aggregate the neighborhood information hidden in the attributed graph structure, we can directly extend our ST-MSC model by leveraging the k-order low-pass filter. The corresponding new model can be formulated as follows:

$$\begin{aligned}\min_{\{\mathrm{P}^{(v)},\mathrm{A}^{(v)},\mathrm{B}^{(v)}\}_{v=1}^{V}} & \sum_{v=1}^{V} \left\|\mathrm{X}^{(v)}(\mathrm{W}^{(v)})^k - \mathrm{P}^{(v)}\mathrm{F}^{(v)}\mathrm{B}^{(v)}\right\|_F^2 + \lambda \sum_{v=1}^{V} \Omega(\mathrm{B}^{(v)}) \\ \text{s.t.} \ & (\mathrm{P}^{(v)^T})\mathrm{P}^{(v)} = \mathrm{I}, \end{aligned} \tag{26}$$

Here $\mathrm{X}^{(v)}(\mathrm{W}^{(v)})^k$ can be viewed as a smooth representation in the vth view. According to Lemma 1, the above model can be formulated into two parts:

$$\begin{aligned}\min_{\{\mathrm{P}^{(v)},\mathrm{F}^{(v)},\mathrm{B}^{(v)}\}_{v=1}^{V}} & \underbrace{\sum_{v=1}^{V} \left\|\mathrm{X}^{(v)}(\mathrm{W}^{(v)})^k - \mathrm{P}^{(v)}(\mathrm{P}^{(v)})^T\mathrm{X}^{(v)}(\mathrm{W}^{(v)})^k\right\|_F^2}_{\text{Multi-view reconstruction}} \\ & + \underbrace{\sum_{v=1}^{V} \left\|(\mathrm{P}^{(v)})^T\mathrm{X}^{(v)}(\mathrm{W}^{(v)})^k - \mathrm{F}^{(v)}\mathrm{B}^{(v)}\right\|_F^2 + \lambda \parallel \sum_{v=1}^{V} \Omega(\mathrm{B}^{(v)})}_{\text{Multi-view graph subspace clustering}} \\ \text{s.t.} \ & (\mathrm{P}^{(v)^T})\mathrm{P}^{(v)} = \mathrm{I}. \end{aligned} \tag{27}$$

From the deep learning perspective, the first term favors the multi-view data reconstruction of k-hop aggregated matrix $\mathrm{X}^{(v)}(\mathrm{W}^{(v)})^k$, which can also be viewed as a linear graph autoenconder term. Meanwhile, the second term favors the multi-view graph subspace clustering network, whose aim is to perform scalable subspace clustering in the latent embedding space. Following the studies in deep clustering and attributed graph clustering, the potential extensions of our ST-MSC approach might be further explored in the future work.

3.3. Optimization algorithm

The alternating direction method of multipliers (ADMM) [46] is a widely-adopted algorithm for solving constrained optimization problems. In this section, based on ADMM, we design a fast minimization algorithm to optimize the objective function of ST-MSC. Let $\mathcal{C}$ denote the auxiliary variable and $\mathcal{Y}$ denote the Lagrange multiplier, we have the augmented Lagrangian function as follows:

$$\begin{aligned} & \mathcal{L}\left(\mathcal{C}, \{\mathrm{B}^{(v)}\}_{v=1}^{V}, \{\mathrm{P}^{(v)}\}_{v=1}^{V}, \{\mathrm{F}^{(v)}\}_{v=1}^{V}, \mathcal{Y}\right) \\ & = \sum_{v=1}^{V} \left\|\mathrm{X}^{(v)} - \mathrm{P}^{(v)}\mathrm{F}^{(v)}\mathrm{B}^{(v)}\right\|_F^2 + \lambda\|\mathcal{C}\|_{\circledast} \\ & \quad + (\mathcal{Y}, \mathcal{B} - \mathcal{C}) + \frac{\mu}{2}\|\mathcal{B} - \mathcal{C}\|_F^2 \\ \text{s.t.} \ & (\mathrm{P}^{(v)^T})\mathrm{P}^{(v)} = \mathrm{I}, \ \mathcal{B} = \Psi(\mathrm{B}^{(1)}, \mathrm{B}^{(2)}, \ldots, \mathrm{B}^{(V)}), \\ & \mathcal{C} = \Psi(\mathrm{C}^{(1)}, \mathrm{C}^{(2)}, \ldots, \mathrm{C}^{(V)}), \ \mathcal{Y} = \Psi(\mathrm{Y}^{(1)}, \mathrm{Y}^{(2)}, \ldots, \mathrm{Y}^{(V)}). \end{aligned} \tag{28}$$

where μ is a positive parameter. Following this, we can design an iterative algorithm to solve the subproblems w.r.t. $\{\mathrm{F}^{(v)}\}_{v=1}^{V}$, $\mathcal{C}$, $\{\mathrm{B}^{(v)}\}_{v=1}^{V}$ and $\{\mathrm{P}^{(v)}\}_{v=1}^{V}$, respectively, by turns.

3.3.1. Solving the $\mathrm{F}^{(v)}$-subproblem

Keeping variables $\{\mathrm{P}^{(v)}\}_{v=1}^{V}$, $\mathcal{C}$ and $\{\mathrm{B}^{(v)}\}_{v=1}^{V}$ as constraints, then the subproblem of Eq. (28) w.r.t. $\mathrm{F}^{(v)}$ becomes following form:

$$\min_{\mathrm{F}^{(v)}} \left\| \mathrm{X}^{(v)} - \mathrm{P}^{(v)}\mathrm{F}^{(v)}\mathrm{B}^{(v)} \right\|_F^2 \tag{29}$$

The above problem can be equivalently rewritten into the following form:

$$\min_{\mathrm{F}^{(v)}} -2\mathrm{Tr}(\mathrm{F}^{(v)}\mathrm{B}^{(v)}\mathrm{X}^{(v)^T}\mathrm{P}^{(v)}) + \mathrm{Tr}(\mathrm{F}^{(v)^T}\mathrm{F}^{(v)}\mathrm{B}^{(v)}\mathrm{B}^{(v)^T}) \tag{30}$$

If we take the derivative of this formulation w.r.t. variable $\mathrm{F}^{(v)}$, then the corresponding solution can be achieved as follows:

$$\mathrm{F}^{(v)} = (\mathrm{P}^{(v)^T}\mathrm{X}^{(v)}\mathrm{B}^{(v)^T})(\mathrm{B}^{(v)}\mathrm{B}^{(v)^T} + \theta\mathrm{I})^{-1} \tag{31}$$

where θ is a very small positive parameter like $1e^{-6}$.

3.3.2. Solving the $\mathcal{C}$-subproblem

Keeping variables $\{\mathrm{F}^{(v)}\}_{v=1}^{V}$, $\{\mathrm{P}^{(v)}\}_{v=1}^{V}$ and $\{\mathrm{B}^{(v)}\}_{v=1}^{V}$ as constraints, then the subproblem of Eq. (28) w.r.t. $\mathcal{C}$ becomes following form:

$$\min_{\mathcal{C}} \lambda\|\mathcal{C}\|_{\circledast} + \frac{\mu}{2}\|\mathcal{C} - \mathcal{Q}\|_F^2 \tag{32}$$

in which $\mathcal{Q} = \mathcal{B} + \frac{\mathcal{Y}}{\mu}$. Guided by the original paper in [24], we aim to address this subproblem via the following tensor operator [47]:

$$\mathcal{D} = \mathcal{C}_{n_3\pi}(\mathcal{Q}) = \mathcal{U} * \mathcal{C}_{n_3\pi}(\mathcal{O}) * \mathcal{V}^T. \tag{33}$$

Note that $\pi = \mu/\lambda$ and $\mathcal{Q} = \mathcal{U} * \mathcal{O} * \mathcal{V}^T$. Additionally, $\mathcal{C}_{n_3}(\mathcal{Q}) = \mathcal{Q} * \mathcal{W}$, where $\mathcal{W}$ is a f-diagonal tensor, while $\mathcal{W}_f(i,i,j) = (1 - \frac{n_3\pi}{\mathcal{O}_f^{(j)}(i,i)})_+$ indicates the diagonal element of $\mathcal{W}$ in the Fourier domain.

3.3.3. Solving the $\mathrm{B}^{(v)}$-subproblem

Keep variables $\{\mathrm{F}^{(v)}\}_{v=1}^{V}$, $\{\mathrm{P}^{(v)}\}_{v=1}^{V}$ and $\mathcal{C}$ as constraints, then the subproblem of Eq. (28) w.r.t. $\mathrm{B}^{(v)}$ becomes following form:

$$\min_{\mathrm{B}^{(v)}} \mathrm{Tr}\{\mathrm{B}^{(v)^T}(\mathrm{F}^{(v)^T}\mathrm{F}^{(v)} + \frac{\mu}{2}\mathrm{I})\mathrm{B}^{(v)}\} - 2\mathrm{Tr}\{\mathrm{B}^{(v)^T}\mathrm{M}^{(v)}\} \tag{34}$$

Here $\mathrm{M}^{(v)} = \mathrm{F}^{(v)^T}\mathrm{P}^{(v)^T}\mathrm{X}^{(v)} - \frac{1}{2}\mathrm{Y}^{(v)} + \frac{\mu}{2}\mathrm{C}^{(v)}$. Thus, the optimal solution w.r.t. $\mathrm{B}^{(v)}$ can be given by:

$$\mathrm{B}^{(v)} = (\mathrm{F}^{(v)^T}\mathrm{F}^{(v)} + \frac{\mu}{2}\mathrm{I})^{-1}\mathrm{M}^{(v)}. \tag{35}$$

3.3.4. Solving the $\mathrm{P}^{(v)}$-subproblem

Keep variables $\{\mathrm{F}^{(v)}\}_{v=1}^{V}$, $\mathcal{C}$ and $\{\mathrm{B}^{(v)}\}_{v=1}^{V}$ as constraints, the subproblem of Eq. (28) w.r.t. $\mathrm{P}^{(v)}$ becomes following form:

$$\max_{\mathrm{P}^{(v)}} \mathrm{Tr}(\mathrm{P}^{(v)^T}\mathrm{R}^{(v)}) \quad \text{s.t. } \mathrm{P}^{(v)^T}\mathrm{P}^{(v)} = \mathrm{I} \tag{36}$$

Here $\mathrm{R}^{(v)} = \mathrm{X}^{(v)}\mathrm{B}^{(v)^T}\mathrm{F}^{(v)^T}$. Similarly, the above subproblem can be efficiently solved in what follows:

$$\mathrm{P}^{(v)} = \mathrm{U}_{\mathrm{R}}^{(v)}\mathrm{W}_{\mathrm{R}}^{(v)^T} \tag{37}$$

where we can represent the SVD decomposition of matrix $\mathrm{R}^{(v)}$ as $\mathrm{R}^{(v)} = \mathrm{U}_{\mathrm{R}}^{(v)}\Lambda_{\mathrm{R}}^{(v)}\mathrm{W}_{\mathrm{R}}^{(v)^T}$.

Also, the multiplier $\mathcal{Y}$ can be updated by solving the equation in what follows:

$$\mathcal{Y} = \mathcal{Y} + \mu(\mathcal{B} - \mathcal{C}) \tag{38}$$

For clarity, the overall workflow of our algorithm is provided in Algorithm 1. When the iterations are finished, we can concatenate the multi-view anchor representations into a unified anchor representation $\overline{\mathrm{B}} \in \mathbb{R}^{pV\times n}$, which is formulated as

$$\overline{\mathrm{B}} = \sqrt{V}\begin{bmatrix} \mathrm{B}^{(1)} \\ \mathrm{B}^{(2)} \\ \vdots \\ \mathrm{B}^{(V)} \end{bmatrix} \tag{39}$$

Then the SVD is performed on the unified anchor representation $\overline{\mathrm{B}}$, denoted as $\overline{\mathrm{B}} = \mathrm{SVD}(\mathrm{U}_{\mathrm{B}}\Lambda_{\mathrm{B}}\mathrm{W}_{\mathrm{B}}{}^T)$. After that, we can run the k-means clustering on the obtained matrix W_{B} and thus achieve the final clustering result with k clusters.

Algorithm 1 ST-MSC

Input: Multi-view data with V views $\{\mathrm{X}^{(v)}\}_{v=1}^{V}$, the maximum number of iterations it_{max}, the number of selected anchors α and the balancing hyper-parameter $\lambda > 0$.

Parameter Setup: Set $\mu = 10^3$, $\rho = 1$, $\eta = 1.3$, $\mu_{max} = = 10^6$ and $\varepsilon = 10^{-4}$.

Initialization: Initialize variables $\mathcal{Y}$, $\mathrm{P}^{(v)}, \forall v$ and $\mathrm{B}^{(v)}, \forall v$. Set $it = 1$.
repeat
  Obtain $\{\mathrm{F}^{(v)}\}_{v=1}^{V}$ by Eq. (31).
  Obtain $\mathcal{C}$ by Eq. (33).
  Obtain $\{\mathrm{B}^{(v)}\}_{v=1}^{V}$ by Eq. (35).
  Obtain $\{\mathrm{P}^{(v)}\}_{v=1}^{V}$ by Eq. (37).
  Obtain $\mathcal{Y}$ by Eq. (38).
  Update μ by $\mu = \min(\eta\mu, \mu_{\max})$.
  $it = it + 1$.
until The convergence conditions are met **OR** $it > it_{max}$.
Output: $\{\mathrm{F}^{(v)}\}_{v=1}^{V}$, $\{\mathrm{P}^{(v)}\}_{v=1}^{V}$, $\mathcal{C}$ and $\{\mathrm{B}^{(v)}\}_{v=1}^{V}$.

3.4. Time and space complexity analysis

In this section, the time complexity and the space complexity of our ST-MSC approach is provided, as described in the following Sections 3.4.1 and 3.4.2, respectively.

3.4.1. Time complexity

The Algorithm 1 involves four iterative steps during each iteration. In the first place, the update for $\{\mathrm{F}^{(v)}\}_{v=1}^{V}$ requires $\mathrm{O}\{\sum_{v=1}^{V}(\gamma_v d_v n + \gamma_v np + \gamma_v p^2)\}$ time cost, with V, n, p, d_v, r_v representing the number of views, data samples, anchors, dimensions of the vth view and the reduced dimensions of the vth embedded space, respectively. In addition, the update for $\{\mathrm{P}^{(v)}\}_{v=1}^{V}$ and $\mathcal{C}$ require $\mathrm{O}\{\sum_{v=1}^{V}(d_v np + d_v p\gamma_v + d_v\gamma_v{}^2)\}$ and $\mathrm{O}(Vnp\log(Vn) + V^2np)$ time cost, respectively. Also, the update for $\{\mathrm{B}^{(v)}\}_{v=1}^{V}$ requires $\mathrm{O}(p\gamma_v d_v + pd_v n + Vnp^2)$ time cost. Suppose Algorithm 1 runs T iterations until convergence, given that conditions $p, T, V, \gamma_v \ll n, d_v$ are satisfied, the final time complexity lies in $\mathrm{O}(n\log n + n\sum_{v=1}^{V} d_v)$, resulting in a highly-efficient computational speed.

3.4.2. Space complexity

The Algorithm 1 relates to five variables in terms of space complexity. The calculations of $\{\mathrm{F}^{(v)}\}_{v=1}^{V}$ and $\{\mathrm{P}^{(v)}\}_{v=1}^{V}$ require $\mathrm{O}(p\sum_{v=1}^{V}\gamma_v)$ and $\mathrm{O}(\sum_{v=1}^{V}\gamma_v d_v)$ space cost, respectively. Furthermore, the calculations of $\mathcal{C}$ and $\{\mathrm{B}^{(v)}\}_{v=1}^{V}$ and $\mathcal{Y}$ require the same space cost of $\mathrm{O}(Vnp)$. Thus, the overall space complexity of the optimization algorithm lies in $\mathrm{O}(n + \sum_{v=1}^{V} d_v)$, whose storage is linear w.r.t. data samples or dimensions of all views.

3.5. Theoretical convergence analysis

In this section, the convergence property of Algorithm 1 is analyzed from a theoretical perspective. According to the relevant works [48–50], we have the following theorems.

Theorem 1. *In Algorithm 1, the sequences* $\{\mathcal{C}, \{\mathrm{B}^{(v)}\}_{v=1}^{V}, \{\mathrm{P}^{(v)}\}_{v=1}^{V}, \{\mathrm{F}^{(v)}\}_{v=1}^{V}, \mathcal{Y}\}$ *are bounded.*

Proof. At the $(t+1)$th iteration, by considering the first-order optimality of $\mathcal{C}$, we can have the following equation:

$$\lambda\frac{\partial\|\mathcal{C}_{t+1}\|_{\circledast}}{\partial\mathcal{C}_{t+1}}+\mu_t(\mathcal{C}_{t+1}-\mathcal{B}_{t+1}-\frac{1}{\mu_t}\mathcal{Y}_t)=0 \tag{40}$$

In the Fourier domain, suppose $\mathcal{C}_f$ is the corresponding Fourier tensor of variable $\mathcal{C}$, and for the kth frontal slice of $\mathcal{C}_f$ (i.e., ${\mathcal{C}_f}^{(k)}$), we have

$$\begin{aligned}\frac{\partial\|{\mathcal{C}_f}^{(k)}\|_*}{\partial{\mathcal{C}_f}^{(k)}}&=\frac{\partial\mathrm{Tr}({\mathcal{O}_f}^{(k)})}{\partial{\mathcal{C}_f}^{(k)}}=\frac{\partial\mathrm{Tr}({\mathcal{U}_f}^{(k)^T}{\mathcal{C}_f}^{(k)}{\mathcal{V}_f}^{(k)})}{\partial{\mathcal{C}_f}^{(k)}}\\&=\frac{\partial\mathrm{Tr}({\mathcal{C}_f}^{(k)}{\mathcal{V}_f}^{(k)}{\mathcal{U}_f}^{(k)^T})}{\partial{\mathcal{C}_f}^{(k)}}={\mathcal{U}_f}^{(k)}{\mathcal{V}_f}^{(k)^T}\end{aligned} \tag{41}$$

where ${\mathcal{C}_f}^{(k)}={\mathcal{U}_f}^{(k)}{\mathcal{O}_f}^{(k)}{\mathcal{V}_f}^{(k)^T}$ is the SVD decomposition of matrix ${\mathcal{C}_f}^{(k)}$ and $\|{\mathcal{C}_f}^{(k)}\|_*$ is the nuclear norm of matrix ${\mathcal{C}_f}^{(k)}$. As the size of ${\mathcal{C}_f}^{(k)}$ is $n\times p$, and it holds that ${\mathcal{U}_f}^{(k)^T}{\mathcal{U}_f}^{(k)}=\mathrm{I}$ and ${\mathcal{V}_f}^{(k)^T}{\mathcal{V}_f}^{(k)}=\mathrm{I}$, we can have

$$\|\frac{\partial\|{\mathcal{C}_f}^{(k)}\|_*}{\partial{\mathcal{C}_f}^{(k)}}\|_F=\|{\mathcal{U}_f}^{(k)}{\mathcal{V}_f}^{(k)^T}\|_F\le\|{\mathcal{U}_f}^{(k)}\|_F\|{\mathcal{V}_f}^{(k)^T}\|_F=p \tag{42}$$

Therefore, the term $\frac{\partial\|{\mathcal{C}_f}^{(k)}\|_*}{\partial{\mathcal{C}_f}^{(k)}}$ is bounded. According to the observation in Eq. (3) and the chain rule, we have

$$\begin{aligned}\|\frac{\partial\|\mathcal{C}_{t+1}\|_{\circledast}}{\partial\mathcal{C}_{t+1}}\|_F&=\|\frac{\partial\sum_{k=1}^{V}\|({\mathcal{C}_f}^{(k)})_{t+1}\|_*}{\partial\mathcal{C}_{t+1}}\|_F=\|\sum_{k=1}^{V}\frac{\partial\|({\mathcal{C}_f}^{(k)})_{t+1}\|_*}{\partial\mathcal{C}_{t+1}}\|_F\\&\le\sum_{k=1}^{V}\|\frac{\partial\|({\mathcal{C}_f}^{(k)})_{t+1}\|_*}{\partial\mathcal{C}_{t+1}}\|_F=\sum_{k=1}^{V}\|\frac{\partial\|({\mathcal{C}_f}^{(k)})_{t+1}\|_*}{\partial({\mathcal{C}_f}^{(k)})_{t+1}}\frac{\partial({\mathcal{C}_f}^{(k)})_{t+1}}{\partial(\mathcal{C}_f)_{t+1}}\frac{\partial(\mathcal{C}_f)_{t+1}}{\partial\mathcal{C}_{t+1}}\|_F\\&\le\sum_{k=1}^{V}\|\frac{\partial\|({\mathcal{C}_f}^{(k)})_{t+1}\|_*}{\partial({\mathcal{C}_f}^{(k)})_{t+1}}\|_F\|\frac{\partial({\mathcal{C}_f}^{(k)})_{t+1}}{\partial(\mathcal{C}_f)_{t+1}}\|_F\|\widehat{\Gamma_V}\|_F\\&=p\|\widehat{\Gamma_V}\|_F\sum_{k=1}^{V}\|\frac{\partial({\mathcal{C}_f}^{(k)})_{t+1}}{\partial(\mathcal{C}_f)_{t+1}}\|_F\end{aligned} \tag{43}$$

where $(\mathcal{C}_f)_{t+1}$ is the corresponding Fourier tensor of variable $\mathcal{C}$ at the $t+1$th iteration, and $({\mathcal{C}_f}^{(k)})_{t+1}$ denotes the kth frontal slice of $(\mathcal{C}_f)_{t+1}$. Besides, it holds that $\widehat{\Gamma_V}=\frac{1}{\sqrt{V}}\Gamma_V$, where Γ_V is the discrete Fourier transform matrix with a size of $V\times V$.

According to the analysis result of Eq. (43), it can be seen that $\frac{\partial\|\mathcal{C}_{t+1}\|_{\circledast}}{\partial\mathcal{C}_{t+1}}$ is bounded. Moreover, by combining Eqs. (38) and (40), it is clear that:

$$\begin{aligned}&\mathcal{Y}_{t+1}=\mathcal{Y}_t+\mu_t(\mathcal{B}_{t+1}-\mathcal{C}_{t+1})\\&\Leftrightarrow\lambda\frac{\partial\|\mathcal{C}_{t+1}\|_{\circledast}}{\partial\mathcal{C}_{t+1}}=\mathcal{Y}_{t+1}\end{aligned} \tag{44}$$

Thus, we can see that the $\mathcal{Y}_{t+1}$ is also bounded.

According to the augmented Lagrangian function in Eq. (28), the subsequent chains of equalities can be achieved as

$$\begin{aligned}&\mathcal{L}\left(\mathcal{C}_{t+1},\mathcal{B}_{t+1},\{\mathrm{P}^{(v)}{}_{t+1}\}_{v=1}^{V},\{\mathrm{F}^{(v)}{}_{t+1}\}_{v=1}^{V},\mathcal{Y}_{t+1},\mu_{t+1}\right)\\=&\mathcal{L}\left(\mathcal{C}_{t+1},\mathcal{B}_{t+1},\{\mathrm{P}^{(v)}{}_{t+1}\}_{v=1}^{V},\{\mathrm{F}^{(v)}{}_{t+1}\}_{v=1}^{V},\mathcal{Y}_t,\mu_t\right)\\&+\frac{\mu_{t+1}}{2}\|\mathcal{B}_{t+1}-\mathcal{C}_{t+1}+\frac{1}{\mu_{t+1}}\mathcal{Y}_{t+1}\|_F^2-\frac{\mu_t}{2}\|\mathcal{B}_{t+1}-\mathcal{C}_{t+1}+\frac{1}{\mu_t}\mathcal{Y}_t\|_F^2\\=&\mathcal{L}\left(\mathcal{C}_{t+1},\mathcal{B}_{t+1},\{\mathrm{P}^{(v)}{}_{t+1}\}_{v=1}^{V},\{\mathrm{F}^{(v)}{}_{t+1}\}_{v=1}^{V},\mathcal{Y}_t,\mu_t\right)\\&+\frac{\mu_{t+1}+\mu_t}{2{\mu_t}^2}\|\mathcal{Y}_{t+1}-\mathcal{Y}_t\|_F^2+\frac{1}{2\mu_{t+1}}\|\mathcal{Y}_{t+1}\|_F^2-\frac{1}{2\mu_t}\|\mathcal{Y}_t\|_F^2\end{aligned} \tag{45}$$

Note that $\mathcal{C}_{t+1},\mathcal{B}_{t+1},\{\mathrm{P}^{(v)}{}_{t+1}\}_{v=1}^{V}$ and $\{\mathrm{F}^{(v)}{}_{t+1}\}_{v=1}^{V}$ represent the optimal solutions w.r.t. their corresponding subproblems at the $t+1$th iterations, which implies that:

$$\begin{aligned}&\mathcal{L}\left(\mathcal{C}_{t+1},\mathcal{B}_{t+1},\{\mathrm{P}^{(v)}{}_{t+1}\}_{v=1}^{V},\{\mathrm{F}^{(v)}{}_{t+1}\}_{v=1}^{V},\mathcal{Y}_t,\mu_t\right)\\&\le\mathcal{L}\left(\mathcal{C}_t,\mathcal{B}_t,\{\mathrm{P}^{(v)}{}_t\}_{v=1}^{V},\{\mathrm{F}^{(v)}{}_t\}_{v=1}^{V},\mathcal{Y}_t,\mu_t\right)\end{aligned} \tag{46}$$

By combining Eqs. (45) and (45), the following inequality can be obtained:

$$\begin{aligned}&\mathcal{L}\left(\mathcal{C}_{t+1},\mathcal{B}_{t+1},\{\mathrm{P}^{(v)}{}_{t+1}\}_{v=1}^{V},\{\mathrm{F}^{(v)}{}_{t+1}\}_{v=1}^{V},\mathcal{Y}_{t+1},\mu_{t+1}\right)\\&\le\mathcal{L}\left(\mathcal{C}_t,\mathcal{B}_t,\{\mathrm{P}^{(v)}{}_t\}_{v=1}^{V},\{\mathrm{F}^{(v)}{}_t\}_{v=1}^{V},\mathcal{Y}_t,\mu_t\right)\\&+\frac{\mu_{t+1}+\mu_t}{2{\mu_t}^2}\|\mathcal{Y}_{t+1}-\mathcal{Y}_t\|_F^2+\frac{1}{2\mu_{t+1}}\|\mathcal{Y}_{t+1}\|_F^2-\frac{1}{2\mu_t}\|\mathcal{Y}_t\|_F^2\end{aligned} \tag{47}$$

Hence, by iteratively summing the above equality t times, the following inequality can be achieved:

$$\begin{aligned}&\mathcal{L}\left(\mathcal{C}_{t+1},\mathcal{B}_{t+1},\{\mathrm{P}^{(v)}{}_{t+1}\}_{v=1}^{V},\{\mathrm{F}^{(v)}{}_{t+1}\}_{v=1}^{V},\mathcal{Y}_{t+1},\mu_{t+1}\right)\\&\le\mathcal{L}\left(\mathcal{C}_0,\mathcal{B}_0,\{\mathrm{P}^{(v)}{}_0\}_{v=1}^{V},\{\mathrm{F}^{(v)}{}_0\}_{v=1}^{V},\mathcal{Y}_0,\mu_0\right)\\&+\sum_{l=0}^{t}\frac{\mu_{l+1}+\mu_l}{2{\mu_l}^2}\|\mathcal{Y}_{l+1}-\mathcal{Y}_l\|_F^2+\frac{1}{2\mu_{t+1}}\|\mathcal{Y}_{t+1}\|_F^2-\frac{1}{2\mu_0}\|\mathcal{Y}_0\|_F^2\end{aligned} \tag{48}$$

According to the relevant research in [48], the term $\sum_{l=0}^{t}\frac{\mu_{l+1}+\mu_l}{2{\mu_l}^2}$ is bounded. Note that $\mathcal{L}\left(\mathcal{C}_0,\mathcal{B}_0,\{\mathrm{P}^{(v)}{}_0\}_{v=1}^{V},\{\mathrm{F}^{(v)}{}_0\}_{v=1}^{V},\mathcal{Y}_0,\mu_0\right)$ is also bounded for any initialization. Thereafter, we can conclude that the augmented Lagrangian function at the $t+1$th iteration is finite, that is

$$\begin{aligned}&\mathcal{L}\left(\mathcal{C}_{t+1},\mathcal{B}_{t+1},\{\mathrm{P}^{(v)}{}_{t+1}\}_{v=1}^{V},\{\mathrm{F}^{(v)}{}_{t+1}\}_{v=1}^{V},\mathcal{Y}_{t+1},\mu_{t+1}\right)\\=&\sum_{v=1}^{V}\left\|\mathrm{X}^{(v)}-\mathrm{P}^{(v)}{}_{t+1}\mathrm{F}^{(v)}{}_{t+1}\mathrm{B}^{(v)}{}_{t+1}\right\|_F^2+\lambda\|\mathcal{C}_{t+1}\|_{\circledast}\\&+\frac{\mu_{t+1}}{2}\|\mathcal{B}_{t+1}-\mathcal{C}_{t+1}-\frac{1}{\mu_{t+1}}\mathcal{Y}_{t+1}\|_F^2\end{aligned} \tag{49}$$

Based on the boundedness of $\mathcal{L}\left(\mathcal{C}_{t+1},\mathcal{B}_{t+1},\{\mathrm{P}^{(v)}{}_{t+1}\}_{v=1}^{V},\{\mathrm{F}^{(v)}{}_{t+1}\}_{v=1}^{V},\mathcal{Y}_{t+1},\mu_{t+1}\right)$, we can see that each term on the right side is also bounded. The term $\|\mathcal{C}_{t+1}\|_{\circledast}$ is bounded, which implies that all singular values of $\mathcal{C}_{t+1}$ is bounded. While the term $\|\mathcal{C}_{t+1}\|_F^2$ is bounded (i.e., the sum of squares of singular values), the $\mathcal{C}_{t+1}$ is also bounded.

Furthermore, since the term $\|\mathcal{B}_{t+1}-\mathcal{C}_{t+1}-\frac{1}{\mu_{t+1}}\mathcal{Y}_{t+1}\|_F^2$ is bounded, in light of the boundedness of $\mathcal{C}_{t+1}$ and $\mathcal{Y}_{t+1}$, we can see that the $\mathcal{B}_{t+1}$ (i.e., $\{\mathrm{B}^{(v)}{}_{t+1}\}_{v=1}^{V}$) is also bounded. Similarity, it is clear that $\sum_{v=1}^{V}\left\|\mathrm{X}^{(v)}-\mathrm{P}^{(v)}{}_{t+1}\mathrm{F}^{(v)}{}_{t+1}\mathrm{B}^{(v)}{}_{t+1}\right\|_F^2$ is bounded, hence $\{\mathrm{P}^{(v)}{}_{t+1}\}_{v=1}^{V}$ and $\{\mathrm{F}^{(v)}{}_{t+1}\}_{v=1}^{V}$ are also bounded. □

Theorem 2. *In Algorithm 1, the sequences* $\{\mathcal{C}, \{\mathrm{B}^{(v)}\}_{v=1}^{V}, \{\mathrm{P}^{(v)}\}_{v=1}^{V}, \{\mathrm{F}^{(v)}\}_{v=1}^{V}, \mathcal{Y}\}$ *has at least one accumulation point. Suppose that when* $t\to\infty$, *then* $(\mathcal{C}_{t+1}-\mathcal{C}_t)\to 0$, $(\mathcal{B}_{t+1}-\mathcal{B}_t)\to 0$, $(\mathrm{F}^{(v)}{}_{t+1}-\mathrm{F}^{(v)}{}_t,\forall v)\to 0$, $(\mathrm{P}^{(v)}{}_{t+1}-\mathrm{P}^{(v)}{}_t,\forall v)\to 0$ *and* $(\mathcal{Y}_{t+1}-\mathcal{Y}_t)\to 0$ *are satisfied. For any accumulation point* $\{\mathcal{C}_*,\{\mathrm{B}^{(v)}{}_*\}_{v=1}^{V},\{\mathrm{P}^{(v)}{}_*\}_{v=1}^{V},\{\mathrm{F}^{(v)}{}_*\}_{v=1}^{V},\mathcal{Y}_*\}$, *the corresponding subsequence* $\{\mathcal{C}_*,\{\mathrm{B}^{(v)}{}_*\}_{v=1}^{V},\{\mathrm{P}^{(v)}{}_*\}_{v=1}^{V},\{\mathrm{F}^{(v)}{}_*\}_{v=1}^{V}\}$ *is also a stationary point of the optimization problem in Eq. (28).*

Proof. According to the Bolzano–Weierstrass theorem, the sequence $\{\mathcal{C},\{\mathrm{B}^{(v)}\}_{v=1}^{V},\{\mathrm{P}^{(v)}\}_{v=1}^{V},\{\mathrm{F}^{(v)}\}_{v=1}^{V},\mathcal{Y}\}$ has at least one accumulation point $\{\mathcal{C}_*,\{\mathrm{B}^{(v)}{}_*\}_{v=1}^{V},\{\mathrm{P}^{(v)}{}_*\}_{v=1}^{V},\{\mathrm{F}^{(v)}{}_*\}_{v=1}^{V},\mathcal{Y}_*\}$. Without loss of generality, we assume that the sequence $\{\mathcal{C}_t,\{\mathrm{B}^{(v)}{}_t\}_{v=1}^{V},\{\mathrm{P}^{(v)}{}_t\}_{v=1}^{V},\{\mathrm{F}^{(v)}{}_t\}_{v=1}^{V},\mathcal{Y}_t\}$ itself converges to $\{\mathcal{C}_*,\{\mathrm{B}^{(v)}{}_*\}_{v=1}^{V},\{\mathrm{P}^{(v)}{}_*\}_{v=1}^{V},\{\mathrm{F}^{(v)}{}_*\}_{v=1}^{V},\mathcal{Y}_*\}$.

First, due to the assumption of $\mathcal{Y}$ and the update rule in Eq. (38), we have the following equalities:

$$\begin{aligned}&\mathcal{B}_* - \mathcal{C}_* \\ =&\lim_{t\to\infty}\{\mathcal{B}_t - \mathcal{C}_t\} \\ =&\lim_{t\to\infty}\{\mathcal{Y}_{t+1} - \mathcal{Y}_t\}/\mu_t \\ =&0\end{aligned} \tag{50}$$

At $t+1$ iterations, based on the first-order optimality of the $\mathrm{F}^{(v)}$, we have

$$\frac{\partial \mathcal{L}(\mathcal{C}_t, \mathcal{B}_t, \mathrm{P}^{(v)}{}_t, \mathrm{F}^{(v)}{}_{t+1}, \mathcal{Y}_t, \mu_t)}{\partial \mathrm{F}^{(v)}{}_{t+1}} = 0 \tag{51}$$

Under the assumption that $\lim_{t\to\infty}(\mathrm{F}^{(v)}{}_{t+1} - \mathrm{F}^{(v)}{}_t) \to 0$ and the condition that the sequence $\{\mathcal{C}_t, \{\mathrm{B}^{(v)}{}_t\}_{v=1}^V, \{\mathrm{P}^{(v)}{}_t\}_{v=1}^V, \{\mathrm{F}^{(v)}{}_t\}_{v=1}^V, \mathcal{Y}_t\}$ converges to $\{\mathcal{C}_*, \{\mathrm{B}^{(v)}{}_*\}_{v=1}^V, \{\mathrm{P}^{(v)}{}_*\}_{v=1}^V, \{\mathrm{F}^{(v)}{}_*\}_{v=1}^V, \mathcal{Y}_*\}$, we have

$$\begin{aligned}&\lim_{t\to\infty}\frac{\partial \mathcal{L}(\mathcal{C}_t, \mathcal{B}_t, \mathrm{P}^{(v)}{}_t, \mathrm{F}^{(v)}{}_{t+1}, \mathcal{Y}_t, \mu_t)}{\partial \mathrm{F}^{(v)}{}_{t+1}} \\ =&\frac{\partial \left\|\mathrm{X}^{(v)} - \mathrm{P}^{(v)}{}_t \mathrm{F}^{(v)}{}_* \mathrm{B}^{(v)}{}_t\right\|_F^2}{\partial \mathrm{F}^{(v)}{}_*}\end{aligned} \tag{52}$$

$$=0 \tag{53}$$

By considering the first-order optimality of the $\mathrm{B}^{(v)}$ at the $(t+1)$th iteration, we have

$$\frac{\partial \mathcal{L}(\mathcal{C}_t, \mathrm{B}^{(v)}{}_{t+1}, \mathrm{P}^{(v)}{}_t, \mathrm{F}^{(v)}{}_{t+1}, \mathcal{Y}_t, \mu_t)}{\partial \mathrm{B}^{(v)}{}_{t+1}} = 0 \tag{54}$$

Next, when considering the first-order optimality of the $\mathcal{C}$ at the $(t+1)$th iteration, it can be seen that

$$\begin{aligned}&\lim_{t\to\infty}\frac{\partial \mathcal{L}(\mathcal{C}_{t+1}, \mathrm{B}^{(v)}{}_t, \mathrm{P}^{(v)}{}_t, \mathrm{F}^{(v)}{}_{t+1}, \mathcal{Y}_t, \mu_t)}{\partial \mathrm{C}^{(v)}{}_{t+1}} \\ =&\lim_{t\to\infty}\{\frac{\partial \left\|\mathcal{C}_{t+1}\right\|_{\circledast}}{\partial \mathcal{C}_{t+1}} + \mu_t(\mathcal{C}_{t+1} - \mathcal{B}_t) - \mathcal{Y}_t\} \\ =&\lim_{t\to\infty}\{\frac{\partial \left\|\mathcal{C}_{t+1}\right\|_{\circledast}}{\partial \mathcal{C}_{t+1}} + \mu_t(\mathcal{C}_{t+1} - \mathcal{B}_{t+1}) - \mathcal{Y}_t + \mu_t(\mathrm{B}^{(v)}{}_{t+1} - \mathrm{B}^{(v)}{}_t)\} \\ =&\lim_{t\to\infty}\{\frac{\partial \left\|\mathcal{C}_{t+1}\right\|_{\circledast}}{\partial \mathcal{C}_{t+1}} - \mathcal{Y}_{t+1}\} \\ =&\frac{\partial \left\|\mathcal{C}_*\right\|_{\circledast}}{\partial \mathcal{C}_*} - \mathcal{Y}_* \\ =&0\end{aligned} \tag{55}$$

Similarity, we can have

$$\begin{aligned}&\lim_{t\to\infty}\frac{\partial \mathcal{L}(\mathcal{C}_{t+1}, \mathrm{B}^{(v)}{}_{t+1}, \mathrm{P}^{(v)}{}_t, \mathrm{F}^{(v)}{}_{t+1}, \mathcal{Y}_t, \mu_t)}{\partial \mathrm{B}^{(v)}{}_{t+1}} \\ =&\lim_{t\to\infty}\{\frac{\partial \left\|\mathrm{X}^{(v)} - \mathrm{P}^{(v)}{}_t \mathrm{F}^{(v)}{}_{t+1} \mathrm{B}^{(v)}{}_{t+1}\right\|_F^2}{\partial \mathrm{B}^{(v)}{}_{t+1}} + \mu_t(\mathrm{B}^{(v)}{}_{t+1} - \mathrm{C}^{(v)}{}_{t+1}) + \mathrm{Y}_t\} \\ =&\lim_{t\to\infty}\{\frac{\partial \left\|\mathrm{X}^{(v)} - \mathrm{P}^{(v)}{}_t \mathrm{F}^{(v)}{}_{t+1} \mathrm{B}^{(v)}{}_{t+1}\right\|_F^2}{\partial \mathrm{B}^{(v)}{}_{t+1}} + \mathrm{Y}_{t+1}\} \\ =&\{\frac{\partial \left\|\mathrm{X}^{(v)} - \mathrm{P}^{(v)}{}_t \mathrm{F}^{(v)}{}_* \mathrm{B}^{(v)}{}_*\right\|_F^2}{\partial \mathrm{B}^{(v)}{}_*} + \mathrm{Y}_*\} \\ =&0\end{aligned} \tag{56}$$

Similar to Eqs. (51) and (51), based on the first-order optimality of $\mathrm{P}^{(v)}$, we have

$$\begin{aligned}&\lim_{t\to\infty}\frac{\partial \mathcal{L}(\mathcal{C}_{t+1}, \mathcal{B}_{t+1}, \mathrm{P}^{(v)}{}_{t+1}, \mathrm{F}^{(v)}{}_{t+1}, \mathcal{Y}_t, \mu_t)}{\partial \mathrm{P}^{(v)}{}_{t+1}} \\ =&\lim_{t\to\infty}\frac{\partial \left\|\mathrm{X}^{(v)} - \mathrm{P}^{(v)}{}_{t+1} \mathrm{F}^{(v)}{}_{t+1} \mathrm{B}^{(v)}{}_{t+1}\right\|_F^2}{\partial \mathrm{P}^{(v)}{}_{t+1}} \\ =&\frac{\partial \left\|\mathrm{X}^{(v)} - \mathrm{P}^{(v)}{}_* \mathrm{F}^{(v)}{}_* \mathrm{B}^{(v)}{}_*\right\|_F^2}{\partial \mathrm{P}^{(v)}{}_*} \\ =&0\end{aligned} \tag{57}$$

Therefore, the sequence $\{\mathcal{C}_*, \{\mathrm{B}^{(v)}{}_*\}_{v=1}^V, \{\mathrm{P}^{(v)}{}_*\}_{v=1}^V, \{\mathrm{F}^{(v)}{}_*\}_{v=1}^V, \mathcal{Y}_*\}$ satisfies the KKT condition, thus the corresponding subsequence $\{\mathcal{C}_*, \{\mathrm{B}^{(v)}{}_*\}_{v=1}^V, \{\mathrm{P}^{(v)}{}_*\}_{v=1}^V, \{\mathrm{F}^{(v)}{}_*\}_{v=1}^V\}$ is a stationary point of original problem in Eq. (28). □

4. Experiments

This section evaluates our ST-MSC approach with regard to three aspects, i.e., comparison against other approaches, hyper-parameter analysis and computational efficiency. Notice that, we run the experiments are on the PC refers to i5-6600 CPU and 64 GB of RAM.

4.1. Datasets and evaluation metrics

To fully prove the effectiveness and efficiency of the ST-MSC approach, we employ nine multi-view datasets for experimental purposes, including five general-scale benchmarks (with $n < 9000$) and four large-scale benchmarks (with $n \geq 9000$). It is worth mentioning that these benchmarks cover diverse domains, and have been commonly adopted in multi-view data analysis.

Five general-scale datasets, namely *ORL* [22], *BBCSport* [51], *COIL-20* [52], *UCI Digits* [30] and *NH-4660* [17] are employed in the experiments. In addition, four large-scale datasets, namely *Caltech101-all* [8], *ALOI-100* [8], *Reuters-lee* [4] and *NUSWIDEOBJ* [4] are deliberately adopted for experiments. The details of these benchmarks are summarized in Table 1.

To empirically measure the clustering quality of different approaches, we employ two widely-used metrics, they are Normalized Mutual Information (NMI) [53] and Accuracy (ACC) [54]. For each dataset, we will run each clustering approach 20 times and report the average score.

4.2. Comparison against other multi-view clustering methods

This subsection evaluates the proposed ST-MSC approach, by comparing it against tweleve typical baselines, which are listed below.

- **DiMSC** [17]: Diversity-guided Multi-view Subspace Clustering.
- **LMSC** [19]: Latent Multi-View Subspace Clustering.
- **SwMC** [52]: Self-weighted multi-view clustering with multiple graphs.
- **MLAN** [3]: Multi-view clustering via Leaning Adaptive Neighbors.
- **MVSC** [30]: Scalable multi-view spectral clustering via bipartite graph.
- **SFMC** [5]: Scalable and parameter-Free Multi-view graph Clustering.
- **LMVSC** [31]: Large-scale Multi-view Subspace Clustering.
- **SMVSC** [33]: Scalable Multi-view Subspace Clustering with Unified Anchors.
- **FPMVS** [34]: Fast Parameter-free Multi-view Subspace Clustering with Consensus Anchor Guidance.
- **OMSC** [55]: Orthogonal and fast Multi-view Subspace Clustering.
- **TBGL** [6]: Tensorized Bipartite Graph Learning for multi-view clustering.
- **S^{2}MVTC** [56]: Simple yet Efficient Scalable Multi-View Tensor Clustering

Table 1
The statistics of used multi-view benchmarks.

Dataset		#Sample	#Class	Views & dimensions
General	*ORL*	400	40	Intensity(4096), LBP(3304), Gabor(6750)
	BBCSport	544	5	View1(3183), View2(3203)
	COIL20	1440	20	Intensity (1024), LBP (3304), Gabor (6750)
	UCI Digits	2000	10	Pix(240), Fou(76), MOR(6)
	NH-4660	4660	5	Intensity(2000), LBP(3304), Gabor(6750)
Large	*Caltech101-all*	9144	102	Gabor(48), WM(40), CENTRIST(254), HOG(1984), GIST(512), LBP(928)
	ALOI-100	10,800	100	Similarity(77), Haralick(13), HSV(64), RGB(125)
	Reuters-lee	18,758	6	English(21,531), French(24,892), German(34,251), Spanish(15,506), Italian(11,547)
	NUSWIDEOBJ	30,000	31	CH(64), CM(255), CORR(144), EDH (73), WT (128)

Table 2
The average NMI (%) scores over 20 runs by thirteen multi-view clustering approaches. On each dataset, the best scores are highlighted in **bold**, while the second best one in [brackets].

Datasets	DiMSC	LMSC	SwMC	MLAN	MVSC	SFMC	LMVSC	SMVSC	FPMVS	OMSC	TBGL	S^2MVTC	ST-MSC
ORL	90.8	92.2	83.3	78.6	84.7	80.6	76.4	75.3	74.3	78.4	[97.2]	NA	**97.5**
BBCSport	69.0	75.2	46.9	51.4	69.3	5.2	[82.9]	14.8	10.4	12.4	**96.9**	NA	[96.0]
COIL20	83.6	89.4	**94.3**	92.4	75.7	79.2	77.3	73.5	74.6	74.8	88.9	89.2	[92.3]
UCI Digits	79.2	77.9	69.8	93.4	64.4	75.2	75.6	80.4	73.5	83.6	93.7	[95.2]	**98.6**
NH-4660	78.6	67.4	8.8	55.5	54.0	[84.1]	68.1	66.6	67.2	72.2	**92.4**	84.1	[91.0]
Caltech101-all	OM	OM	15.9	OM	23.5	33.4	37.2	37.0	35.9	36.5	31.2	[83.0]	**85.2**
ALOI-100	OM	OM	OM	OM	30.5	60.8	75.0	57.4	55.5	49.3	77.4	[92.0]	**92.4**
Reuters-lee	OM	OM	OM	OM	OM	5.8	44.5	32.2	34.6	32.3	OM	[68.6]	**97.0**
NUSWIDEOBJ	OM	OM	OM	OM	OM	8.1	8.6	12.0	12.4	[12.7]	OM	[12.7]	**20.1**
Average score	OM	OM	OM	OM	OM	43.5	[61.2]	49.9	48.7	54.9	OM	NA	**85.6**
Average rank	7.0	6.8	8.0	7.3	8.3	7.1	5.5	7.3	8.0	6.3	5.4	[5.3]	**1.4**

* Note that "OM" indicates the out-of-memory error, and "NA" indicates the baseline is not applicable to this dataset.

These baseline approaches are chosen because of their excellent clustering performance. They can be treated as two groups: traditional approaches and scalable approaches. The traditional approaches cover two graph-based approaches (e.g., SwMC and MLAN), and two subspace-based approaches (i.e., DiMSC and LMSC). The scalable approaches include two anchor graph-based approaches (e.g, MVSC and SFMC), four anchor subspace-based approaches (i.e., SMVSC, LMVSC, FPMVS and OMSC), and two tensorized anchor-based approaches (i.e., TBGL and S^2MVTC).

In the experiments, we adopt the experimental settings as follows:

- For the baseline approaches, their hyper-parameters are tuned in the range of $[10^{-5}, 10^{-4}, \ldots, 10^{4}, 10^{5}]$, unless some suggested setting is given in the corresponding papers.
- For the ST-MSC approach, the balancing hyper-parameter λ is set in the range of $[1, 10^{1}, \ldots, 10^{3}, 10^{5}]$, while the hyper-parameter that controls the number of anchors α is set in the range of $[k, 100, 200, 300, 400]$. For some small-scale datasets (with $n < 1,000$), the anchor number is set in the range of $[k, 50, 100, 150, 200]$.

Specialized, the comparison results with two stand metrics are recorded in Tables 2 and 3, respectively. From the experimental results, several key observations are provided below.

- In terms of the general-scale datasets, the traditional approaches can achieve quite robust clustering results. But in terms of the large-scale datasets, these traditional approaches become computationally infeasible for the datasets with more than ten thousand samples. Notably, out of the thirteen multi-view clustering approaches, only six approaches (including our proposed approach) are capable of dealing with all benchmark datasets.
- In terms of the high-dimensional datasets, such as ORL, BBCSport, COIL20, NH-4660, and Reuters-lee, our proposed approach achieves the highest or the second highest NMI score on each of these high-dimensional datasets.
- The two tensorized anchor-based approaches have achieved impressive results against the other competitors. Specifically, the NMI scores of TBGL rank in the top two on three out of the nine datasets, while the NMI scores of ST-MSC rank in the top two on all the nine datasets.

In summary, these experimental results have comprehensively confirmed the advantages of the proposed approach in terms of effectiveness.

Meanwhile, we carry out the visual experiments for clearer illustration and comparison. Note that four competitive baselines (i.e., LMVSC, SMVSC, FPMVSC and OMSC) are selected for comparison, while the plotted affinity matrix is computed via the following equation:

$$S = W_B W_B{}^T \tag{58}$$

where W_B represents the right orthogonal matrix of anchor presentation $\overline{B}$ (i.e., $\overline{B} = \mathrm{SVD}(U_B \Lambda_B W_B{}^T)$), as shown in Eq. (39). The visual results on three benchmark datasets (i.e., $ORL, BBCSport$ and $UCIDigits$) are shown in Figs. 4 to 6, respectively. From these three figures, it can be observed that the affinity matrices achieved by our proposed approach generally reveal comparatively clearer block diagonal structures over other test approaches. These visualization results are well consistent with the clustering results in Tables 2 and 3.

4.3. Hyper-parameter analysis

This subsection investigates the practical impacts of parameters λ and α across all benchmark datasets. λ is the balancing hyper-parameter, while α decides the number of selected anchors. Following the experimental settings in Sections 4.1 and 4.2, we vary the parameter λ in $\{1, 10^{1}, \ldots, 10^{4}, 10^{5}\}$. At the same time, the parameter p is tuned within the set of $\{k, 100, 200, 300, 400\}$ for large-scale datasets, and within the set of $\{k, 50, 100, 150, 200\}$ for general-scale datasets).

The sensitivity analysis over two metrics is depicted in Figs. 7 and 8, respectively. As we can see, larger values of λ and α often generate favorable results for general-scale cases. Especially, as hyper-parameter α goes from 100 to 200, our ST-MSC approach exhibits competitive clustering results on *ORL* and *UCI-3views* datasets. When it comes to large-scale cases, we find that the proposed approach performs relatively stable on *ALOI-100* and *NUSWIDEOBJ* datasets. On the other two

Table 3
The average Accuracy (%) scores over 20 runs by thirteen multi-view clustering approaches. On each dataset, the best scores are highlighted in **bold**, while the second best one in [brackets].

Datasets	DiMSC	LMSC	SwMC	MLAN	MVSC	SFMC	LMVSC	SMVSC	FPMVS	OMSC	TBGL	S^2MVTC	ST-MSC
ORL	80.9	81.3	70.8	68.4	72.4	63.5	57.0	56.3	56.0	61.5	[89.7]	NA	**91.0**
BBCSport	86.5	85.5	62.9	67.4	77.7	34.2	[94.7]	32.4	40.6	35.7	**99.1**	NA	[97.4]
COIL20	76.3	80.9	[86.4]	84.4	61.7	64.4	65.5	60.6	63.8	63.8	80.5	68.9	**89.1**
UCI Digits	86.5	86.3	65.5	[97.0]	62.0	75.2	80.3	83.4	72.2	91.4	89.1	89.4	**97.8**
NH-4660	84.3	74.6	33.9	59.2	62.9	[91.9]	73.8	71.6	71.3	87.2	**97.0**	88.2	[96.4]
Caltech101-all	OM	OM	16.1	OM	20.1	16.5	11.7	26.7	29.5	31.9	33.9	[52.4]	**53.2**
ALOI-100	OM	OM	OM	OM	12.4	45.4	56.6	34.3	31.5	26.1	72.8	[75.6]	**83.4**
Reuters-lee	OM	OM	OM	OM	OM	32.9	59.5	55.	57.5	48.0	OM	[71.5]	**98.9**
NUSWIDEOBJ	OM	OM	OM	OM	OM	11.9	12.1	18.9	[19.2]	18.7	OM	14.0	**22.3**
Average score	OM	OM	OM	OM	OM	43.0	50.1	48.8	52.8	[54.1]	OM	NA	**81.1**
Average rank	6.8	6.8	8.1	7.3	8.4	7.2	6.3	7.6	7.1	6.0	[5.6]	5.9	**1.2**

* Note that "OM" indicates the out-of-memory error, and "NA" indicates the baseline is not applicable to this dataset.

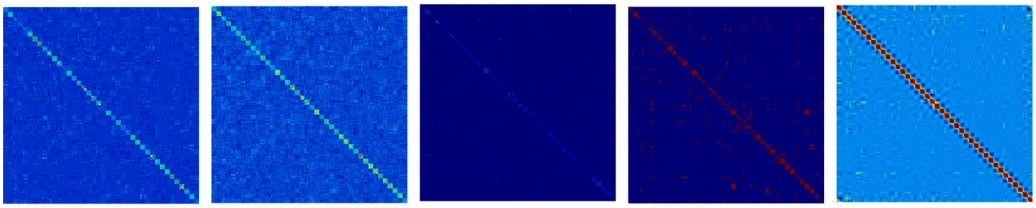


Fig. 4. The visualization of affinity matrices w.r.t. five comparison approaches on *ORL* dataset.

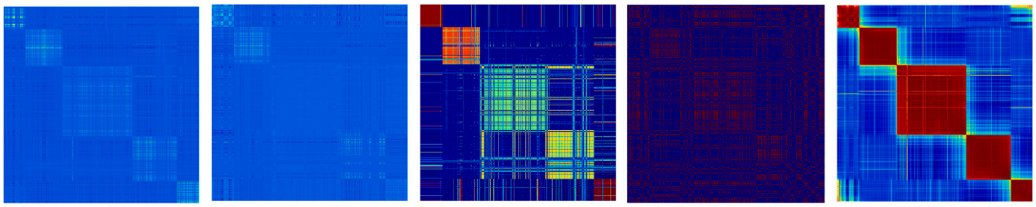


Fig. 5. The visualization of affinity matrices w.r.t. five comparison approaches on *BBCSport* dataset.

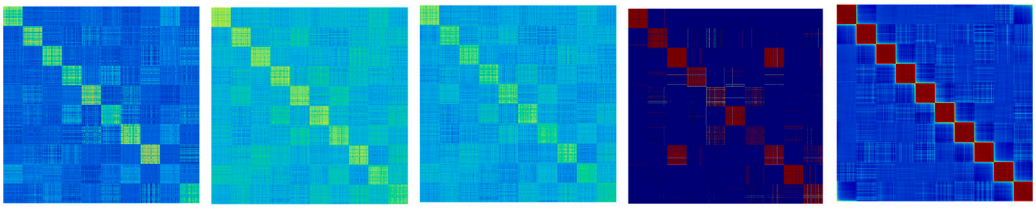


Fig. 6. The visualization of affinity matrices w.r.t. five comparison approaches on *UCI Digits* dataset.

Table 4
The average running time over 20 runs by thirteen multi-view clustering approaches.

Datasets	DiMSC	LMSC	SwMC	MLAN	MVSC	SFMC	LMVSC	SMVSC	FPMVS	OMSC	TBGL	S^2MVTC	ST-MSC
ORL	4.0	69.1	7.4	0.5	2.8	1.0	1.5	10.7	22.9	19.5	4.3	NA	1.4
BBCSport	6.6	21.5	16.5	25.1	1.8	0.5	0.8	3.9	7.1	13.1	5.2	NA	1.3
COIL20	71.8	228.7	112.5	5.9	8.3	2.7	5.3	16.9	32.7	26.4	506.2	2.0	4.9
UCI Digits	116.9	344.2	521.3	12.4	3.8	3.7	4.4	14.0	22.0	22.7	915.3	1.2	[1.7]
NH-p4660	1340.3	4732.4	5709.7	255.79	53.0	5.8	23.8	34.0	66.7	321.2	715.3	5.6	46.5
Caltech101-all	OM	OM	5012.0	OM	231.8	21.6	72.6	682.3	400.5	289.5	15 401.0	9.7	367.9
ALOI-100	OM	OM	OM	OM	205.7	25.7	25.5	124.9	241.0	192.3	12 011.4	4.0	300.2
Reuters-lee	OM	OM	OM	OM	OM	95.6	43.4	7374.0	1205.1	19 379.7	OM	118.6	1381.5
NUSWIDEOBJ	OM	OM	OM	OM	OM	60.0	92.0	203.2	386.5	310.0	OM	14.2	607.8

* Note that "OM" indicates the out-of-memory error, and "NA" indicates the baseline is not applicable to this dataset.

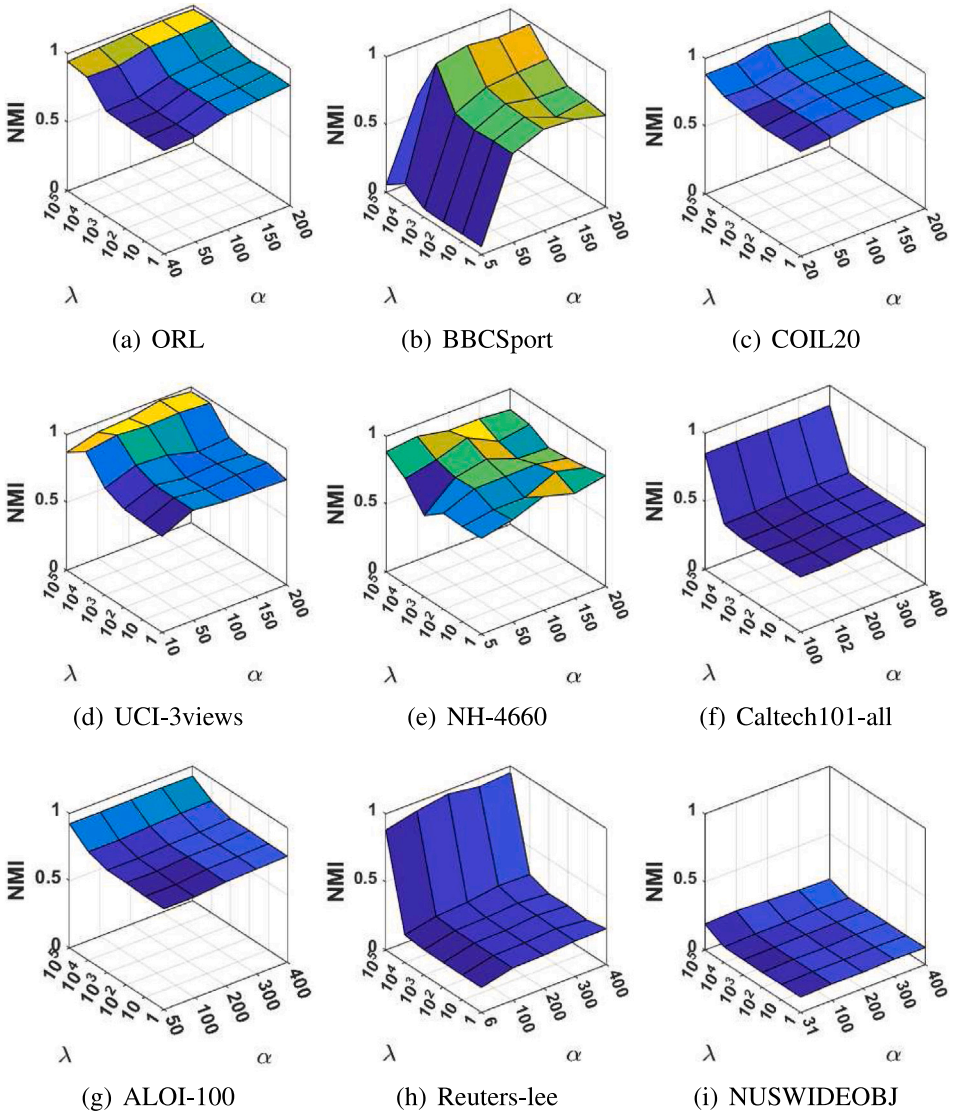


Fig. 7. Hyper-parameter analysis (in terms of NMI) on the parameters λ and α.

datasets, like *Reuters-lee*, the proposed approach is capable to obtain fairly better clustering results when a large λ is selected. Consequently, we argue that a practical combination of these two parameters is crucial in our ST-MSC approach for various applications.

4.4. Computational efficiency

This subsection conducts the computational analysis to reveal the time efficiency of all comparison approaches. The corresponding computational results are reported in Table 4. Note the "OM" indicates the out-of-memory error, while "NA" indicates other running errors.

When dealing with general-scale cases, the traditional approaches usually require more computational time compared to the scalable approaches, and the incorporation of anchor graph learning can substantially improve the running efficiency of multi-view clustering. When dealing with large-scale cases, most traditional approaches become computationally infeasible due to their high computational complexity. It is worth noting that our approach is capable of achieving time efficiency that is comparable to or even better than many scalable approaches (as shown in Table 4), while producing overall better clustering quality than most of baseline approaches (as shown in Tables 2 and 3).

4.5. Ablation study

This subsection investigates the effectiveness of two key components in the proposed model, namely the tri-factorization scheme and tensorized bipartite graph learning. Notably, two comparison baselines can be treated as special cases of our ST-MSC approach. The brief descriptions of them are given as follows.

- **Tensorized Fast Multi-view Subspace Clustering (TF-MSC)**: In terms of the traditional large-scale subspace clustering framework, this approach helps to exploit the high-order correlations across multiple anchor graphs. The target model of TF-MSC approach is provided in the following.

$$\min_{\{\mathrm{F}^{(v)},\mathrm{B}^{(v)}\}_{v=1}^{V}} \sum_{v=1}^{V} \left\| \mathrm{X}^{(v)} - \mathrm{F}^{(v)}\mathrm{B}^{(v)} \right\|_F^2 + \lambda \|\mathcal{B}\|_{\circledast} \quad \text{s.t. } \mathcal{B} = \Psi(\mathrm{B}^{(1)}, \mathrm{B}^{(2)}, \ldots, \mathrm{B}^{(V)}). \tag{59}$$

  where $\mathrm{F}^{(v)}$ is the anchor set, $\mathrm{Z}^{(v)}$ is the similarity matrix of the bipartite graph of the vth view. In particular, if we remove the tri-factorization component in ST-MSC, then our proposed approach is equivalent to this TF-MSC baseline.
- **Concept Factorization Based Multiview Clustering (CFMC)** [38]: This is a newly proposed tri-factorization based approach,

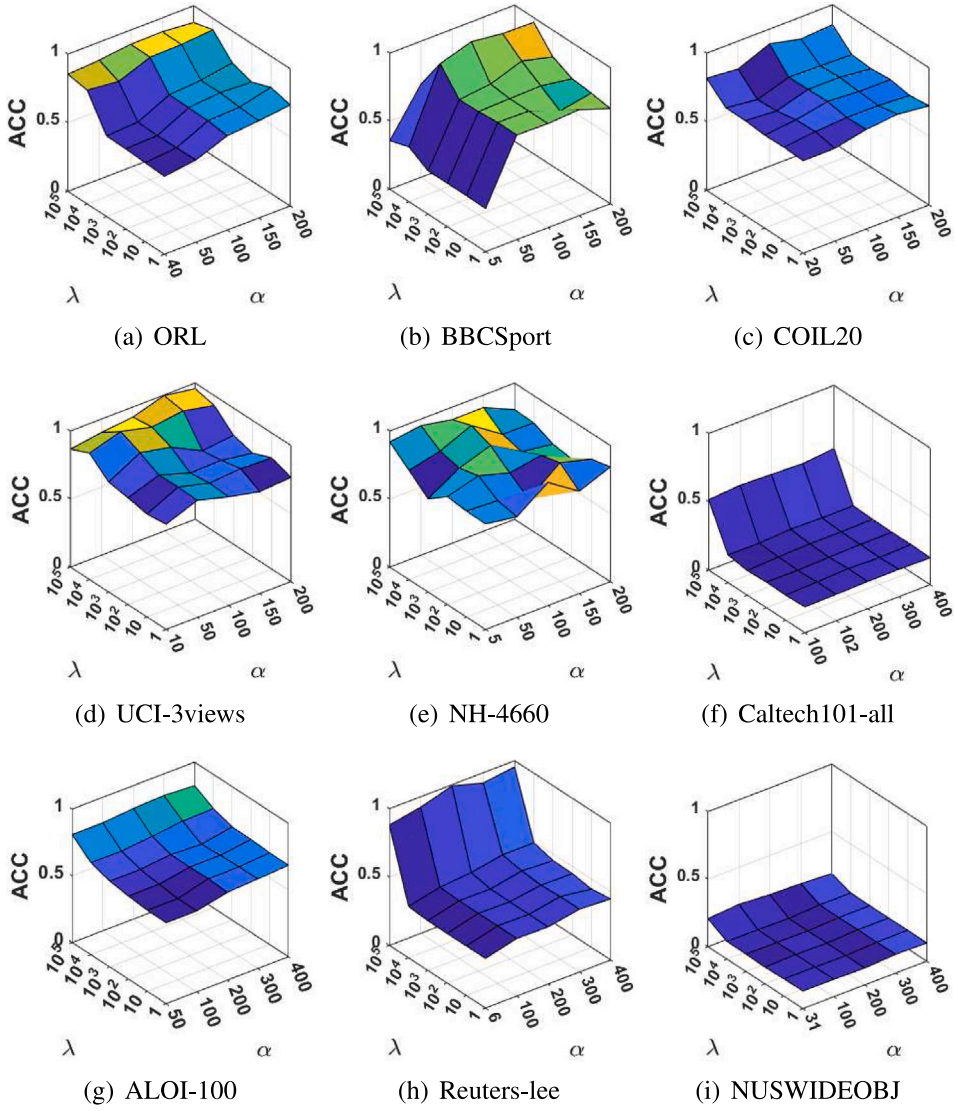


Fig. 8. Hyper-parameter analysis (in terms of Accuracy) on the parameters λ and α.

which aims to capture the semantic correlations between anchor bases and clustering assignments. The target model of CFMC is provided as follows.

$$\min_{\{\mathrm{A}^{(v)},\mathrm{W}^{(v)}\}_{v=1}^{V},\mathrm{Y}} \sum_{v=1}^{V} \left\| \mathrm{X}^{(v)} - \mathrm{A}^{(v)}\mathrm{W}^{(v)}\mathrm{Y} \right\|_F^2 \tag{60}$$
$$\text{s.t. } {\mathrm{A}^{(v)}}^T \mathrm{A}^{(v)} = \mathrm{I}, \mathrm{W}^{(v)} \geq 0, \forall v, \mathrm{Y} \in Ind.$$

where $\mathrm{A}^{(v)}$ is the anchor base in the vth view, Y is the consensus clustering assignment across multiple sources, and $\mathrm{W}^{(v)}$ is the coefficient matrix that exploits the underlying relationship between anchor sets and the consensus clustering assignment.

The clustering performances of these three approaches are reported in Table 5. According to these ablation results, TF-MSC usually produces more promising clustering results than CFMC. Also, we can observe that our ST-MSC approach consistently outperforms TF-MSC and CFMC in most of the comparisons. Especially, when it comes to the large-scale datasets, our ST-MSC approach has obtained significant improvements compared to TF-MSC.

Table 5
The Ablation study over NMI and Accuracy on nine benchmarks, where the best scores are highlighted in **bold** for each dataset.

Dataset	NMI			Acc		
	TF-MSC	CFMC	ST-MSC	TF-MSC	CFMC	ST-MSC
ORL	**97.9**	82.8	97.5	**91.4**	64.0	91.0
BBCSport	93.0	21.5	**96.0**	95.6	53.6	**97.4**
COIL20	**96.6**	83.9	92.3	88.2	68.7	**89.1**
UCI Digits	98.2	83.0	**98.6**	97.5	90.5	**97.8**
NH-p4660	83.2	73.6	**91.0**	90.4	80.9	**96.4**
Caltech101-all	**87.3**	49.5	85.2	**55.7**	28.9	53.2
ALOI-100	**93.6**	78.2	92.4	83.1	56.1	**83.4**
Reuters-lee	88.6	32.5	**97.0**	87.3	52.4	**98.9**
NUSWIDEOBJ	14.4	14.9	**20.1**	15.6	15.4	**22.3**
Average Score	83.6	57.8	**85.6**	78.3	56.7	**81.1**

* Note that "Avg. Score" indicates the average results for NMI or Accuracy metrics.

4.6. Empirical convergence analysis

In this subsection, we further conduct empirical experiments to analyze the convergence property of our proposed ST-MSC approach. In particular, the convergence condition in Algorithm 1 are guaranteed by three error terms, i.e., the match error (ME), the iteration error for $\mathcal{B}$ ($\mathrm{IE}_{\mathcal{B}}$), and the iteration error for C (IE_C). Concretely, these terms can be defined in what follows:

$$\mathrm{ME} = \|\mathcal{B} - C\|_F \tag{61}$$

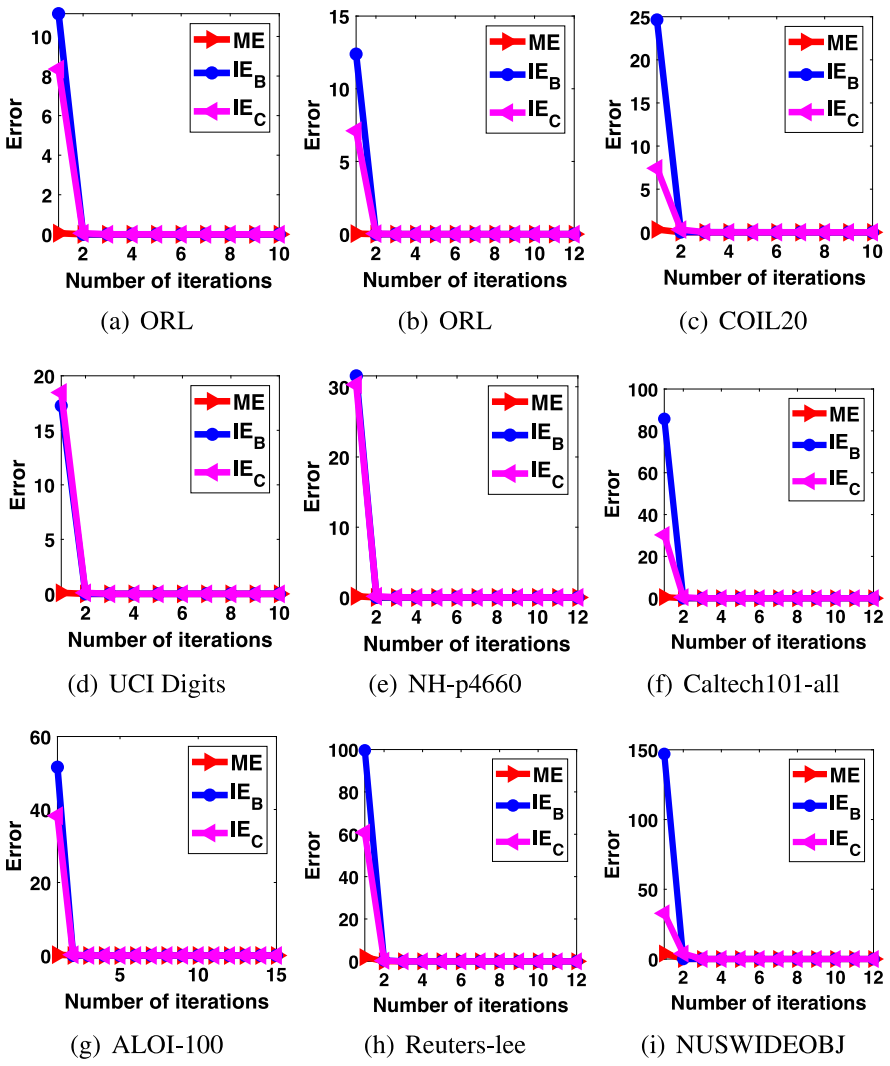


Fig. 9. Convergence analysis on nine benchmark datasets.

$$\mathrm{IE}_{\mathcal{B}} = \|\mathcal{B}_{current} - \mathcal{B}_{pre}\|_F \tag{62}$$

$$\mathrm{IE}_{C} = \|C_{current} - C_{pre}\|_F \tag{63}$$

The convergence cures w.r.t. different benchmarks are plotted in Fig. 9. As can be observed in Fig. 9, these three error values decrease rapidly and tend to 0 within a few iterations (usually less than 15 iterations), which indicates that our proposed approach enjoys relatively robust rapid convergence for different benchmark datasets.

5. Conclusion

This paper develops a new MSC approach termed ST-MSC, which inherits the advantage of the tri-factorization mechanism. Distinct from the existing approaches, our ST-MSC approach aims at discovering the deep anchor-sample relationship for scalable multi-view subspace clustering. In particular, the proposed approach is capable of flexibly exploiting the embedded low-dimensional spaces from various views, while simultaneously and robustly discovering high-order complementarity by tensorized anchor subspace learning. Extensive experiments have demonstrated the advantageous clustering performance of our ST-MSC approach on both general-scale and large-scale datasets. In the future work, a promising direction is to delve deeper into extending our proposed approach to the deep graph clustering scenarios.

CRediT authorship contribution statement

Guang-Yu Zhang: Writing – original draft, Methodology, Investigation, Data curation, Conceptualization. **Chang-Bin Guan:** Writing – review & editing, Software, Methodology, Investigation. **Dong Huang:** Writing – review & editing, Visualization, Supervision, Funding acquisition, Formal analysis. **Zihao Wen:** Writing – review & editing, Project administration, Methodology, Formal analysis. **Chang-Dong Wang:** Writing – review & editing, Validation, Project administration, Funding acquisition. **Lei Xiao:** Supervision, Investigation, Funding acquisition.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: The manuscript, or part of it, has neither been published nor is currently under consideration for publication by any other journal. All authors have read and approved the content. No disclosure, ethical or legal conflicts and no competing interests exist in this study. For the manuscript, there is not any potential financial or substantive conflict. We hope that it is in accordance with your publishing regulations and under consideration for publication once got approved.

Acknowledgments

This work was supported by the NSFC (62206099, 62476102 & 62276277), the Science and Technology Program of Guangzhou, China (2024A04J4451), and the Opening research fund of State Key Laboratory of Swine and Poultry Breeding Industry (2022GZ05).

Data availability

No data was used for the research described in the article.

References

[1] Anil K. Jain, Data clustering: 50 years beyond K-means, Pattern Recognit. Lett. 31 (8) (2010) 651–666.
[2] Dong Huang, Chang-Dong Wang, Jiansheng Wu, Jian-Huang Lai, Chee Keong Kwoh, Ultra-scalable spectral clustering and ensemble clustering, IEEE Trans. Knowl. Data Eng. 32 (6) (2020) 1212–1226.
[3] Feiping Nie, Guohao Cai, Xuelong Li, Multi-view clustering and semi-supervised classification with adaptive neighbours, in: AAAI Conference on Artificial Intelligence, 2017, pp. 2408–2414.
[4] You-Wei Liang, Dong Huang, Chang-Dong Wang, Philip S. Yu, Multi-view graph learning by joint modeling of consistency and inconsistency, IEEE Trans. Neural Netw. Learn. Syst. 35 (2) (2024) 2848–2862.
[5] Xuelong Li, Han Zhang, Rong Wang, Feiping Nie, Multiview clustering: A scalable and parameter-free bipartite graph fusion method, IEEE Trans. Pattern Anal. Mach. Intell. 44 (1) (2020) 330–344.
[6] Wei Xia, Quanxue Gao, Qianqian Wang, Xinbo Gao, Chris Ding, Dacheng Tao, Tensorized bipartite graph learning for multi-view clustering, IEEE Trans. Pattern Anal. Mach. Intell. 45 (4) (2023) 5187–5202.
[7] Suyuan Liu, Xinwang Liu, Siwei Wang, Xin Niu, En Zhu, Fast incomplete multi-view clustering with view-independent anchors, IEEE Trans. Neural Netw. Learn. Syst. (2022).
[8] Guang-Yu Zhang, Dong Huang, Chang-Dong Wang, Facilitated low-rank multi-view subspace clustering, Knowl.-Based Syst. (2022) 110141.
[9] Ben Yang, Xuetao Zhang, Zhiping Lin, Feiping Nie, Badong Chen, Fei Wang, Efficient and robust multi-view clustering with anchor graph regularization, IEEE Trans. Circuits Syst. Video Technol. (2022).
[10] Xinlei Chen, Deng Cai, Large scale spectral clustering with landmark-based representation, in: AAAI Conference on Artificial Intelligence, vol. 25, 2011, pp. 313–318, 1.
[11] Charu C. Aggarwal, Philip S. Yu, Outlier detection for high dimensional data, in: ACM SIGMOD International Conference on Management of Data, 2001, pp. 37–46.
[12] Gutha Jaya Krishna, Vadlamani Ravi, Outlier detection using evolutionary computing, in: International Conference on Informatics and Analytics, 2016, pp. 1–6.
[13] Gutha Jaya Krishna, Vadlamani Ravi, Keystroke based user authentication using modified differential evolution, in: TENCON 2019-2019 IEEE Region 10 Conference, TENCON, 2019, pp. 739–744.
[14] Gutha Jaya Krishna, Vadlamani Ravi, Anomaly detection using modified differential evolution: an application to banking and insurance, in: International Conference on Soft Computing and Pattern Recognition, 2021, pp. 102–111.
[15] Xiaosha Cai, Dong Huang, Guang-Yu Zhang, Chang-Dong Wang, Seeking commonness and inconsistencies: A jointly smoothed approach to multi-view subspace clustering, Inf. Fusion (ISSN: 1566-2535) 91 (2023) 364–375.
[16] Hongchang Gao, Feiping Nie, Xuelong Li, Heng Huang, Multi-view subspace clustering, in: IEEE International Conference on Computer Vision, 2015, pp. 4238–4246.
[17] Xiaochun Cao, Changqing Zhang, Huazhu Fu, Si Liu, Hua Zhang, Diversity-induced multi-view subspace clustering, in: IEEE Conference on Computer Vision and Pattern Recognition, 2015, pp. 586–594.
[18] Arthur Gretton, Olivier Bousquet, Alex Smola, Bernhard Schölkopf, Measuring statistical dependence with Hilbert–Schmidt norms, in: International Conference on Algorithmic Learning Theory, 2005, pp. 63–77.
[19] Changqing Zhang, Qinghua Hu, Huazhu Fu, Pengfei Zhu, Xiaochun Cao, Latent multi-view subspace clustering, in: IEEE Conference on Computer Vision and Pattern Recognition, 2017, pp. 4279–4287.
[20] Xiaobo Wang, Xiaojie Guo, Zhen Lei, Changqing Zhang, Stan Z Li, Exclusivity-consistency regularized multi-view subspace clustering, in: IEEE Conference on Computer Vision and Pattern Recognition, 2017, pp. 923–931.
[21] Shudong Huang, Zhao Kang, Ivor W. Tsang, Zenglin Xu, Auto-weighted multi-view clustering via kernelized graph learning, Pattern Recognit. 88 (2019) 174–184.
[22] Man-Sheng Chen, Ling Huang, Chang-Dong Wang, Dong Huang, Multi-view clustering in latent embedding space, in: AAAI Conference on Artificial Intelligence, 2020.
[23] Yuan Xie, Dacheng Tao, Wensheng Zhang, Yan Liu, Lei Zhang, Yanyun Qu, On unifying multi-view self-representations for clustering by tensor multi-rank minimization, Int. J. Comput. Vis. 126 (11) (2018) 1157–1179.
[24] Yuan Xie, Wensheng Zhang, Yanyun Qu, Longquan Dai, Dacheng Tao, Hyper-Laplacian regularized multilinear multiview self-representations for clustering and semisupervised learning, IEEE Trans. Cybern. 50 (2) (2018) 572–586.
[25] Qianqian Wang, Jiafeng Cheng, Quanxue Gao, Guoshuai Zhao, Licheng Jiao, Deep multi-view subspace clustering with unified and discriminative learning, IEEE Trans. Multimed. 23 (2020) 3483–3493.
[26] Zhibin Gu, Songhe Feng, Zhendong Li, Jiazheng Yuan, Jun Liu, NOODLE: Joint cross-view discrepancy discovery and high-order correlation detection for multi-view subspace clustering, ACM Trans. Knowl. Discov. Data 18 (6) (2024) 1–23.
[27] Dong Huang, Chang-Dong Wang, Jian-Huang Lai, Fast multi-view clustering via ensembles: Towards scalability, superiority, and simplicity, IEEE Trans. Knowl. Data Eng. 35 (11) (2023) 11388–11402.
[28] Jinghuan Lao, Dong Huang, Chang-Dong Wang, Jian-Huang Lai, Towards scalable multi-view clustering via joint learning of many bipartite graphs, IEEE Trans. Big Data 10 (1) (2024) 77–91.
[29] Si-Guo Fang, Dong Huang, Xiao-Sha Cai, Chang-Dong Wang, Chaobo He, Yong Tang, Efficient multi-view clustering via unified and discrete bipartite graph learning, IEEE Trans. Neural Netw. Learn. Syst. 35 (8) (2024) 11436–11447.
[30] Yeqing Li, Feiping Nie, Heng Huang, Junzhou Huang, Large-scale multi-view spectral clustering via bipartite graph, in: AAAI Conference on Artificial Intelligence, 2015, pp. 2750–2756.
[31] Zhao Kang, Wangtao Zhou, Zhitong Zhao, Junming Shao, Meng Han, Zenglin Xu, Large-scale multi-view subspace clustering in linear time, in: AAAI Conference on Artificial Intelligence, vol. 34, 2020, pp. 4412–4419, 04.
[32] Zhao Kang, Zhiping Lin, Xiaofeng Zhu, Wenbo Xu, Structured graph learning for scalable subspace clustering: From single view to multiview, IEEE Trans. Cybern. (2021).
[33] Mengjing Sun, Pei Zhang, Siwei Wang, Sihang Zhou, Wenxuan Tu, Xinwang Liu, En Zhu, Changjian Wang, Scalable multi-view subspace clustering with unified anchors, in: ACM International Conference on Multimedia, 2021, pp. 3528–3536.
[34] Siwei Wang, Xinwang Liu, Xinzhong Zhu, Pei Zhang, Yi Zhang, Feng Gao, En Zhu, Fast parameter-free multi-view subspace clustering with consensus anchor guidance, IEEE Trans. Image Process. 31 (2021) 556–568.
[35] Zuyuan Yang, Naiyao Liang, Wei Yan, Zhenni Li, Shengli Xie, Uniform distribution non-negative matrix factorization for multiview clustering, IEEE Trans. Cybern. 51 (6) (2020) 3249–3262.
[36] Xinhang Wan, Jiyuan Liu, Xinbiao Gan, Xinwang Liu, Siwei Wang, Yi Wen, Tianjiao Wan, En Zhu, One-step multi-view clustering with diverse representation, IEEE Trans. Neural Netw. Learn. Syst. (2024).
[37] Zhibin Gu, Songhe Feng, Ruiting Hu, Gengyu Lyu, ONION: Joint unsupervised feature selection and robust subspace extraction for graph-based multi-view clustering, ACM Trans. Knowl. Discov. Data 17 (5) (2023) 1–23.
[38] Man-Sheng Chen, Chang-Dong Wang, Dong Huang, Jian-Huang Lai, S Yu Philip, Concept factorization based multiview clustering for large-scale data, IEEE Trans. Knowl. Data Eng. (2024).
[39] Naiyao Liang, Zuyuan Yang, Shengli Xie, Incomplete multi-view clustering with sample-level auto-weighted graph fusion, IEEE Trans. Knowl. Data Eng. 35 (6) (2022) 6504–6511.
[40] Misha E. Kilmer, Karen Braman, Ning Hao, Randy C. Hoover, Third-order tensors as operators on matrices: A theoretical and computational framework with applications in imaging, SIAM J. Matrix Anal. Appl. 34 (1) (2013) 148–172.
[41] Oguz Semerci, Ning Hao, Misha E. Kilmer, Eric L. Miller, Tensor-based formulation and nuclear norm regularization for multienergy computed tomography, IEEE Trans. Image Process. 23 (4) (2014) 1678–1693.
[42] Zhibin Gu, Zhendong Li, Songhe Feng, Topology-driven multi-view clustering via tensorial refined sigmoid rank minimization, in: ACM SIGKDD Conference on Knowledge Discovery and Data Mining, 2024, pp. 920–931.
[43] Feiping Nie, Xiaoqian Wang, Cheng Deng, Heng Huang, Learning a structured optimal bipartite graph for co-clustering, Adv. Neural Inf. Process. Syst. 30 (2017).
[44] Shuai Chang, Jie Hu, Tianrui Li, Hao Wang, Bo Peng, Multi-view clustering via deep concept factorization, Knowl.-Based Syst. 217 (2021) 106807.
[45] Zhao Kang, Zhanyu Liu, Shirui Pan, Ling Tian, Fine-grained attributed graph clustering, in: SIAM International Conference on Data Mining, 2022, pp. 370–378.
[46] Zhouchen Lin, Huan Li, Cong Fang, Alternating Direction Method of Multipliers for Machine Learning, Springer, 2022.
[47] Wenrui Hu, Dacheng Tao, Wensheng Zhang, Yuan Xie, Yehui Yang, The twist tensor nuclear norm for video completion, IEEE Trans. Neural Netw. Learn. Syst. 28 (12) (2016) 2961–2973.
[48] Haizhou Yang, Quanxue Gao, Wei Xia, Ming Yang, Xinbo Gao, Multiview spectral clustering with bipartite graph, IEEE Trans. Image Process. 31 (2022) 3591–3605.
[49] Zhibin Gu, Zhendong Li, Songhe Feng, EDISON: Enhanced dictionary-induced tensorized incomplete multi-view clustering with Gaussian error rank minimization, in: International Conference on Machine Learning, 2024.

[50] Jing Li, Quanxue Gao, Qianqian Wang, Ming Yang, Wei Xia, Orthogonal non-negative tensor factorization based multi-view clustering, Adv. Neural Inf. Process. Syst. 36 (2024).
[51] Youwei Liang, Dong Huang, Chang-Dong Wang., Consistency meets inconsistency: A unified graph learning framework for multi-view clustering, in: IEEE International Conference on Data Mining, 2019.
[52] Feiping Nie, Jing Li, Xuelong Li, Self-weighted multiview clustering with multiple graphs, in: International Joint Conference on Artificial Intelligence, 2017, pp. 2564–2570.
[53] Zi-Feng Zhou, Dong Huang, Chang-Dong Wang, Pyramid contrastive learning for clustering, Neural Netw. 185 (2025) 107217.
[54] Dong Huang, Chang-Dong Wang, Hongxing Peng, Jianhuang Lai, Chee-Keong Kwoh, Enhanced ensemble clustering via fast propagation of cluster-wise similarities, IEEE Trans. Syst. Man Cybern.: Syst. 51 (1) (2021) 508–520.
[55] Man-Sheng Chen, Chang-Dong Wang, Dong Huang, Jian-Huang Lai, Philip S Yu, Efficient orthogonal multi-view subspace clustering, in: ACM SIGKDD Conference on Knowledge Discovery and Data Mining, 2022, pp. 127–135.
[56] Zhen Long, Qiyuan Wang, Yazhou Ren, Yipeng Liu, Ce Zhu, S2MVTC: a simple yet efficient scalable multi-view tensor clustering, in: IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2024, pp. 26213–26222.

Unified and Tensorized Incomplete Multi-View Kernel Subspace Clustering

Guang-Yu Zhang, Dong Huang, *Member, IEEE*, and Chang-Dong Wang, *Senior Member, IEEE*

***Abstract*—Incomplete multi-view clustering (IMC) has recently received widespread attention in the field of clustering analysis. In spite of the great success, we observe that the current IMC approaches are still faced with three common demerits. First, they mostly fail to recover the inherent (especially nonlinear) subspace structure during incomplete clustering procedure. Second, these approaches tend to design the objective function by some specific matrix norms, yet often overlook the high-level correlation embedded in heterogeneous views. Third, many of them follow a two-stage framework, which inevitably leads to the sub-optimal clustering result due to the lack of the ability of joint optimization. To overcome these demerits, we develop a novel approach termed Unified and Tensorized Incomplete Multi-view Kernel Subspace Clustering (UT-IMKSC) in this paper. Specifically, a kernelized incomplete subspace clustering framework is formulated to exploit the inherent subspace structure from multiple views. In this framework, we aim to impute the incomplete kernels and perform incomplete subspace clustering simultaneously, upon which the low-rank tensor representations as well as their affinity matrix can be seamlessly achieved in a one-step manner. This unified formulation enables our approach to recover the latent relationship among observed and unobserved samples, while capturing the high-level correlation for strengthened subspace clustering. To the best of our knowledge, our approach is the first attempt to formulate incomplete multi-view kernel subspace clustering from unified and tensorized perspectives. Extensive experiments are conducted on various incomplete multi-view datasets, which have demonstrated the superiority of our approach over the state-of-the-art.**

***Index Terms*—Data clustering, incomplete multi-view clustering, tensorized kernel subspace clustering, one-step framework.**

I. Introduction

RECENT years have witnessed the explosive emergence of multi-view data in scientific and engineering problems. Specifically, the collected multi-view data may suffer from missing views (or partial views) for various reasons, such as human errors and environmental noise, which naturally leads to the incomplete multi-view data in practice. For instance, in document processing, some news stories are only reported by either Chinese or English, while some may have both two forms of languages. In pattern recognition, some object images may have both visual and textual features, while some other object images perhaps only have one of these two feature sets. The incomplete multi-view data has brought great interests to the incomplete multi-view clustering (IMC) research. Compared with the traditional tools of clustering analysis, IMC aims to obtain more satisfactory performance by mining both complete and incomplete information across multiple views.

In the present development, the prominent directions for IMC research broadly lie in three categories, i.e., the matrix factorization based approaches, the graph-based approaches as well as the kernel k-means based approaches. Among these three categories, the matrix factorization based approaches have been widely studied due to their scalable and well-defined framework. These approaches rely on pursuing a shared representation contained in complete and incomplete views. By utilizing different techniques, a variety of matrix factorization-based approaches have been proposed, typically including [1], [2], [3], [4], among which the incomplete multi-view clustering problem is generally transformed into a low-dimensional subspace learning problem. The existing matrix factorization-based approaches have made impressive progress, which, however, still have one vital drawback in common. That is, they usually ignore the underlying geometry structure (especially the topology graph structure) embedded in heterogeneous views, which limits their ability to model the underlying relationship among observed and missing samples.

Beyond the aforementioned approaches, the graph-based approaches provide another promising tool for solving the IMC problem. For these approaches, the key objective is to capture the topology similarity (also known as spectral embedding) for robust incomplete clustering. To this end, several endeavors are devoted to the development of graph-based approaches [5], [6], [7], [8]. Benefiting from the well-known graph spectral theory, such graph-based approaches have aroused widespread research interest, especially for the nonlinearly separable IMC scenarios. Following this line, some researchers attempt to incorporate the kernel clustering scheme and present a series of kernel k-means based approaches, mainly involving [9], [10], [11]. Compared to the traditional two-stage approaches, these kernel-based approaches have exhibited the significant advantages in

Manuscript received 5 March 2023; revised 28 June 2023 and 10 August 2023; accepted 17 September 2023. Date of publication 30 January 2024; date of current version 27 March 2024. This work was supported in part by NSFC under Grants 62206099 and 62276277, in part by the Guangdong Basic and Applied Basic Research Foundation under Grant 2022B1515120059, in part by the Science and Technology Program of Guangzhou, China under Grant 2024A04J4451, and in part by the Guangdong Provincial Key Laboratory of Intellectual Property and Big Data under Grant 2018B030322016. *(Correspond author: Dong Huang.)*

Guang-Yu Zhang and Dong Huang are with the College of Mathematics and Informatics, South China Agricultural University, Guangzhou 510642, China (e-mail: guangyuzhg@foxmail.com; huangdonghere@gmail.com).

Chang-Dong Wang is with the School of Computer Science and Engineering, Sun Yat-sen University, Guangzhou 510006, China, and also with the Guangdong Provincial Key Laboratory of Intellectual Property and Big Data Guangzhou 510665, China (e-mail: changdongwang@hotmail.com).

Recommended for acceptance by M. Zhang.

Digital Object Identifier 10.1109/TETCI.2024.3353576

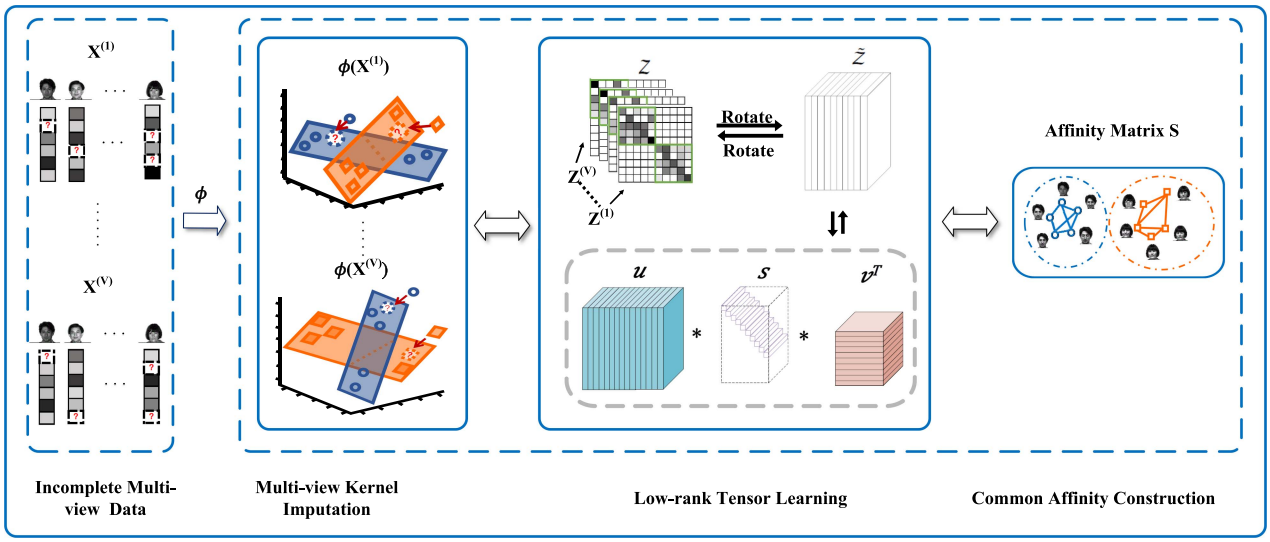


Fig. 1. Flowchart of the proposed approach.

many challenging IMC tasks. By deriving from three viewpoints, considerable advances have been achieved in the recent IMC literature. Nevertheless, there still remains following three vital questions. First of all, how to break through the traditional practice, by devising a new paradigm for incomplete kernel subspace clustering? Second, how to model the inherent high-level correlation contained in complete and incomplete views? Last but not least, how to unify the isolated learning process of subspace representations and common affinity matrix, so as to enhance the incomplete multi-view clustering under a one-step framework?

In light of these vital questions, we present a novel approach for incomplete multi-view clustering, named **U**nified and **T**ensorized **I**ncomplete **M**ulti-view **K**ernel **S**ubspace **C**lustering (UT-IMKSC), as shown in Fig. 1. Notably, a kernelized incomplete subspace clustering framework is formulated in our approach, which ensures the versatile subspace structure among complete and incomplete views can be well preserved. With this framework formulated, the proposed approach unifies the multi-view kernel imputation with incomplete kernel subspace clustering jointly, where the low-rank tensor as well as common affinity matrix are harmlessly learned in a one-step manner. By means of this, our UT-IMKSC approach can extract the hidden relationship among observed and unobserved samples, while discovering the high-level complementary information for better incomplete clustering. Starts from the unified and tensorized perspectives, as far as we know, this would be the first practice for incomplete multi-view kernel subspace clustering. In summary, this paper has following three contributions:

- Instead of following the conventional thoughts, this paper designs a novel paradigm to bridge the gap between incomplete multi-view clustering and tensorized kernel subspace clustering.
- Going beyond the traditional practice, this paper integrates the kernel imputation, low-rank tensor learning and common affinity construction into a unified objective function, to solve which an efficient optimization algorithm is further designed.
- Experimental results on six incomplete multi-view datasets demonstrate the clustering robustness and effectiveness of our proposed approach over the state-of-the-art IMC approaches.

The rest of this paper is organized in what follows. Section II briefly reviews the related works on IMC approaches. Then, Section III is devised to describe the proposed UT-IMKSC approach. Next, Section IV reports the experimental results on several incomplete multi-view datasets. Finally, Section V provides the conclusion of this paper.

II. Related Work

This section will introduce several representatives and recent developments on multi-view subspace clustering (MSC) and incomplete multi-view clustering (IMC).

In the literature, many MSC works [12], [13] [14] originated from the singe-view subspace (or graph) learning models [15], [16]. For example, Wang et al. [17] made an attempt to solve the multi-view subspace clustering problem, which extends the single-view subspace clustering model to multi-view domain by enhancing the low-rank structure across diverse views. Additionally, Zhang et al. [18] proposed another novel subspace-based approach for multi-view clustering that learns the shared latent representation from multiple views. Inspired by the success of one-pass scheme in feature learning [19], [20], Zhang et al. [21] developed a one-step kernel subspace approach for multi-view clustering, whose idea is to integrate the kernel subspace clustering and affinity matrix construction into a joint optimization procedure. Along this line, Zhang et al. [22] proposed a consensus multi-view subspace approach to directly generate the discrete clustering label. By introducing the

structured graph learning technique, Kang et al. [23] designed a novel subspace clustering approach for large-scale multi-view problem. Inspired by the great success of deep neural networks in computer version [24], [25], Wang et al. [26] further developed a fast multi-view subspace clustering approach guided by the consensus anchors of multiple views. Recently, Kang et al. [27] presented a novel multi-view subspace clustering approach by means of the attributed graph filter. Wang et al. [28] designed a deep multi-view subspace clustering approach via unified and discriminative learning. It can not only discover the local structure from individual views, but also extract the discriminative constraints of all views. Further, Wang et al. [29] proposed a fast self-supervised scheme for deep multi-view subspace clustering. This approach can capture both low-level and high-level structured features to fully explore the comprehensive complementarity across different views. For more detailed review on multi-view subspace clustering, please see the literature review in [30].

The increasing availability of incomplete multi-view data gives rise to the emergence of incomplete multi-view clustering (IMC). Many existing IMC researches concentrate on the matrix factorization-based approaches. In [1], Li et al. design a novel matrix factorization approach named PVC. As an early attempt, the PVC approach seeks to learn a common representation for paired samples and obtain two individual representations for unpaired samples. Inspired by this work, Zhao et al. [2] devised an NMF-style approach for incomplete multi-view clustering, upon which the global graph structure is captured within the low-dimensional embedded subspace. Inheriting from the two-stage strategy, Shao et al. [31] attempted to fill the missing values with the help of observed samples, and then extract a latent subspace by the weighted NMF technique. In [4], Hu et al. introduced a novel matrix factorization approach, so as to pursue a nonnegative consensus representation by Laplacian graph regularizer. In the meantime, Hu et al. [3] proposed a one-pass framework to directly obtain the clustering labels, without needing any postprocessing step. Recently, Liu et al. [32] devised a consensus learning approach, which can efficiently uncover the complementary information from the observed and missing samples. Besides, Yin et al. [33] incorporated the cosine similarity to a matrix factorization model, so as to preserve the manifolds of complete and partial views. However, most of them still ignore the similarity graph structure embedded in multiple views, making them inconvenient to uncover the latent relationship (or topology structure) between observed and missing samples.

Besides the matrix-based approaches, several researchers intend to leverage the latent graph learning technique, which renders another effective tool for the clustering analysis of incomplete multi-view data. In [5], Wen et al. designed a pioneering work for graph-based approach termed IMSC-AGL. This approach incorporates the co-regularization constraint to adaptively obtain a consistent spectral embedding from heterogenous views. By flexibly combining the graph learning and matrix factorization techniques, Wen et al. [8] proposed a novel graph-based approach from generalized perspective. Following this representative, Li et al. [34] jointly employed graph learning and partition fusion for enhancing incomplete clustering performance. Very lately, Yu et al. [35] devised to solve the large-scale IMC problem, by efficiently characterizing the sample-level relationship hidden in incomplete multi-view data. Different from the matrix factorization based approaches as well as graph-based approaches, another notable technique used in IMC scenario relies on kernel k-means clustering. By means of the novel one-stage strategy, to the best of our knowledge, Liu et al. [9] devised the first work on incomplete multiple kernel clustering. Based on this work, Li et al. further consider the local structure hidden in incomplete kernels, and meanwhile present a robust approach termed IK-MKKM [11]. In addition, Liu et al. [36] derived an effective and efficient IMC approach at the partition level. Though remark progress has been achieved, yet the performances of current IMC approaches can be further facilitated by following three considerations. (i) explore the latent subspace structure embedded in the incomplete multi-view data. (ii) enhance the high-level complementarity for better incomplete clustering performance. (iii) formulate a unified framework to directly obtain the clustering-friendly affinity matrix.

III. Proposed UT-IMKSC Approach

This section describes our proposed approach in a stepwise way. Particularly, the primary notations (including t-SVD tensor norm) are introduced in Section III-A. Following this line, the objective formulation of our UT-IMKSC approach is provided in Section III-B. Then an efficient optimization algorithm is derived in Section III-D. Finally, the time complexity is analyzed in Section III-E.

A. Primary Notations

To simplify notation, we represent the calligraphy letters for 3-mode tensors (i.e., $\mathcal{J}$) and upper-case letters for matrices (i.e., J). In particular, the i-th row, j-th column, and ij-th element of the matrix J are deployed by $\mathrm{J}_{i:}$, $\mathrm{J}_{:j}$ and J_{ij}, respectively. In addition, the i-th horizontal slice, j-th lateral slice and k-th frontal slice of the 3-mode tensor $\mathcal{J} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$ are deployed by $\mathcal{J}(i,:,:)$, $\mathcal{J}(:,j,:)$ and $\mathcal{J}(:,:,k)$, respectively. The operator $\overline{\mathcal{J}} = \mathrm{fft}(\mathcal{J}, [], 3)$ indicates the DFT transform towards the 3-th dimension of tensor $\mathcal{J}$, while the inverse DFT of tensor $\mathcal{J}$ can be deployed by $\mathcal{J} = \mathrm{ifft}(\overline{\mathcal{J}}, [], 3)$. Notice that, the matrix $\mathrm{J}^{(k)}$ is utilized to represent the frontal slice $\mathcal{J}(:,:,k)$.

In terms of the aforementioned notations, necessary block-based operators and the basic definitions that are relevant to t-SVD are given in detail. To begin with, we can define the following block circulant matrix for any given tensor $\mathcal{J} \in \mathcal{R}^{n_1 n_3 \times n_2 n_3}$.

$$\mathrm{bcirc}(\mathcal{J}) = \begin{bmatrix} \mathrm{J}^{(1)} & \mathrm{J}^{(n_3)} & \cdots & \mathrm{J}^{(2)} \\ \mathrm{J}^{(2)} & \mathrm{J}^{(1)} & \cdots & \mathrm{J}^{(3)} \\ \vdots & \ddots & \ddots & \vdots \\ \mathrm{J}^{(n_3)} & \mathrm{J}^{(n_3-1)} & \cdots & \mathrm{J}^{(1)} \end{bmatrix}.$$

The *unfold* and *fold* operations of tensor $\mathcal{J}$ could be given in what follows.

$$\text{unfold}(\mathcal{J}) = \begin{bmatrix} \mathrm{J}^{(1)} \\ \mathrm{J}^{(2)} \\ \vdots \\ \mathrm{J}^{(n_3)} \end{bmatrix}, \ \text{fold}(\text{unfold}(\mathcal{J})) = \mathcal{J}.$$

Furthermore, the *bdiag* and *bdfold* operations of tensor $\mathcal{J}$ are given in the following.

$$\text{bdiag}(\mathcal{J}) = \begin{bmatrix} \mathrm{J}^{(1)} & & \\ & \mathrm{J}^{(2)} & \\ & & \ddots \\ & & \mathrm{J}^{(n_3)} \end{bmatrix},$$

$$\text{bdfold}(\text{bdaig}(\mathcal{J})) = \mathcal{J}. \tag{1}$$

Next, we give some important definitions that are related to t-SVD in what follows.

Definition 1. (f-diagonal tensor [37]): A tensor is called f-diagonal if each of its frontal slices is diagonal matrix.

Definition 2. (Identity tensor [37]): For the identity tensor $\mathcal{I} \in \mathbb{R}^{n\times n\times n_3}$, its first frontal slice is the identity matrix with size $n \times n$, and all other frontal slices are zero.

Definition 3. (Tensor transpose [37]): For any tensor $\mathcal{X} \in \mathbb{R}^{n_1\times n_2\times n_3}$, its transpose tensor $\mathcal{X}^T \in \mathbb{R}^{n_2\times n_1\times n_3}$ can be obtained by transposing each frontal slice of $\mathcal{X}$ and then reversing the order of the transposed frontal slices 2 through n_3.

Definition 4. (Orthogonal tensor [37]): A tensor $\mathcal{J} \in \mathbb{R}^{n\times n\times n_3}$ is orthogonal if it satisfies

$$\mathcal{J}^T * \mathcal{J} = \mathcal{J} * \mathcal{J}^T = \mathcal{I}. \tag{2}$$

where "$*$" denotes the t-product.

Definition 5. (t-product [37]): Given any two tensors $\mathcal{P} \in \mathbb{R}^{n_1\times n_2\times n_3}$ and $\mathcal{Q} \in \mathbb{R}^{n_2\times n_4\times n_3}$, the corresponding t-product $\mathcal{J} = \mathcal{P} * \mathcal{Q}$ is a tensor of size $n_1 \times n_4 \times n_3$,

$$\mathcal{J} = \mathcal{P} * \mathcal{Q} = \text{fold}(\text{bcirc}(\mathcal{P}) \cdot \text{unfold}(\mathcal{Q})). \tag{3}$$

Definition 6. (t-SVD [37]): Given a tensor $\mathcal{J} \in \mathbb{R}^{n_1\times n_2\times n_3}$, the corresponding t-SVD could be denoted as

$$\mathcal{J} = \mathcal{W} * \mathcal{O} * \mathcal{Q}^T, \tag{4}$$

in which $\mathcal{W} \in \mathbb{R}^{n_1\times n_1\times n_3}$ and $\mathcal{Q} \in \mathbb{R}^{n_2\times n_2\times n_3}$ are two orthogonal tensors, and $\mathcal{O} \in \mathbb{R}^{n_1\times n_2\times n_3}$ is a f-diagonal tensor.

Definition 7. (t-SVD based tensor nuclear norm [37]): Given a tensor $\mathcal{J} \in \mathbb{R}^{n_1\times n_2\times n_3}$, the corresponding t-SVD based tensor nuclear norm (i.e., $\| \mathcal{J} \|_{\circledast}$) is defined by the sum of the singular values of all frontal slices, which are as follows:

$$\| \mathcal{J} \|_{\circledast} = \sum_{k=1}^{n_3} \| \mathcal{J}_f^{(k)} \|_* = \sum_{i=1}^{\min(n_1,n_2)} \sum_{k=1}^{n_3} | O_f^{(k)}(i,i) | \tag{5}$$

in which $O_f^{(k)}(i,i)$ is computed by the t-SVD of frontal slices of $\mathcal{J}_f$, i.e., $\mathcal{J}_f^{(k)} = \mathcal{W}_f^{(k)} \mathcal{O}_f^{(k)} \mathcal{Q}_f^{(k)T}$. Especially, Fig. 2 presents the t-SVD sample of a tensor $\mathcal{J} \in \mathbb{R}^{n_1\times n_2\times n_3}$.

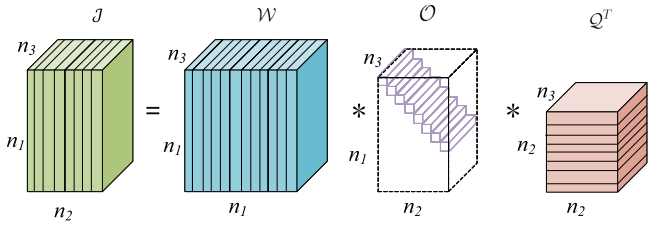


Fig. 2. t-SVD sample of a tensor $\mathcal{J} \in \mathbb{R}^{n_1\times n_2\times n_3}$.

B. The Objective Formulation

Before our objective formulation is described, we briefly review the multi-view subspace clustering (MSC) in advance. In the MSC setting, given a multi-view dataset $\mathrm{X}_{\text{data}} = [\mathrm{X}^{(1)}; \ldots; \mathrm{X}^{(V)}]$ with V views and n data samples, where the v-th view data sub-matrix is denoted as $\mathrm{X}^{(v)} \in \mathbb{R}^{d_v\times n}$. By extending from single-view subspace learning to multiple views, the current MSC studies use all the n data samples as the dictionary. This means that the data samples lie in a collection of low-dimension subspaces extracted from original high-dimension spaces. Therefore, the data samples of the v-th view can be described by a linear combination of remaining samples, which leads to the following self-expressive formulation:

$$\mathrm{X}^{(v)} \approx \mathrm{X}^{(v)}(\mathrm{Z}^{(v)}), \forall v, \tag{6}$$

in which $\mathrm{Z}^{(v)} \in \mathbb{R}^{n\times n}$ indicates the subspace representation of the v-th view. Following this formulation, the basic model for MSC could be formulated in following form:

$$\min_{\{\mathrm{Z}^{(v)}\}_{v=1}^V} \sum_{v=1}^V \left\| \mathrm{X}^{(v)} - \mathrm{X}^{(v)}(\mathrm{Z}^{(v)}) \right\|_F^2 + \sum_{v=1}^V \Omega(\mathrm{Z}^{(v)}) \tag{7}$$

where $\Omega(.)$ is a specially-designed regularization term for $\{\mathrm{Z}^{(v)}\}_{v=1}^V$. As indicated by related literatures, the general MSC approaches have ability to preserve the structural subspace within each individual view. Yet their clustering robustness may be degraded when handling nonlinearly separable tasks. For this reason, we provide a flexible solution by leveraging with the kernel learning, which comes to the multi-view kernel subspace clustering (MKSC) scenario. First of all, if a nonlinear mapping $\phi(.)$ is guided by a predefined kernel function, the whole data samples (of the v-th view) could be linearly characterized in the kernel space, that is

$$\phi(\mathrm{X}^{(v)}) \approx \phi(\mathrm{X}^{(v)})(\mathrm{Z}^{(v)}), \forall v, \tag{8}$$

Furthermore, the kernel extension of the MSC model can be given as follows:

$$\min_{\{\mathrm{Z}^{(v)}\}_{v=1}^V} \sum_{v=1}^V \left\| \phi(\mathrm{X}^{(v)}) - \phi(\mathrm{X}^{(v)})(\mathrm{Z}^{(v)}) \right\|_F^2 + \sum_{v=1}^V \Omega(\mathrm{Z}^{(v)}) \tag{9}$$

Notice that the reconstruction error term could be rewritten in the trace form:

$$\begin{aligned}&\sum_{v=1}^{V}||\phi(\mathrm{X}^{(v)})-\phi(\mathrm{X}^{(v)})\mathrm{Z}^{(v)}||_F^2\\&=\sum_{v=1}^{V}\mathrm{Tr}\{\mathrm{K}^{(v)}-2\mathrm{K}^{(v)}\mathrm{Z}^{(v)}+{\mathrm{Z}^{(v)}}^T\mathrm{K}^{(v)}\mathrm{Z}^{(v)}\}\\&=\sum_{v=1}^{V}\mathrm{Tr}\{\mathrm{K}^{(v)}(\mathrm{I}-\mathrm{Z}^{(v)}-{\mathrm{Z}^{(v)}}^T+\mathrm{Z}^{(v)}{\mathrm{Z}^{(v)}}^T)\}\end{aligned}\tag{10}$$

Here $\mathrm{K}^{(v)}$ stands for the kernel matrix of the v-th view, which can be calculated by the kernel trick, i.e., $\mathrm{K}^{(v)}=\phi(\mathrm{X}^{(v)})^T\phi(\mathrm{X}^{(v)})$. As shown in (9), this model has the advantage in handling more complex clustering problem, i.e., multi-view data with nonlinear distributions.

Following this line, we proceed to the incomplete multi-view kernel subspace clustering (IMKSC). Formally, suppose p_v be the index set for the v-th view observed samples. Similarity, let $\mathrm{K}_p^{(v)}$ be the v-th view kernel sub-matrix that computed by these observed samples. It is noteworthy that one-stage based approaches yield more robust clusterings over the traditional two-stage based approaches. Thereby, we propose to integrate kernel imputation as well as kernel subspace clustering into a unified optimization framework. Mathematically, this presentation leads to the following optimization problem:

$$\begin{aligned}\min_{\{\mathrm{Z}^{(v)},\mathrm{K}^{(v)}\}_{v=1}^{V}}\ &\sum_{v=1}^{V}\mathrm{Tr}\{\mathrm{K}^{(v)}(\mathrm{I}-\mathrm{Z}^{(v)}-{\mathrm{Z}^{(v)}}^T+\mathrm{Z}^{(v)}{\mathrm{Z}^{(v)}}^T)\}\\&+\sum_{v=1}^{V}\Omega(\mathrm{Z}^{(v)})\\\text{s.t. }&\mathrm{K}^{(v)}(p_v,p_v)=\mathrm{K}_p^{(v)},\mathrm{K}^{(v)}\succeq\mathbf{0},\forall v,\end{aligned}\tag{11}$$

in which $\mathbf{0}\in\mathbb{R}^{n\times n}$ represents the zero matrix. The constraint $\mathrm{K}^{(v)}(p_v,p_v)=\mathrm{K}_p^{(v)}$ is used to ensure that $\mathrm{K}^{(v)}$ maintains the observed entries during the optimization. Inspired by the great success of low-rank tensor learning in recent years [38], [39], [40], we further employ the t-SVD-based tensor constraint on the subspace representations, aiming to explore the high-level complementarity among complete and incomplete views. The corresponding model turns out to be:

$$\begin{aligned}\min_{\{\mathrm{Z}^{(v)},\mathrm{K}^{(v)}\}_{v=1}^{V}}\ &\sum_{v=1}^{V}\mathrm{Tr}\{\mathrm{K}^{(v)}(\mathrm{I}-\mathrm{Z}^{(v)}-{\mathrm{Z}^{(v)}}^T+\mathrm{Z}^{(v)}{\mathrm{Z}^{(v)}}^T)\}\\&+\lambda_1\|\mathcal{Z}\|_{\circledast}\\\text{s.t. }&\mathrm{K}^{(v)}(p_v,p_v)=\mathrm{K}_p^{(v)},\mathrm{K}^{(v)}\succeq\mathbf{0},\forall v,\\&\mathcal{Z}=\Phi(\mathrm{Z}^{(1)},\mathrm{Z}^{(2)},\ldots,\mathrm{Z}^{(V)}).\end{aligned}\tag{12}$$

The $\Psi(\cdot)$ is an operator that transforms the multi-view subspace representations $\{\mathrm{Z}^{(v)}\}_{v=1}^{V}$ into a 3-order tensor $\mathcal{Z}$, and then rotates its size to $n\times V\times n$, as illustrated in Fig. 3. In the graph learning context, it has been proven that the cluster (or topology) structure can be significantly enhanced by removing

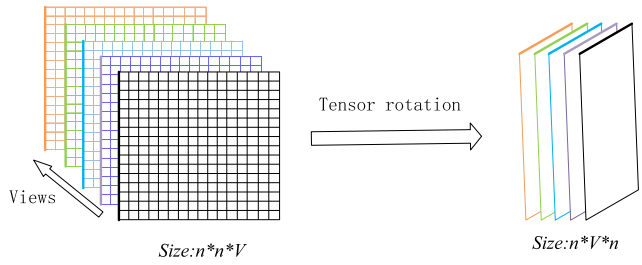


Fig. 3. Example of rotation operation.

the redundant links of fully-connected graph. Inheriting from this, we expect the common affinity matrix should obey the sparse property of near neighbor graph, by revealing the neighborhood relationship among multi-view subspace representations. Consequently, the regularizer term for common affinity matrix can be formulated as follows:

$$\begin{aligned}&\min_{\mathrm{S}}\sum_{v=1}^{V}\sum_{i,j=1}^{n}\|{\mathrm{Z}_{:i}}^{(v)}-{\mathrm{Z}_{:j}}^{(v)}\|_2^2\mathrm{S}_{ij}\\\Leftrightarrow&\min_{\mathrm{S}}\sum_{v=1}^{V}\mathrm{Tr}(\mathrm{Z}^{(v)}\mathrm{L}_S{\mathrm{Z}^{(v)}}^T)\\&\text{s.t. }\mathrm{S}\mathbf{1}=\mathbf{1},\mathrm{S}\succeq\mathbf{0},\|\mathrm{S}_{i:}\|_0=k,\forall i.\end{aligned}\tag{13}$$

Here $\mathrm{S}\mathbf{1}=\mathbf{1},\mathrm{S}\succeq\mathbf{0}$ indicates the simplex constraint on the affinity matrix S, and $\mathbf{1}\in\mathbb{R}^{n\times 1}$ denotes an all-one column vector. The constraint $\|\mathrm{S}_{i:}\|_0=k,\forall i$ is used to impose the sparsity of affinity matrix S, which ensures every row has exactly k nonzero elements (also known as the number of near neighbors). Besides, $\mathrm{L}_S=\mathrm{W}-\mathrm{S}$ is the Laplacian matrix for common affinity matrix S, here W is the diagonal matrix with its diagonal element $\mathrm{W}_{ii}=\sum_{j=1}^{n}\mathrm{S}_{ij}$.

Moving forward, we propose to directly obtain a clustering-friendly affinity matrix for better clustering performance, upon which the learning process of low-rank tensor representation and their affinity matrix are mutually enhanced in a one-step manner. According to this consideration, we could formulate the objective function of our UT-IMKSC approach as follows:

$$\begin{aligned}\min_{\{\mathrm{Z}^{(v)},\mathrm{K}^{(v)}\}_{v=1}^{V},\mathrm{S}}\ &\sum_{v=1}^{V}\mathrm{Tr}\{\mathrm{K}^{(v)}(\mathrm{I}-\mathrm{Z}^{(v)}-{\mathrm{Z}^{(v)}}^T+\mathrm{Z}^{(v)}{\mathrm{Z}^{(v)}}^T)\}\\&+\lambda_1\|\mathcal{Z}\|_{\circledast}+\lambda_2\sum_{v=1}^{V}\mathrm{Tr}(\mathrm{Z}^{(v)}\mathrm{L}_S{\mathrm{Z}^{(v)}}^T)\\\text{s.t. }&\mathrm{K}^{(v)}(p_v,p_v)=\mathrm{K}_p^{(v)},\mathrm{K}^{(v)}\succeq\mathbf{0},\forall v,\\&\mathcal{Z}=\Phi(\mathrm{Z}^{(1)},\mathrm{Z}^{(2)},\ldots,\mathrm{Z}^{(V)}),\\&\mathrm{S}\mathbf{1}=\mathbf{1},\mathrm{S}\succeq\mathbf{0},\|\mathrm{S}_{i:}\|_0=k,\forall i\end{aligned}\tag{14}$$

where λ_1 and λ_2 are two positive hyper-parameters, k is a positive integer parameter for common affinity matrix S. As shown in (14), there are three terms that devote to the different merits of our objective function. Specifically, the first term (i.e., *the nonlinear reconstruction term*) formulates a general

paradigm for incomplete multi-view kernel subspace clustering. The second term concerns with *the low-rank tensor learning*, aiming to enhance the inherent high-level complementarity across complete and incomplete views. The third term concerns with *the common affinity matrix learning*, whose purpose relies on capturing the discriminative cluster structure from the tensor subspace representations. With our objective function being minimized, the multi-view kernel imputation, the tensor subspace representation and their common affinity matrix learning are beneficial to each other in a mutual promotion manner. By this means, the learned affinity matrix (can be viewed as a latent neighborhood graph) has a great potential to yield promising incomplete clustering performance. These advantageous points contribute to the superior performance of our proposed approach, which have been experimentally validated on several incomplete multi-view datasets. (Please refer to Section IV for detailed information).

C. Further Analysis

In this section, we attempt to analyze the proposed one-step scheme from a theoretical perspective. First, the definition of *group effect* is given as follows:

Definition 8. (Group Effect [41]): Given a set of data samples $\mathrm{X} = [\mathrm{X}_{:1}, \mathrm{X}_{:2}, \ldots, \mathrm{X}_{:n}] \in \mathbb{R}^{d\times n}$, the corresponding self-representation matrix $\mathrm{Z} = [\mathrm{Z}_{:1}, \mathrm{Z}_{:2}, \ldots, \mathrm{Z}_{:n}] \in \mathbb{R}^{n\times n}$ has the grouping effect if $||\mathrm{X}_{:i} - \mathrm{X}_{:j}||_2 \to 0 \Rightarrow ||\mathrm{Z}_{:i} - \mathrm{Z}_{:j}||_2 \to 0, \forall i \neq j$.

In the single-view subspace clustering, the smooth regularizer has shown its advantage in preserving the local group effect. Following this idea, we can enhance the view-specific group effect within individual view, which gives rise to the view-specific smooth regularizer, that is

$$\sum_{i,j=1}^{n} \|\mathrm{Z}_{:i}{}^{(v)} - \mathrm{Z}_{:j}{}^{(v)}\|_2^2 \mathrm{S}_{ij}{}^{(v)} = \mathrm{Tr}(\mathrm{Z}^{(v)}\mathrm{L}^{(v)}\mathrm{Z}^{(v)T}), \tag{15}$$

where $\mathrm{S}^{(v)}$ denotes the v-th similarity matrix, which can be derived from a k-nearest neighbor ($k$nn) graph. Besides, $\mathrm{L}^{(v)} = \mathrm{D}^{(v)} - \mathrm{S}^{(v)}$ is the Laplacian matrix of $\mathrm{S}^{(v)}$. In order to facilitate the consensus across different views, we can extend the smooth regularizer from view-specific to view-consensus by constructing a global affinity matrix:

$$\sum_{v=1}^{V} \mathrm{Tr}(\mathrm{Z}^{(v)}\mathrm{L}_S\mathrm{Z}^{(v)T}) \tag{16}$$

where L_S is the global affinity matrix of multiple views, which can be constructed according to Algorithm 1. Notice that the above view-consensus smooth regularizer is consistent with the third term (i.e., (14)) of our objective function. Therefore, by incorporating this graph term with the objective model in ((12) can benefit to explore the locality awareness across multiple views (via view-consensus group effect).

Moving forward, we analyze the benefit of our proposed one-step scheme from the statistical perspective. In the first palace, we consider the following optimization problem:

$$\min_{\mathrm{F}^T\mathrm{F}=\mathrm{I}} ||\mathrm{S} - \mathrm{F}\mathrm{F}^T|||_F^2 \tag{17}$$

where S is a given normalized similarity matrix. Moreover, a useful lemma w.r.t. this optimization problem is given as follows:

Lemma 1: The optimization problem in (17) is equivalent to the spectral clustering problem.

Proof:

$$\begin{aligned} &\min_{\mathrm{F}^T\mathrm{F}=\mathrm{I}} ||\mathrm{S} - \mathrm{F}\mathrm{F}^T|||_F^2 \\ \Leftrightarrow\ &\min_{\mathrm{F}^T\mathrm{F}=\mathrm{I}} \{\mathrm{Tr}(\mathrm{S}^T\mathrm{S}) - 2\mathrm{Tr}(\mathrm{S}\mathrm{F}\mathrm{F}^T) + \mathrm{Tr}(\mathrm{F}\mathrm{F}^T\mathrm{F}\mathrm{F}^T)\} \\ \Leftrightarrow\ &\max_{\mathrm{F}^T\mathrm{F}=\mathrm{I}} \mathrm{Tr}(\mathrm{F}^T\mathrm{S}\mathrm{F}) \end{aligned} \tag{18}$$

Since the input similarity matrix is normalized (i.e., $\sum_{j=1}^{n} \mathrm{S}_{ij} = 1, \forall i$), then the correspond diagonal matrix D becomes an identity matrix. Hence, the optimization in (18) can be rewritten in the following form:

$$\begin{aligned} &\max_{\mathrm{F}^T\mathrm{F}=\mathrm{I}} \mathrm{Tr}(\mathrm{F}^T\mathrm{S}\mathrm{F}) \\ \Leftrightarrow\ &\max_{\mathrm{F}^T\mathrm{F}=\mathrm{I}} \mathrm{Tr}(\mathrm{F}^T(\mathrm{I} - \mathrm{L})\mathrm{F}) \\ \Leftrightarrow\ &\min_{\mathrm{F}^T\mathrm{F}=\mathrm{I}} \mathrm{Tr}(\mathrm{F}^T\mathrm{L}\mathrm{F}) \end{aligned} \tag{19}$$

Here variable $\mathrm{F} \in \mathbb{R}^{n\times k}$ represents the spectral embedding of similarity matrix S. Obviously, the optimization problem in (19) is a standard spectral clustering problem. ■

Based on the above lemma, we further investigate the statistical property of regularizer term in (13). Formally, an important lemma is provided in what follows.

Lemma 2: Minimizing the optimization problem in (13) is equivalent to maximize the statistical correlation between multi-view representations and spectral embedding of global similarity matrix (affinity matrix).

Proof: Consider a set of multi-view representations $\sum_{v=1}^{V}\{\mathrm{Z}^{(v)}\}$ and a global affinity matrix that can be represented as $\mathrm{S} = \mathrm{F}\mathrm{F}^T$, we construct the following linear kernels, i.e., $\mathrm{K}_{\mathrm{Z}}{}^{(v)} = \mathrm{Z}^{(v)T}\mathrm{Z}^{(v)}, \forall v$ and $\mathrm{K}_{\mathrm{S}} = \{\mathrm{F}^T\}^T\mathrm{F}^T = \mathrm{S}$. Accordingly, by utilizing Hilbert Schmidt Independence Criterion (HSIC) as the empirical estimation, then we have:

$$\begin{aligned} &\max \sum_{v=1}^{V} \mathrm{HSIC}(\mathrm{Z}^{(v)}, \mathrm{F}^T) \\ \Leftrightarrow\ &\max \sum_{v=1}^{V} \mathrm{Tr}(\mathrm{K}_{\mathrm{Z}}{}^{(v)}\mathrm{H}\mathrm{K}_{\mathrm{S}}\mathrm{H}) \\ \Leftrightarrow\ &\max_{\mathrm{S}\mathbf{1}=\mathbf{1},\mathrm{S}\succeq\mathbf{0}} \sum_{v=1}^{V} \mathrm{Tr}(\mathrm{Z}^{(v)T}\mathrm{Z}^{(v)}\mathrm{H}\mathrm{S}\mathrm{H}) \\ \Leftrightarrow\ &\max_{\mathrm{S}\mathbf{1}=\mathbf{1},\mathrm{S}\succeq\mathbf{0}} \sum_{v=1}^{V} \mathrm{Tr}(\{\mathrm{Z}^{(v)}\mathrm{H}\}\mathrm{S}\{\mathrm{Z}^{(v)}\mathrm{H}\}^T) \end{aligned} \tag{20}$$

where $\mathrm{H} = \mathrm{I} - n^{-1}\mathbf{1}\mathbf{1}^T$ is a centering matrix. If we assume that the multi-view representations $\{\mathrm{Z}^{(v)}\}_{v=1}^{V}$ are already centered,

then $\mathrm{Z}^{(v)} = \mathrm{Z}^{(v)}\mathrm{H}$ is satisfied. Besides, due to the simplex constraint on global similarity matrix, then $\mathrm{S} = \mathrm{I} - \mathrm{L_S}$ is met. Thereby, the optimization problem in (20) can be reduced to the following form:

$$\begin{aligned} &\max_{\mathrm{S}\mathbf{1}=\mathbf{1},\mathrm{S}\succeq\mathbf{0}} \sum_{v=1}^{V} \mathrm{Tr}(\{\mathrm{Z}^{(v)}\mathrm{H}\}\mathrm{S}\{\mathrm{Z}^{(v)}\mathrm{H}\}^T) \\ \Leftrightarrow\ &\max_{\mathrm{S}\mathbf{1}=\mathbf{1},\mathrm{S}\succeq\mathbf{0}} \sum_{v=1}^{V} \mathrm{Tr}(\mathrm{Z}^{(v)}(\mathrm{I}-\mathrm{L_S})\mathrm{Z}^{(v)^T}) \\ \Leftrightarrow\ &\min_{\mathrm{S}\mathbf{1}=\mathbf{1},\mathrm{S}\succeq\mathbf{0}} \sum_{v=1}^{V} \mathrm{Tr}(\mathrm{Z}^{(v)}\mathrm{L_S}\mathrm{Z}^{(v)^T}) \end{aligned} \tag{21}$$

■

According to the above analysis, by integrating the regularizer term (13) within the one-step scheme can seamlessly maximize the correlation between multi-view subspace representations and latent spectral embedding, which is beneficial for constructing an informative similarity matrix for incomplete multi-view clustering. As far as we know, how to theoretically analyze the benefit of one-step framework remains a challenging problem. In the future, we expect to provide deeper insights towards the research of the one-step scheme in incomplete multi-view clustering.

D. Optimization Algorithm

To efficiently minimize our objective function, an ADMM-based optimization algorithm is derived in this subsection. Firstly, by introducing a set of auxiliary variables $\{\mathrm{C}^{(v)}\}_{v=1}^{V}$, the augmented Lagrangian function should be considered in the following form:

$$\begin{aligned} &\mathcal{L}\left(\{\mathrm{Z}^{(v)}\}_{v=1}^{V}, \mathcal{C}, \mathrm{S}, \{\mathrm{K}^{(v)}\}_{v=1}^{V}\right) \\ &= \sum_{v=1}^{V} \mathrm{Tr}\{\mathrm{K}^{(v)}(\mathrm{I}-\mathrm{Z}^{(v)}-\mathrm{Z}^{(v)^T}+\mathrm{Z}^{(v)}\mathrm{Z}^{(v)^T})\} \\ &\quad + \lambda_1\|\mathcal{C}\|_{\circledast} + \lambda_2 \sum_{v=1}^{V} \mathrm{Tr}(\mathrm{Z}^{(v)}\mathrm{L_S}\mathrm{Z}^{(v)^T}) \\ &\quad + \sum_{v=1}^{V}\left(\left\langle \mathrm{Y}^{(v)}, \mathrm{Z}^{(v)}-\mathrm{C}^{(v)}\right\rangle + \frac{\mu}{2}\left\|\mathrm{Z}^{(v)}-\mathrm{C}^{(v)}\right\|_F^2\right) \\ &\text{s.t. } \mathrm{K}^{(v)}(p_v,p_v) = \mathrm{K}_p^{(v)}, \mathrm{K}^{(v)} \succeq \mathbf{0}, \forall v, \\ &\quad \mathcal{C} = \Phi(\mathrm{C}^{(1)}, \mathrm{C}^{(2)}, \ldots, \mathrm{C}^{(V)}), \\ &\quad \mathrm{S}\mathbf{1} = \mathbf{1}, \mathrm{S} \succeq \mathbf{0}, \|\mathrm{S}_{i:}\|_0 = k, \forall i \end{aligned} \tag{22}$$

in which matrices $\{\mathrm{Y}^{(v)}\}_{v=1}^{V}$ are Lagrange multipliers, while μ is a penalty parameter. Next, an iterative algorithm is adopted to solve the subproblems w.r.t. $\{\mathrm{Z}^{(v)}\}_{v=1}^{V}$, $\mathcal{C}$, S and $\{\mathrm{K}^{(v)}\}_{v=1}^{V}$ in proper order.

1) Solving the $\mathrm{Z}^{(v)}$-Subproblem: If the other variables are fixed, then the subproblem of (22) w.r.t. $\mathrm{Z}^{(v)}$ has the following form:

$$\begin{aligned} \min_{\mathrm{Z}^{(v)}}\ &\mathrm{Tr}\left\{\mathrm{Z}^{(v)^T}\left(\mathrm{K}^{(v)}+\frac{\mu}{2}\mathrm{I}\right)\mathrm{Z}^{(v)}\right\} + \lambda_2\mathrm{Tr}\left\{\mathrm{Z}^{(v)}\mathrm{L_S}\mathrm{Z}^{(v)^T}\right\} \\ &+ \mathrm{Tr}\{(-2\mathrm{K}^{(v)}+\mathrm{Y}^{(v)^T}-\mu\mathrm{C}^{(v)^T})\mathrm{Z}^{(v)}\} \end{aligned} \tag{23}$$

If we take the derivative of the subproblem w.r.t. $\mathrm{Z}^{(v)}$ and let it vanish, then the following equation is given:

$$\mathrm{Z}^{(v)}\lambda_2\mathrm{L_S} + \left(\mathrm{K}^{(v)}+\frac{\mu}{2}\mathrm{I}\right)\mathrm{Z}^{(v)} = \mathrm{K}^{(v)} + \frac{\mu}{2}\mathrm{C}^{(v)} - \frac{1}{2}\mathrm{Y}^{(v)}. \tag{24}$$

This is a standard Sylvester Equation [42], thus it can be solved by Bartels-Stewart algorithm [42].

2) Solving the $\mathcal{C}$-Subproblem: If the other variables are fixed, then the subproblem of (22) w.r.t. $\mathcal{C}$ has the following form:

$$\min_{\mathcal{C}}\ \lambda_1\|\mathcal{C}\|_{\circledast} + \tfrac{\mu}{2}\|\mathcal{C}-(\mathcal{Z}+\tfrac{\mathcal{Y}}{\mu})\|_F^2 \tag{25}$$

Following closely by the research in [43], this subproblem is solved by tensor tubal-shrinkage operator [44]:

$$\mathcal{C} = \mathcal{D}_{n_3\pi}(\mathcal{Q}) = \mathcal{U} * \mathcal{C}_{n_3\pi}(\mathcal{O}) * \mathcal{V}^T. \tag{26}$$

Here $\pi = \mu/\lambda_1$ and $\mathcal{Q} = \mathcal{U} * \mathcal{O} * \mathcal{V}^T$. In addition, $\mathcal{D}_{n_3}(\mathcal{Q}) = \mathcal{Q} * \mathcal{P}$, herein, tensor $\mathcal{P}$ is f-diagonal, and $\mathcal{P}_f(i,i,j) = (1 - \frac{n_3\pi}{\mathcal{O}_f^{(j)}(i,i)})_+$ means the diagonal element of $\mathcal{P}$ in the Fourier domain.

3) Solving the S*-Subproblem:* If the other variables are fixed, then the subproblem of (22) w.r.t. S has the following form:

$$\begin{aligned} &\min_{\mathrm{S}} \sum_{v=1}^{V}\sum_{i,j=1}^{n} \|\mathrm{Z}_{:i}^{(v)} - \mathrm{Z}_{:j}^{(v)}\|_2^2 \mathrm{S}_{ij} \\ &\text{s.t. } \mathrm{S}\mathbf{1} = \mathbf{1}, \mathrm{S} \succeq \mathbf{0}, \|\mathrm{S}_{i:}\|_0 = k, \forall i. \end{aligned} \tag{27}$$

Thus, this subproblem can be solved by:

$$\mathrm{S}_{ij} = \begin{cases} \frac{\mathrm{H}_{i,k+1}-\mathrm{H}_{ij}}{k\mathrm{H}_{i,k+1}-\sum_{j=1}^{k}\mathrm{H}_{ij}} & j \le k \\ 0 & j > k \end{cases} \tag{28}$$

Here $\mathrm{H}_{ij} = \sum_{v=1}^{V}\|\mathrm{Z}_{:i}^{(v)} - \mathrm{Z}_{:j}^{(v)}\|_2^2$. According to this equation, we can determine the sparsity of S by the number of neighbors k, where the similar update solution can be found in the relevant literatures [45], [46].

4) The Rule for Updating $\{\mathrm{K}^{(v)}\}_{v=1}^{V}$: If the other variables are fixed, then the subproblem of (22) w.r.t. $\mathrm{K}^{(v)}$ has the following form:

$$\begin{aligned} &\min_{\mathrm{K}^{(v)}} \mathrm{Tr}(\mathrm{K}^{(v)}\mathrm{H}^{(v)}) \\ &\text{s.t. } \mathrm{K}^{(v)}(p_v,p_v) = \mathrm{K}_p^{(v)}, \mathrm{K}^{(v)} \succeq \mathbf{0} \end{aligned} \tag{29}$$

where $\mathrm{H}^{(v)} = (\mathrm{I}-\mathrm{Z}^{(v)}-\mathrm{Z}^{(v)^T}+\mathrm{Z}^{(v)}\mathrm{Z}^{(v)^T})$. This is a standard kernel imputation problem. Borrowing the idea of self-expressive property in subspace clustering, we can parameterize the kernel matrix $\mathrm{K}^{(v)}$ into the following form:

$$\mathrm{K}^{(v)} = \begin{bmatrix} \mathrm{K}_p^{(v)} & \mathrm{K}_p{}^{(v)}\mathrm{B}^{(v)} \\ \mathrm{B}^{(v)^T}\mathrm{K}_p{}^{(v)} & \mathrm{B}^{(v)^T}\mathrm{K}_p{}^{(v)}\mathrm{B}^{(v)} \end{bmatrix} \tag{30}$$

where $\mathrm{B}^{(v)}$ is a self-expressive matrix of the v-th view, such that the missing kernel entries can be linearly represented by the

Algorithm 1: UT-IMKSC.

Input: Incomplete multi-view dataset with V views, the maximum iterations it_{max}, the number of near neighbors k and two balancing hyper-parameter $\lambda_1, \lambda_2 > 0$.

Parameter Setup: Set $\mu = 10^{-2}$, $\eta = 1.5$, $\mu_{max} = 10^6$ and $\varepsilon = 10^{-4}$.

Initialization: Initialize $\mathrm{C}^{(v)} = \mathrm{I}$, $\mathrm{Y}^{(v)} = \mathbf{0}, \forall v$. Calculate the partial kernels $\{\mathrm{K}_p{}^{(v)}\}_{v=1}^{V}$ for observed samples. Impute the $\{\mathrm{K}^{(v)}\}_{v=1}^{V}$ by zero filling method. Adopt the (28) to initialize the variable S by replacing $\mathrm{H}_{ij} = \sum_{v=1}^{V} \|\mathrm{X}_{:i}{}^{(v)} - \mathrm{X}_{:j}{}^{(v)}\|_2^2$. Set $it = 1$.

repeat

Obtain $\{\mathrm{Z}^{(v)}\}_{v=1}^{V}$ by solving the problem in (24).

Obtain $\mathcal{C}$ by (26).

Obtain S by (28).

Obtain $\{\mathrm{K}^{(v)}\}_{v=1}^{V}$ by (33) and (30).

Obtain $\{\mathrm{Y}^{(v)}\}_{v=1}^{V}$ by $\mathrm{Y}^{(v)} = \mathrm{Y}^{(v)} + \mu(\mathrm{Z}^{(v)} - \mathrm{C}^{(v)})$.

Update μ by using $\mu = \min(\eta\mu, \mu_{\max})$.

$it = it + 1$.

until The convergence conditions in (34) to (36) are met or $it > it_{max}$.

Output: the common affinity matrix S for postprocessing Ncut algorithm.

observed entries. Meanwhile, by integrating (30), we can rewrite the above problem (see (29)) in the following form:

$$\min_{\mathrm{B}^{(v)}} \mathrm{Tr}\left(\begin{bmatrix} \mathrm{K}_p^{(v)} & \mathrm{K}_p{}^{(v)}\mathrm{B}^{(v)} \\ \mathrm{B}^{(v)T}\mathrm{K}_p{}^{(v)} & \mathrm{B}^{(v)T}\mathrm{K}_p{}^{(v)}\mathrm{B}^{(v)} \end{bmatrix} \begin{bmatrix} \mathrm{H}_{\mathrm{cc}}^{(v)} & \mathrm{H}_{\mathrm{cm}}{}^{(v)} \\ \mathrm{H}_{\mathrm{cm}}{}^{(v)T} & \mathrm{H}_{\mathrm{mm}}{}^{(v)} \end{bmatrix}\right) \tag{31}$$

where $\mathrm{H}^{(v)}$ is blocked as following forms:

$$\mathrm{H}^{(v)} = \begin{bmatrix} \mathrm{H}_{\mathrm{cc}}^{(v)} & \mathrm{H}_{\mathrm{cm}}{}^{(v)} \\ \mathrm{H}_{\mathrm{cm}}{}^{(v)T} & \mathrm{H}_{\mathrm{mm}}{}^{(v)} \end{bmatrix} \tag{32}$$

If we take the derivative of this subproblem w.r.t. $\mathrm{B}^{(v)}$ and set it to be zero, the corresponding optimal solution is given as follows:

$$\mathrm{B}^{(v)} = -\mathrm{H}_{\mathrm{mm}}{}^{(v)}(\mathrm{H}_{\mathrm{cm}}{}^{(v)})^{-1}. \tag{33}$$

After combining the $\mathrm{B}^{(v)}$ in (33) with (30), we can achieve the optimal solution for variable $\mathrm{K}^{(v)}$.

In brief, the detailed process of optimization algorithm is listed in Algorithm 1.

E. Computational Complexity Analysis

The time complexity of Algorithm 1 mainly comes from four phases during iterations. In the first phase, the calculation of $\{\mathrm{Z}^{(v)}\}_{v=1}^{V}$ requires $\mathrm{O}(Vn^3)$ time cost, with V, n standing for the number of views and data samples, respectively. In the second phase, the calculation of $\mathcal{C}$ needs $\mathrm{O}(Vn^2 longn)$ time cost. In the third stage, the calculation of S requires $\mathrm{O}(nk)$, with k standing for the number of neighbors. In the fourth phase, the calculation of $\{\mathrm{K}^{(v)}\}_{v=1}^{V}$ needs $\mathrm{O}(\min\{(n_m{}^{(v)})^3, n_m{}^{(v)}(n_c{}^{(v)})^2\})$ time cost, with $n_m{}^{(v)}$ and $n_c{}^{(v)}$ denote the number of observed and missing objects of the v-th view. Suppose Algorithm 1 runs T times to convergence, since the condition $k, T, V \ll n$ is satisfied, the time complexity of our algorithm lies in $\mathrm{O}(VTn^3 + T\sum_{v=1}^{V}\min\{(n_m{}^{(v)})^3, n_m{}^{(v)}(n_c{}^{(v)})^2\})$. Moreover, according to the condition $n_m{}^{(v)} + n_c{}^{(v)} = n, \forall v$, the overall time complexity of the proposed algorithm reduced to $\mathrm{O}(VTn^3)$.

IV. Experiments

This section evaluates the proposed UT-IMKSC approach over five different aspects, e.g., comparison against other IMC approaches, parameter sensitivity, convergence analysis, computational time analysis, as well as ablation study. It is noteworthy that, all the experiments are built on a PC with i5-6600 CPU and 64 GB of RAM.

A. Experimental Settings

To begin with, we employ six practical multi-view datasets that have been commonly adopted for experimental purpose. Specifically, these datasets are collected with different characteristics, involving handwritten digit dataset (i.e., UCI-3view [47]), news report datasets (i.e, Reuters [21] and BBCSport [48]), object image dataset (i.e., COIL20 [49]) and face image datasets (i.e., Yale [50] and ORL). For clarity, a brief description of them is listed in Table I.

On this basis, we follow the relevant papers to generate the incomplete multi-view datasets. For each incomplete multi-view dataset, a parameter β is utilized to indicate the *missing rate*, which controls the percentage that data samples are removed in the partial views. Notice that, the clustering performance could be affected by the setting of parameter β. As a result, we adjust parameter β within the same range in our experiments (i.e., $\beta \in [0.1, \ldots, 0.5]$ with step 0.1).

To empirically evaluate the clustering quality of different approaches in comparison, two widely-used metrics are employed for evaluation, namely, the Normalized Mutual Information (NMI) [51] and the Accuracy (ACC) [52].

B. Comparison Against Other IMC Approaches

To verify the effectiveness and robustness of our UT-IMKSC approach, this subsection empirically compares it against various IMC baselines, they are:

- *MKKM-ZF [53]:* Multi-view kernel k-mean clustering with zero filling algorithm.
- *MKKM-MF [53]:* Multi-view kernel k-mean clustering with mean filling algorithm.
- *MKKM-KNN [53]:* Multi-view kernel k-mean clustering with k-nearest neighbor algorithm.
- *BSV [54]:* Impute individual view with mean filling algorithm, and then select the best single-view result by performing spectral clustering.
- *Concat [54]:* Concatenate all views and impute the missing values with mean filling algorithm, and then perform spectral clustering to achieve the final result.

TABLE I
DETAILS OF SIX BENCHMARKS USED IN THE EXPERIMENTS

Dataset	#Sample	#View	#Class	Dimension
UCI-3view	2000	3	10	PIX (240), FOU (76), MOR (6)
COIL20	1440	3	20	Intensity (1024), LBP (3304), Gabor (6750)
Reuters	1200	5	6	English (2,000), German (2,000), French (2,000), Spanish (2,000), Italian(2,000)
BBCSports	544	2	5	View1 (3183), View2 (3203)
ORL	400	3	40	LBP (3304), Gabor(6750), Intensity (4096)
Yale	165	3	15	LBP (3304), Gabor(6750), Intensity (4096)

TABLE II
KERNEL CONSTRUCTIONS AND PARAMETER SETTINGS OF UT-IMKSCL ON THE TEN BENCHMARK DATA SETS

	UCI-3view	COIL20	Reuters	BBCSports	ORL	Yale
View 1	Gaussian	Gaussian	Gaussian	Linear	Linear	Linear
View 2	Linear	Gaussian	Gaussian	Linear	Linear	Linear
View 3	Linear	Gaussian	Gaussian	-	Linear	Gaussian
View 4	-	-	Gaussian	-	-	-
View 5	-	-	Linear	-	-	-
λ_1	100	100	100	100	10	10
λ_2	1	1	0.01	1	0.1	0.01
k	30	10	40	15	5	25

- *UEAF [55]:* Incomplete multi-view spectral clustering based on unified embedding alignment framework.
- *IMC-AGL [5]:* Incomplete multi-view spectral clustering via adaptive graph learning.
- *UIMC [56]:* Unbalanced incomplete multi-view clustering.
- *IK-MKKM [9]:* Incomplete multiple kernel k-means clustering.
- *EE-IMVC [36], [57]:* Effective and efficient incomplete multi-view clustering.

Starting from the different perspectives, these ten IMC baselines can be generally divided into two classes, namely two-stage based IMC approaches and one-stage based IMC approaches. One the one hand, the two-stage based IMC approaches involve MKKM-ZF, MKKM-MF, MKKM-KNN, BSV and Concat. On the other hand, the one-stage based IMC approaches include UEAF, IMC-AGL, UIMC, IK-MKKM and EE-IMVC. Their detailed descriptions can be found in the original papers.

It should be emphasized that, the following experimental setting should be obeyed:

- The source code of IMC baselines is directly downloaded in public access. In addition, the corresponding hyper-parameters will be varied within a wide range, i.e., $[10^{-5}, 10^{-4}, \ldots, 10^{4}, 10^{5}]$, except some specific range (or value) is suggested in the original paper.
- Regarding to our UT-IMKSC approach, two positive hyper-parameters λ_1 and λ_2 are varied within the same range of $[10^{-4}, 10^{-3}, \ldots, 10^{3}, 10^{4}]$, while the number of near neighbor is varied within the range of $[5, \ldots, 40]$ with step 5.
- For the proposed method, we use two types of kernels (i.e., the linear kernel as well as the Gaussian kernel) in different views. The corresponding kernel constructions and the default parameters of our method are listed in Table II.
- For each tested dataset, we repeat every IMC approach for 20 times, and meanwhile report the best clustering results by cross-validation strategy.

Specifically, the empirical performance (w.r.t. NMI and ACC) are respectively displayed in Tables III and IV. Notice that the aggregated results and average rank are also presented for comprehensiveness. By analyzing the experimental results, we can conclude some valuable observations as below.

- Among the eleven IMC competitors, it turns out that the one-stage based approaches are generally superior to that of the two-stage based approaches. In terms of the aggregated NMI, one-stage approach IK-MKKM achieves the improvements of 12.7%, 10.8%, 9.5% and 6.6% on four datasets (e.g., COIL20, BBCSports, ORL and Yale), by comparing with the best two-stage approaches (i.e., MKKM-KNN and MKKM-MF). This comparison indicates that the unified optimization on imputation and clustering is more flexibly suited to IMC problems.
- Regarding to the two-stage IMC cases, we can observe that the kernel k-means based approaches (i.e., MKKM-ZF, MKKM-MF and MKKM-KNN) perform more reliable over that of the traditional k-means approaches (i.e., BSV and Concat). This comparison validates that the incorporation of kernel learning gives rise to the robust IMC performance, by exploiting the non-linear structure embedding in distinct views.
- As for the one-stage IMC cases, we can see that the performance of graph-based approaches (i.e., UEAF and IMSC-AGL) is consistent (or even better) than the newly proposed approach EE-IMVC. This comparison demonstrates that the graph-based approach is capable of capturing the latent relationship (or structural topology) among available and

TABLE III
NMI(%) SCORES OVER 20 RUNS BY ELEVEN IMC APPROACHES

Dataset	Missing rate	MKKM-ZF	MKKM-MF	MKKM-KNN	BSV	Concat	UEAF	IMC-AGL	UIMC	IK-MKKM	EE-IMVC	UT-IMKSC
UCI-3views	0.1	42.3	43.1	48.5	40.6	65.5	63.9	84.9	[85.3]	45.9	65.6	**92.8**
	0.2	38.6	38.3	45.1	35.2	57.7	65.0	80.7	[81.4]	42.4	59.4	**91.8**
	0.3	34.0	33.8	40.3	30.0	43.0	57.0	74.4	[74.5]	38.2	53.2	**92.1**
	0.4	29.3	29.1	34.1	25.7	37.5	46.6	66.3	[67.8]	35.3	45.8	**89.6**
	0.5	24.3	24.0	28.0	21.7	30.5	35.7	57.0	[58.1]	30.9	38.0	**91.5**
	Agg NMI	33.7	33.7	39.2	30.6	46.8	53.6	72.7	[73.4]	38.5	52.4	**91.5**
	Avg Rank	9.2	9.8	7.3	11.0	6.0	4.5	3.0	[2.0]	7.5	4.7	**1.0**
COIL20	0.1	70.6	79.2	77.5	63.1	68.3	77.8	[89.3]	70.1	81.1	75.3	**99.8**
	0.2	64.1	69.8	72.1	56.4	49.0	77.6	[87.5]	68.6	79.7	69.5	**97.4**
	0.3	55.5	62.0	64.6	51.0	45.2	76.2	[84.6]	66.4	77.8	61.9	**97.1**
	0.4	49.5	54.6	55.6	41.6	37.9	73.1	[79.2]	63.6	73.6	53.2	**96.5**
	0.5	42.0	45.9	48.0	36.5	28.8	69.3	[71.3]	59.6	69.3	42.8	**97.0**
	Agg NMI	56.3	62.3	63.6	49.7	45.8	74.8	[82.4]	65.7	76.3	60.5	**97.6**
	Avg Rank	8.8	6.3	5.8	10.2	10.8	4.0	[2.0]	6.2	3.0	7.7	**1.0**
Reuters	0.1	2.9	29.0	3.4	1.7	0.4	[29.7]	14.3	18.8	29.2	26.7	**66.7**
	0.2	1.4	28.6	2.0	0.8	0.4	[32.8]	14.4	17.8	29.1	26.2	**66.2**
	0.3	1.0	28.0	2.1	16.2	0.4	[30.4]	14.1	16.7	28.9	26.3	**70.4**
	0.4	2.7	26.3	3.1	0.4	0.6	[31.0]	14.5	17.1	28.3	24.8	**74.1**
	0.5	2.4	20.6	3.6	2.7	0.4	[28.1]	12.6	15.7	27.1	22.0	**77.4**
	Agg NMI	2.1	26.5	2.9	4.3	0.5	[30.4]	14.0	17.2	28.5	25.2	**70.9**
	Avg Rank	9.5	4.2	8.3	9.2	10.8	[2.0]	7.2	6.0	3.0	4.8	**1.0**
BBCSports	0.1	44.9	55.7	53.4	45.0	0.8	75.9	82.5	[84.7]	60.8	74.5	**95.5**
	0.2	31.7	51.1	35.2	0.7	3.5	75.6	74.0	[77.7]	58.4	68.2	**97.9**
	0.3	21.8	48.3	27.1	0.9	1.2	[69.7]	68.7	68.6	56.6	57.3	**94.4**
	0.4	11.9	34.6	24.2	0.7	1.5	[63.4]	57.3	57.7	51.8	47.3	**94.7**
	0.5	6.6	29.7	14.7	0.7	0.8	[53.6]	46.2	45.9	45.9	37.0	**93.6**
	Agg NMI	23.4	43.9	30.9	9.6	1.5	[67.7]	65.8	66.9	54.7	56.9	**95.2**
	Avg Rank	9.2	7.0	8.0	10.5	10.3	[2.5]	3.5	3.0	5.7	5.3	**1.0**
ORL	0.1	78.8	83.8	81.6	63.1	54.8	73.8	[88.6]	84.4	85.7	86.8	**100.0**
	0.2	74.7	78.0	75.4	42.7	30.6	72.0	[87.8]	83.3	84.9	85.8	**99.8**
	0.3	69.4	71.4	68.6	42.9	25.8	71.3	[85.3]	81.1	82.0	83.6	**99.3**
	0.4	63.0	64.7	63.5	37.2	23.2	65.3	[80.6]	76.4	78.4	79.7	**99.6**
	0.5	57.2	57.7	57.6	34.4	29.0	60.4	73.1	69.7	71.8	[73.2]	**99.3**
	Agg NMI	68.6	71.1	69.3	44.1	32.7	68.5	[83.1]	79.0	80.6	81.8	**99.6**
	Avg Rank	8.3	6.3	7.7	10.0	11.0	7.7	[2.2]	5.0	4.0	2.8	**1.0**
Yale	0.1	53.5	64.6	58.4	38.9	40.8	57.8	67.2	[68.3]	65.8	64.1	**89.8**
	0.2	48.8	60.9	56.1	51.6	34.0	57.3	65.9	[66.0]	65.8	63.4	**86.2**
	0.3	46.7	55.7	50.7	44.5	23.5	56.9	[64.4]	63.0	62.3	60.9	**79.3**
	0.4	42.2	48.4	44.0	22.1	28.6	59.6	59.7	[60.6]	57.1	56.7	**79.3**
	0.5	37.2	40.9	39.6	27.4	18.8	51.2	55.0	53.8	[55.7]	53.2	**66.7**
	Agg NMI	45.7	54.1	49.7	36.9	29.1	56.6	[62.4]	62.3	61.3	59.7	**80.2**
	Avg Rank	9.2	6.5	7.8	10.2	10.7	6.0	[2.7]	[2.7]	3.7	5.2	**1.0**

* Note that "Agg NMI" indicates the aggregated NMI results, "Avg Rank" indicates the average rank.
On each dataset, the best score is highlighted in bold, while the second best one in [brackets].

missing samples, which demonstrates the effectiveness of graph learning for robust incomplete multi-view clustering.

- When the missing rate β is varied under different values, it can be seen that our proposed approach has gained impeccably promising performance on all the six datasets. Especially, our approach achieves very competitive results on the *BBCSports* and *ORL* datasets. Beyond the expectation, our approach has exhibited apparent advantages over the second-best competitor on three challenging datasets, namely *Reuters*, *ORL* and *Yale* (with large margins over 15%). Take the *Reuters* dataset as an example, it obtains the scores of 70.9% and 82.1% by aggregated NMI and aggregated Accuracy, where the second best IMC approaches only achieve 30.4% and 50.8%, respectively. Notice that, our proposed approach is always ranked in the first position on all six datasets over different metrics, as shown in the last line of Tables III and IV.

Following the quantitative analysis, we further conduct the visual experiments by comparing our UT-IMKSC approach with UEAF approach, with the missing rate is adjusted in the range of [01,0.3,0.5]. Particularly, the common affinity matrices w.r.t. two IMC approaches are plotted, whereas two datasets (i.e., BBCSports and ORL) are selected in the experiment. The visualization results in Figs. 4 and 5 clearly show that our proposed approach separates different clusters clean under different settings of missing rate. Besides, we can see that UT-IMKSC approach can better recover the underlying cluster structure in comparison with its competitor. As a result, these visualization results are well consistent with the quantitative results in Tables III and IV. Overall, the comprehensive results

TABLE IV
ACCURACY(%) SCORES OVER 20 RUNS BY ELEVEN IMC APPROACHES

Dataset	Missing rate	MKKM-ZF	MKKM-MF	MKKM-KNN	BSV	Concat	UEAF	IMC-AGL	UIMC	IK-MKKM	EE-IMVC	UT-IMKSC
UCI-3views	0.1	44.0	45.2	53.9	38.6	71.1	70.9	86.5	[89.1]	46.7	76.7	**97.0**
	0.2	40.8	41.6	49.3	35.1	64.7	78.1	84.9	[87.6]	45.4	73.0	**96.3**
	0.3	36.9	37.8	43.7	30.8	48.7	69.9	81.7	[82.8]	41.7	67.5	**96.6**
	0.4	33.0	34.0	36.2	27.7	45.5	56.5	76.0	[79.3]	39.1	59.8	**94.1**
	0.5	29.2	28.9	31.3	26.1	40.9	46.0	68.9	[69.8]	34.6	51.9	**96.0**
	Agg Acc	36.8	37.5	42.9	31.7	54.2	64.3	79.6	[81.7]	41.5	65.8	**96.0**
	Avg Rank	9.8	9.2	7.3	11.0	5.8	4.8	3.0	[2.0]	7.7	4.3	**1.0**
COIL20	0.1	63.6	70.8	69.4	49.5	55.5	66.8	[77.1]	54.2	72.4	66.2	**99.9**
	0.2	55.8	61.4	64.3	47.1	36.9	68.9	[76.2]	51.7	71.7	62.4	**92.2**
	0.3	46.2	52.7	57.0	42.7	37.8	68.0	[76.3]	50.6	69.6	57.6	**92.6**
	0.4	40.3	45.8	46.3	36.5	32.6	66.3	[72.9]	50.4	65.1	51.3	**92.9**
	0.5	34.1	37.4	37.7	31.6	26.1	63.5	[66.5]	47.6	63.1	43.2	**94.0**
	Agg Acc	48.0	53.6	54.9	41.5	37.8	66.7	[73.8]	50.9	68.4	56.2	**94.3**
	Avg Rank	8.7	6.8	6.0	10.2	10.7	3.8	2.0	7.7	3.2	5.7	**1.0**
Reuters	0.1	23.5	51.2	23.4	18.3	16.9	[52.7]	35.8	38.9	51.0	46.6	**77.3**
	0.2	21.7	51.3	22.2	17.4	16.9	[52.0]	36.1	37.4	51.3	46.4	**76.8**
	0.3	20.5	50.9	21.5	34.9	16.9	48.5	35.4	36.8	[51.3]	47.3	**81.8**
	0.4	21.9	49.4	22.7	17.1	17.2	[51.3]	37.2	38.6	50.6	45.8	**85.6**
	0.5	22.2	42.2	23.3	19.2	16.9	48.0	35.7	38.3	[49.8]	43.8	**89.1**
	Agg Acc	22.0	49.0	22.6	21.4	17.0	50.5	36.0	38.0	[50.8]	46.0	**82.1**
	Avg Rank	9.0	3.7	8.3	9.8	10.8	[2.7]	7.0	6.0	2.8	4.8	**1.0**
BBCSports	0.1	64.0	71.9	71.7	68.4	35.5	87.4	93.8	[95.0]	76.6	89.3	**98.7**
	0.2	52.8	65.7	57.0	35.7	34.0	[91.2]	87.8	91.1	73.8	86.7	**99.5**
	0.3	44.6	62.5	51.5	35.8	35.8	[88.1]	87.2	85.3	71.8	78.1	**98**.4
	0.4	36.2	52.8	50.5	35.7	35.8	77.6	78.0	[78.7]	67.6	70.0	**98.4**
	0.5	31.5	46.5	39.8	35.7	35.5	[69.0]	68.1	66.0	63.1	60.4	**97.8**
	Agg Acc	45.8	59.9	54.1	42.2	35.3	82.6	83.0	[83.2]	70.6	76.9	**98.5**
	Avg Rank	9.5	7.0	8.0	9.8	10.7	3.2	3.2	[2.8]	5.8	5.0	**1.0**
ORL	0.1	63.9	69.5	67.9	51.0	38.9	58.1	[76.1]	68.7	72.1	75.8	**100.0**
	0.2	59.6	63.2	60.2	31.7	26.5	56.0	[75.8]	68.5	72.1	74.4	**99.8**
	0.3	52.7	55.7	50.9	34.5	20.7	56.8	[73.6]	67.2	67.4	71.8	**98.3**
	0.4	44.3	47.7	44.6	29.5	16.5	51.4	[68.3]	62.5	64.9	66.8	**99.5**
	0.5	38.3	40.3	38.5	29.5	21.5	46.0	[59.9]	56.0	58.0	[59.9]	**96.3**
	Agg Acc	51.8	55.3	52.4	35.2	24.8	53.7	[70.7]	64.6	66.9	69.7	**98.4**
	Avg Rank	8.5	6.3	7.8	10.0	11.0	7.2	[2.0]	5.2	4.0	3.0	**1.0**
Yale	0.1	49.3	62.2	54.4	31.6	31.9	53.5	65.9	[67.3]	62.5	62.0	**87.3**
	0.2	44.6	57.5	51.4	48.5	30.3	51.5	64.6	[65.1]	62.9	61.4	**85.1**
	0.3	42.7	51.4	45.9	41.8	26.7	51.8	[63.1]	61.8	59.2	58.5	**77.1**
	0.4	37.5	42.9	38.7	21.2	27.3	51.5	58.7	[59.1]	53.6	53.9	**75.2**
	0.5	32.8	37.5	34.5	26.1	19.4	46.1	[53.2]	52.0	51.8	49.9	**62.6**
	Agg Acc	41.4	50.3	45.0	33.8	27.1	50.9	[61.1]	[61.1]	58.0	57.2	**77.5**
	Avg Rank	9.2	6.5	7.8	10.2	10.7	6.5	2.5	[2.3]	4.0	4.8	**1.0**

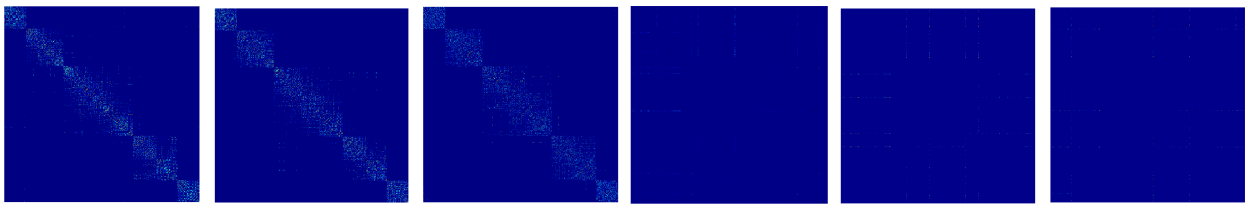
* Note that "Agg Acc" indicates the aggregated Accuracy results, "Avg Rank" indicates the average rank.
On each dataset, the best score is highlighted in bold, while the second best one in [brackets].

have confirmed the robustness and effectiveness of our LTKMSC approach when compared with the other IMC approaches.

C. Parameter Sensitivity

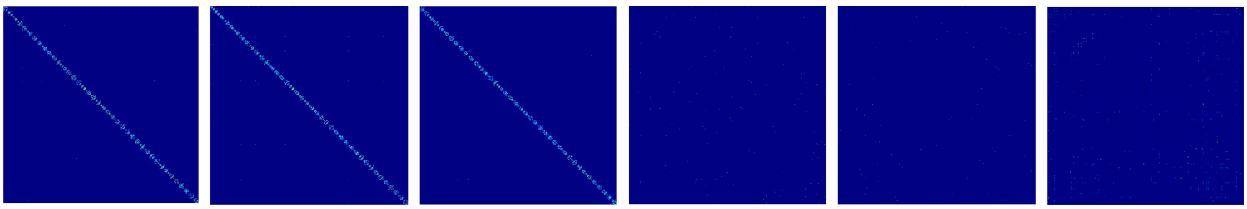
This subsection experimentally studies the influence of λ_1, λ_2 and k in our UT-IMKSC approach, with the missing rate is set within [0.1,0.3,0.5]. In the first palace, we investigate the parameter k while the hyper-parameters λ_1 and λ_2 are fixed as constraint value ($\lambda_1 = 100$ and $\lambda_2 = 1$). The clustering performance (over NMI and ACC metrics) w.r.t. different missing rates are depicted in Figs. 6 to 8, respectively. As observed in these figures, the NMI and ACC matrics remain stable over a wide range of parameter knn. Particularly, as knn goes from 5 to 40, the proposed approach shows very effective clustering performance on *UCI-3views*, *COIL20* and *BBCSports* datasets. This demonstrates that our approach is relatively robust to the variation of parameter knn for many IMC tasks.

Moving forward, we further conduct experiments to test the hyper-parameters λ_1 and λ_2 in our UT-IMKSC approach. As previously, we study their influence while fixing the left parameter knn as a static value, i.e., 15. In the meanwhile, the parameters λ_1 and λ_2 are varied in the same range of $[10^{-4}, 10^{-3}, \ldots, 10^{3}, 10^{4}]$. Besides, we only report the results over NMI in the following experiments since the Accuracy results have a similar trend. The visualization results w.r.t. different missing rate are shown in Figs. 9 to 11, respectively. Based on the figures, we make two crucial observations. In the first palace, our approach behaves relatively stable on the ORL and Yale datasets, with different combinations of parameters λ_1 and λ_2. Furthermore, when relatively large parameter λ_1 (i.e.,



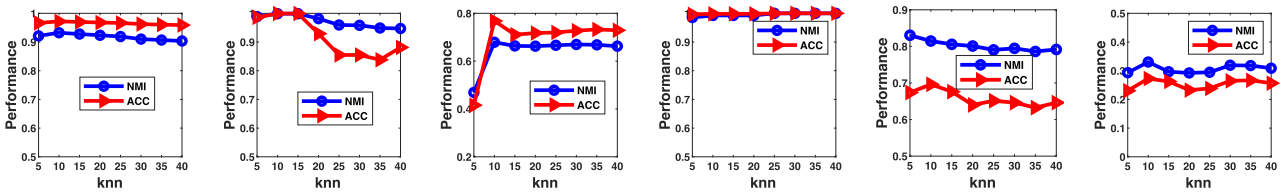
(a) UT-IMKSC, β = 0.1 (b) UT-IMKSC, β = 0.3 (c) UT-IMKSC, β = 0.5 (d) UEAF, β = 0.1 (e) UEAF, β = 0.3 (f) UEAF, β = 0.5

Fig. 4. Visualization of affinity matrices w.r.t. two comparison approaches on BBCSports dataset.



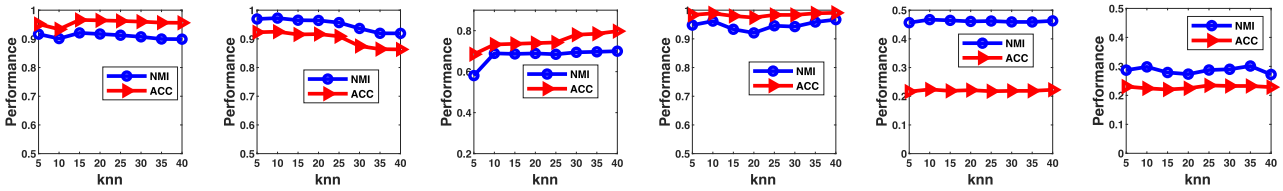
(a) UT-IMKSC, β = 0.1 (b) UT-IMKSC, β = 0.3 (c) UT-IMKSC, β = 0.5 (d) UEAF, β = 0.1 (e) UEAF, β = 0.3 (f) UEAF, β = 0.5

Fig. 5. Visualization of affinity matrices w.r.t. two comparison approaches on ORL dataset.



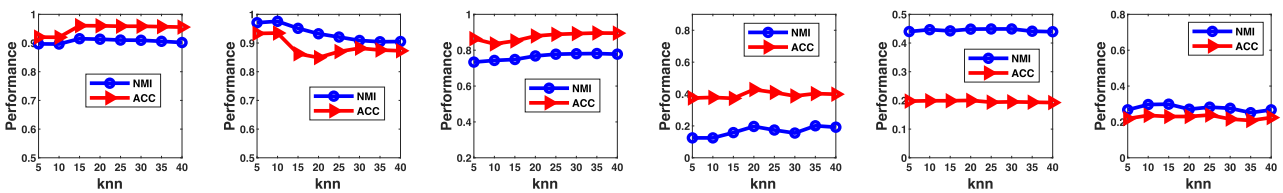
(a) *UCI-3views* (b) *COIL20* (c) *Reuters* (d) *BBCSports* (e) *ORL* (f) *Yale*

Fig. 6. Parameter study on the neighbor number knn while fixing the hyper-parameters λ_1 and λ_2 as static values (missing rate = 0.1).



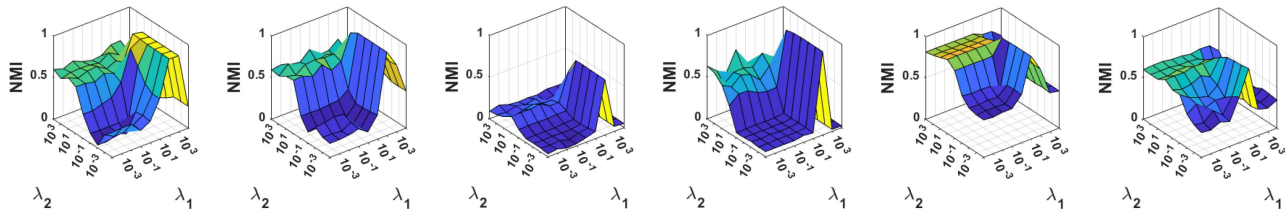
(a) *UCI-3views* (b) *COIL20* (c) *Reuters* (d) *BBCSports* (e) *ORL* (f) *Yale*

Fig. 7. Parameter study on the neighbor number knn while fixing the hyper-parameters λ_1 and λ_2 as static values (missing rate = 0.3).



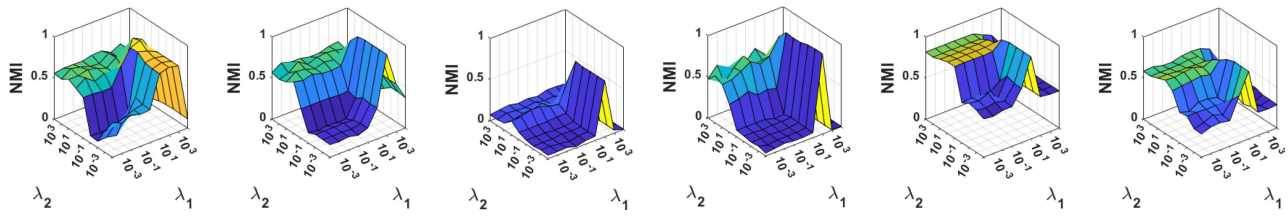
(a) *UCI-3views* (b) *COIL20* (c) *Reuters* (d) *BBCSports* (e) *ORL* (f) *Yale*

Fig. 8. Parameter study on the neighbor number knn while fixing the hyper-parameters λ_1 and λ_2 as static values (missing rate = 0.5).



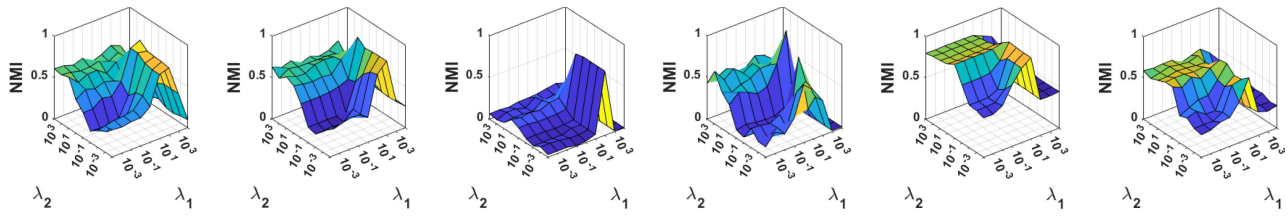
(a) *UCI-3views* (b) *COIL20* (c) *Reuters* (d) *BBCSports* (e) *ORL* (f) *Yale*

Fig. 9. Parameter study (w.r.t. NMI) on the hyper-parameters λ_1 and λ_2 while fixing the neighbor number knn as static values (missing rate $= 0.1$).



(a) *UCI-3views* (b) *COIL20* (c) *Reuters* (d) *BBCSports* (e) *ORL* (f) *Yale*

Fig. 10. Parameter study (w.r.t. NMI) on the hyper-parameters λ_1 and λ_2 while fixing the neighbor number knn as static values (missing rate $= 0.3$).



(a) *UCI-3views* (b) *COIL20* (c) *Reuters* (d) *BBCSports* (e) *ORL* (f) *Yale*

Fig. 11. Parameter study (w.r.t. NMI) on the hyper-parameters λ_1 and λ_2 while fixing the neighbor number knn as static values (missing rate $= 0.5$).

$\lambda_1 \geq 10^{-1}$) and relatively small parameter λ_2 (i.e., $\lambda_2 \leq 10^{-1}$) are selected, our approach always generates favorable performance for most of the datasets. Therefore, it is evident that an appropriate combination of hyper-parameters (λ_1 and λ_2) in our UT-IMKSC approach can benefit to practical IMC applications.

D. Convergence And Computational Time Analysis

This subsection first analyzes the convergence property of our UT-IMKSC approach on six incomplete multi-view datasets. Notably, the convergence conditions in Algorithm 1 are determined by three error terms, namely the match error (ME), the iteration error for variable $\mathcal{Z}$ (IE$_{\mathcal{Z}}$) and the iteration error for variable $\mathcal{C}$ (IE$_{\mathcal{C}}$). Concretely, these terms can be defined as follows:

$$\text{ME} = \sum_{v=1}^{V} ||\text{Z}^{(v)} - \text{C}^{(v)}||_{\infty} \tag{34}$$

and

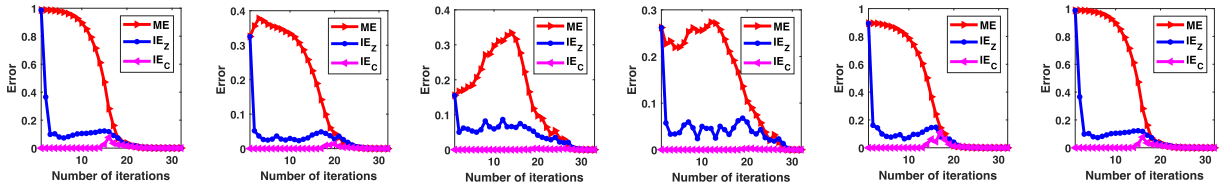
$$\text{IE}_{\mathcal{Z}} = \|\mathcal{Z}_{current} - \mathcal{Z}_{pre}\|_{\infty} \tag{35}$$

and

$$\text{IE}_{\mathcal{C}} = \|\mathcal{C}_{current} - \mathcal{C}_{pre}\|_{\infty} \tag{36}$$

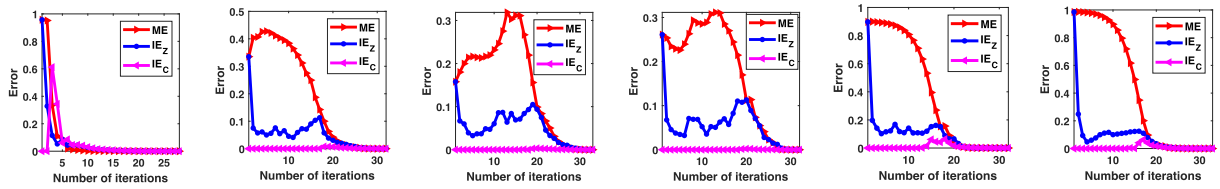
According to the curves displayed in Figs. 12, 13, 14, we can see that our approach can always exhibit a relatively fast convergence speed, with the missing rate varying from 0.1 to 0.5. In general, these three convergence curves will concurrently tend to zero around 30 iterations.

Next, we further analyze the running time of comparison approaches on benchmark data sets. The corresponding experimental results are recorded in Table V. As shown in the table, one-stage based approaches need to take more time compared with two-stage based approaches. However, it should be noted that the advantages of one-stage based approaches in clustering performance make up for the increase in time complexity. For the one-stage cases, it is clear that the IK-MKKM and EE-IMVC approaches require less time compared to the other approaches due to their simplified structure. Further, it is clear that our approach is faster than matrix factorization based approaches (i.e., UEAF, IMC-AGL and UIMC) in most of cases. This is probably because the proposed approach needs fewer times



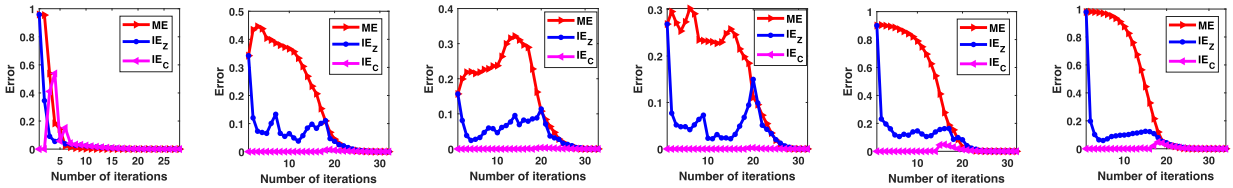
(a) *UCI-3views* (b) *COIL20* (c) *Reuters* (d) *BBCSports* (e) *ORL* (f) *Yale*

Fig. 12. Convergence analysis on the benchmark datasets when setting the missing rate as 0.1.



(a) *UCI-3views* (b) *COIL20* (c) *Reuters* (d) *BBCSports* (e) *ORL* (f) *Yale*

Fig. 13. Convergence analysis on the benchmark datasets when setting the missing rate as 0.3.



(a) *UCI-3views* (b) *COIL20* (c) *Reuters* (d) *BBCSports* (e) *ORL* (f) *Yale*

Fig. 14. Convergence analysis on the benchmark datasets when setting the missing rate as 0.5.

TABLE V
AVERAGE RUNNING TIME OVER DIFFERENT MISSING RATES BY ELEVEN IMC APPROACHES

Datasets	MKKM-ZF	MKKM-MF	MKKM-KNN	BSV	Concat	UEAF	IMC-AGL	UIMC	IK-MKKM	EE-IMVC	UT-IMKSC
UCI-3views	1.08	1.14	2.05	0.95	0.82	16.63	571.31	144.38	19.76	2.26	115.85
COIL20	0.63	0.68	1.10	1.12	1.32	42.48	112.2	194.32	9.25	1.73	42.28
Reuters	0.39	0.46	0.93	0.51	0.62	9.96	24.96	106.64	2.75	0.97	61.12
BBCSports	0.07	0.08	0.09	1.24	1.26	6.54	28.23	34.74	0.34	0.39	5.45
ORL	0.12	0.13	0.16	0.37	0.46	24.71	24.46	134.49	1.01	1.46	6.11
Yale	0.03	0.03	0.04	0.13	0.15	15.80	6.16	124.51	0.17	0.45	0.52

On each dataset, we run every IMC approach 20 times with missing rate from 0.1 to 0.5.

to complete the iteration process (please see Figs. 12 to 14). In the future, we expect to enhance the time efficiency of our approach via bipartite graph technique [58] and facilitated strategy [59].

E. Ablation Study

This subsection conducts the ablation study to verify the effectiveness of the proposed unified one-step framework. In the experiments, three interested components, namely multi-view kernel imputation, tensorized subspace clustering and affinity matrix construction are seamlessly combined as baselines for comparison. These baselines can be treated as special cases of our UT-IMKSC approach, and the corresponding descriptions of them are given as follows:

- *Tensorized Multi-view Kernel Subspace Clustering with Mean Filling and Affinity Matrix Computation (TMKSC-MF-AMC):* This is a three-step case of the proposed UT-IMKSC approach. In this approach, the multi-view imputation (i.e., mean filling), tensorized subspace clustering as well as affinity matrix construction are separately performed in a three-step manner.
- *Tensorized Incomplete Multi-view Kernel Subspace Clustering with Affinity Matrix Computation (TIMKSC-AMC):* This is a two-step case of the proposed UT-IMKSC approach. In the first step, this approach unifies the multi-view imputation and tensorized subspace clustering in a joint manner. In the second step, this approach constructs the final affinity matrix by means of the multi-view representations.

TABLE VI
ABLATION STUDY OVER NMI AND ACCURACY ON THREE BENCHMARK DATASETS

Metric		NMI					Acc				
Dataset	Missing rate	TMKSC-MF-AMC	TIKMSC-AMC	UTKMSC-MF	UT-IMKSC	Impro	TMKSC-MF-AMC	TIKMSC-AMC	UTKMSC-MF	UT-IMKSC	Impro
COIL20	0.1	78.7	[98.0]	95.7	**99.8**	1.84%	68.6	[96.8]	88.1	**99.9**	3.20%
	0.2	82.0	[95.5]	78.4	**97.4**	1.99%	75.2	**95.0**	54.3	[92.2]	-2.95%
	0.3	80.5	[90.3]	75.2	**97.1**	7.53%	72.4	[87.2]	47.2	**92.6**	6.19%
	0.4	[79.1]	76.2	69.4	**96.5**	22.00%	70.3	[71.7]	44.7	**92.9**	29.57%
	0.5	66.5	[67.0]	60.6	**97.0**	44.78%	[56.7]	51.9	36.1	**94.0**	65.78%
	Avg. Score	77.4	[85.4]	75.9	**97.6**	14.29%	68.6	[80.5]	54.1	**94.3**	17.14%
ORL	0.1	83.5	[99.2]	96.8	**100.0**	0.81%	69.7	[96.4]	90.9	**100.0**	3.73%
	0.2	94.3	[98.9]	95.8	**99.8**	0.91%	87.9	[95.8]	90.0	**99.8**	4.18%
	0.3	88.9	[98.7]	95.9	**99.3**	0.61%	80.3	[95.8]	94.3	**98.3**	2.61%
	0.4	77.8	[97.3]	91.4	**99.6**	2.36%	64.6	[94.3]	84.7	**99.5**	5.51%
	0.5	62.4	[96.5]	84.3	**99.3**	2.90%	44.2	[93.1]	76.5	**96.3**	3.44%
	Avg. Score	81.4	[98.1]	92.8	**99.6**	1.53%	69.3	[95.1]	87.3	**98.4**	3.47%
Yale	0.1	63.8	[86.5]	79.6	**89.8**	3.82%	62.3	[82.0]	75.6	**87.3**	6.46%
	0.2	46.9	[83.9]	74.6	**86.2**	2.74%	42.3	[80.8]	71.8	**85.1**	5.32%
	0.3	48.9	**85.5**	74.0	[79.3]	-7.25%	43.4	**85.2**	69.8	[77.1]	-9.51%
	0.4	36.3	[76.9]	66.8	**79.3**	3.12%	30.3	[74.7]	65.5	**75.2**	0.67%
	0.5	47.6	49.9	[64.9]	**66.7**	33.67%	42.9	46.2	**62.7**	[62.6]	-0.16%
	Avg. Score	48.7	[76.5]	72.0	**80.2**	4.84%	44.2	[73.8]	69.1	**77.5**	5.01%

* Note that "Avg. Score" indicates the average results for NMI or Accuracy metrics.
On each dataset, the best scores are highlighted in bold, while the second best one in [brackets].

- *Unified and Tensorized Multi-view Kernel Subspace Clustering with Mean Filling (UTKMSC-MF):* This is a two-step case of the proposed UT-IMKSC approach. In the first step, the UTMKSC-MF approach fills the missing kernels with meaning filling algorithm. In the second step, this approach performs the tensorized subspace clustering and affinity matrix construction in a unified manner.

The NMI and Accuracy results on three benchmark data sets (i.e., COIL20, ORL and Yale) are reported in Table VI. Notice that, "impro" here means the improvements of our approach over the baselines for ablation study, which can be computed by $\frac{\text{The UT–IMKSC result} - \text{The best baseline result}}{\text{The best baseline result}} \times 100\%$. As can be observed, the two-step based approach, namely TIMKSC-AMC has achieved the best performance among three baselines. More importantly, we could see that our proposed UT-IMKSC approach can obtain relatively significant improvements over the three-step or two-step based baselines. Notably, our approach has shown increase of 44.78% and 65.78% improvements compared by TIKMSC-AMC in terms of NMI and Accuracy, respectively. Overall, the necessity and effectiveness of our unified framework have been empirically verified by the ablation results in Table VI.

V. CONCLUSION

This paper develops a novel IMC approach termed UT-IMKSC from the unified and tensorized perspectives. As far as we know, this is the first attempt for incomplete multi-view kernel subspace clustering. Notably, the multi-view kernel imputation as well as incomplete tensorized subspace clustering are jointly leveraged into a unified framework, upon which the low-rank tensor representation and their common affinity matrix can be mutually enhanced in a one-step manner. By means of this mechanism, our approach can flexibly exploit the latent high-level correlation among various views, while robustly discovering the nonlinear subspace structure for robust incomplete multi-view clustering. Experimental results on numerous incomplete multi-view datasets validate the effectiveness and superiority of our UT-IMKSC approach. Following the literatures in deep contrastive learning [60], [61] [62], how to extend the proposed UT-IMKSC approach would be further investigated in the future work.

REFERENCES

[1] S. Li, Y. Jiang, and Z. Zhou, "Partial multi-view clustering," in *Proc. AAAI Conf. Artif. Intell.*, 2014, pp. 1968–1974.
[2] H. Zhao, H. Liu, and Y. Fu, "Incomplete multi-modal visual data grouping," in *Proc. Int. Joint Conf. Artif. Intell.*, 2016, pp. 2392–2398.
[3] M. Hu and S. Chen, "One-pass incomplete multi-view clustering," in *Proc. AAAI Conf. Artif. Intell.*, 2019, pp. 3838–3845.
[4] M. Hu and S. Chen, "Doubly aligned incomplete multi-view clustering," in *Proc. Int. Joint Conf. Artif. Intell.*, 2019, pp. 2262–2268.
[5] J. Wen, Y. Xu, and H. Liu, "Incomplete multiview spectral clustering with adaptive graph learning," *IEEE Trans. Cybern.*, vol. 50, no. 4, pp. 1418–1429, Apr. 2020.
[6] W. Zhuge, H. Tao, T. Luo, L. L. Zeng, C. Hou, and D. Yi, "Joint representation learning and clustering: A framework for grouping partial multiview data," *IEEE Trans. Knowl. Data Eng.*, vol. 34, no. 8, pp. 3826–3840, Aug. 2022.
[7] J. Wen et al., "Adaptive graph completion based incomplete multi-view clustering," *IEEE Trans. Multimedia*, vol. 23, pp. 2493–2504, 2021.
[8] J. Wen, Z. Zhang, Z. Zhang, L. Fei, and M. Wang, "Generalized incomplete multiview clustering with flexible locality structure diffusion," *IEEE Trans. Cybern.*, vol. 51, no. 1, pp. 101–114, Jan. 2021.
[9] X. Liu et al., "Multiple kernel k k-means with incomplete kernels," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 42, no. 5, pp. 1191–1204, May 2020.
[10] X. Zhu et al., "Localized incomplete multiple kernel k-means," in *Proc. Int. Joint Conf. Artif. Intell.*, 2018, pp. 3271–3277.
[11] M. Li, J. Xia, H. Xu, Q. Liao, X. Zhu, and X. Liu, "Localized incomplete multiple kernel k-means with matrix-induced regularization," *IEEE Trans. Cybern.*, vol. 53, no. 6, pp. 3479–3492, Jun. 2023.
[12] Y. Wang, X. Lin, L. Wu, W. Zhang, Q. Zhang, and X. Huang, "Robust subspace clustering for multi-view data by exploiting correlation consensus," *IEEE Trans. Image Process.*, vol. 24, no. 11, pp. 3939–3949, Nov. 2015.

[13] X. Cai, D. Huang, G.-Y. Zhang, and C.-D. Wang, "Seeking commonness and inconsistencies: A jointly smoothed approach to multi-view subspace clustering," *Inf. Fusion*, vol. 91, pp. 364–375, 2023.
[14] Z. Kang et al., "Partition level multiview subspace clustering," *Neural Netw.*, vol. 122, pp. 279–288, 2020.
[15] J. Lv, Z. Kang, X. Lu, and Z. Xu, "Pseudo-supervised deep subspace clustering," *IEEE Trans. Image Process.*, vol. 30, pp. 5252–5263, 2021.
[16] Z. Ma, Z. Kang, G. Luo, L. Tian, and W. Chen, "Towards clustering-friendly representations: Subspace clustering via graph filtering," in *Proc. 28th ACM Int. Conf. Multimedia*, 2020, pp. 3081–3089.
[17] Y. Wang, L. Wu, X. Lin, and J. Gao, "Multiview spectral clustering via structured low-rank matrix factorization," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 29, no. 10, pp. 4833–4843, Oct. 2018.
[18] C. Zhang, Q. Hu, H. Fu, P. Zhu, and X. Cao, "Latent multi-view subspace clustering," in *Proc. IEEE 30th Conf. Comput. Vis. Pattern Recognit.*, 2017, pp. 4279–4287.
[19] C. Hou and Z. H. Zhou, "One-pass learning with incremental and decremental features," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 40, no. 11, pp. 2776–2792, Nov. 2018.
[20] C. Hou, L.-L. Zeng, and D. Hu, "Safe classification with augmented features," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 41, no. 9, pp. 2176–2192, Sep. 2019.
[21] G.-Y. Zhang, Y.-R. Zhou, X.-Y. He, C.-D. Wang, and D. Huang, "One-step Kernel multi-view subspace clustering," *Knowl.-Based Syst.*, vol. 189, 2019, Art. no. 105126.
[22] P. Zhang et al., "Consensus one-step multi-view subspace clustering," *IEEE Trans. Knowl. Data Eng.*, vol. 34, no. 10, pp. 4676–4689, Oct. 2022.
[23] Z. Kang, Z. Lin, X. Zhu, and W. Xu, "Structured graph learning for scalable subspace clustering: From single view to multiview," *IEEE Trans. Cybern.*, vol. 52, no. 9, pp. 8976–8986, Sep. 2022.
[24] Y. Liu, Z. Lu, J. Li, T. Yang, and C. Yao, "Deep image-to-video adaptation and fusion networks for action recognition," *IEEE Trans. Image Process.*, vol. 29, pp. 3168–3182, 2020.
[25] Y. Liu, K. Wang, G. Li, and L. Lin, "Semantics-aware adaptive knowledge distillation for sensor-to-vision action recognition," *IEEE Trans. Image Process.*, vol. 30, pp. 5573–5588, 2021.
[26] S. Wang et al., "Fast parameter-free multi-view subspace clustering with consensus anchor guidance," *IEEE Trans. Image Process.*, vol. 31, pp. 556–568, 2022.
[27] X. Cao, C. Zhang, H. Fu, S. Liu, and H. Zhang, "Diversity-induced multi-view subspace clustering," in *Proc. IEEE 28th Conf. Comput. Vis. Pattern Recognit.*, 2015, pp. 586–594.
[28] Q. Wang, J. Cheng, Q. Gao, G. Zhao, and L. Jiao, "Deep multi-view subspace clustering with unified and discriminative learning," *IEEE Trans. Multimedia*, vol. 23, pp. 3483–3493, 2021.
[29] S. Wang, C. Li, Y. Li, Y. Yuan, and G. Wang, "Self-supervised information bottleneck for deep multi-view subspace clustering," *IEEE Trans. Image Process.*, vol. 32, pp. 1555–1567, 2023.
[30] M.-S. Chen et al., "Representation learning in multi-view clustering: A literature review," *Data Sci. Eng.*, vol. 7, pp. 225–241, 2022.
[31] W. Shao, L. He, and P. S. Yu, "Multiple incomplete views clustering via weighted nonnegative matrix factorization with l2, 1 regularization," in *Proc. Joint Eur. Conf. Mach. Learn. Knowl. Discov. Databases*, 2015, pp. 318–334.
[32] J. Liu et al., "A novel consensus learning approach to incomplete multi-view clustering," *Pattern Recognit.*, vol. 115, 2021, Art. no. 107890.
[33] J. Yin and S. Sun, "Incomplete multi-view clustering with cosine similarity," *Pattern Recognit.*, vol. 123, 2022, Art. no. 108371.
[34] L. Li, Z. Wan, and H. He, "Incomplete multi-view clustering with joint partition and graph learning," *IEEE Trans. Knowl. Data Eng.*, vol. 35, no. 1, pp. 589–602, Jan. 2023.
[35] X. Yu, H. Liu, Y. Lin, Y. Wu, and C. Zhang, "Auto-weighted sample-level fusion with anchors for incomplete multi-view clustering," *Pattern Recognit.*, vol. 130, 2022, Art. no. 108772.
[36] X. Liu et al., "Efficient and effective incomplete multi-view clustering," in *Proc. AAAI Conf. Artif. Intell.*, 2019, pp. 4392–4399.
[37] M. E. Kilmer and C. D. Martin, "Factorization strategies for third-order tensors," *Linear Algebra Appl.*, vol. 435, no. 3, pp. 641–658, 2011.
[38] Y. Xie, D. Tao, W. Zhang, Y. Liu, L. Zhang, and Y. Qu, "On unifying multi-view self-representations for clustering by tensor multi-rank minimization," *Int. J. Comput. Vis.*, vol. 126, no. 11, pp. 1157–1179, 2018.
[39] J. Wu, Z. Lin, and H. Zha, "Essential tensor learning for multi-view spectral clustering," *IEEE Trans. Image Process.*, vol. 28, no. 12, pp. 5910–5922, Dec. 2019.
[40] G.-Y. Zhang, Y.-R. Zhou, C.-D. Wang, D. Huang, and X.-Y. He, "Joint representation learning for multi-view subspace clustering," *Expert Syst. Appl.*, vol. 166, 2021, Art. no. 113913.
[41] H. Hu, Z. Lin, J. Feng, and J. Zhou, "Smooth representation clustering," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2014, pp. 3834–3841.
[42] R. H. Bartels and G. W. Stewart, "Solution of the matrix equation ax xb= c," *Commun. ACM*, vol. 15, no. 9, pp. 820–826, 1972.
[43] Y. Xie, W. Zhang, Y. Qu, L. Dai, and D. Tao, "Hyper-Laplacian regularized multilinear multiview self-representations for clustering and semisupervised learning," *IEEE Trans. Cybern.*, vol. 50, no. 2, pp. 572–586, Feb. 2020.
[44] W. Hu, D. Tao, W. Zhang, Y. Xie, and Y. Yang, "The twist tensor nuclear norm for video completion," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 28, no. 12, pp. 2961–2973, Dec. 2017.
[45] F. Nie, X. Wang, and H. Huang, "Clustering and projected clustering with adaptive neighbors," in *Proc. 20th ACM SIGKDD Int. Conf. Knowl. Discov. Data Mining*, 2014, pp. 977–986.
[46] F. Nie, G. Cai, and X. Li, "Multi-view clustering and semi-supervised classification with adaptive neighbours," in *Proc. 31st AAAI Conf. Artif. Intell.*, 2017, pp. 2408–2414.
[47] Y. Liang, D. Huang, and C.-D. Wang, "Consistency meets inconsistency: A unified graph learning framework for multi-view clustering," in *Proc. IEEE 19th Int. Conf. Data Mining*, 2019, pp. 1204–1209.
[48] L. Huang, H.-Y. Chao, and C.-D. Wang, "Multi-view intact space clustering," *Pattern Recognit.*, vol. 86, pp. 344–353, 2019.
[49] Y. Liang, D. Huang, C.-D. Wang, and P. S. Yu, "Multi-view graph learning by joint modeling of consistency and inconsistency," *IEEE Trans. Neural Netw. Learn. Syst.*, early access, Jul. 27, 2022, doi: 10.1109/TNNLS.2022.3192445.
[50] J. Liu et al., "Self-representation subspace clustering for incomplete multi-view data," in *Proc. 29th ACM Int. Conf. Multimedia*, 2021, pp. 2726–2734.
[51] D. Huang, C.-D. Wang, and J.-H. Lai, "Fast multi-view clustering via ensembles: Towards scalability, superiority, and simplicity," *IEEE Trans. Knowl. Data Eng.*, vol. 35, no. 11, pp. 11388–11402, Nov. 2023.
[52] J. Lao, D. Huang, C.-D. Wang, and J.-H. Lai, "Towards scalable multi-view clustering via joint learning of many bipartite graphs," *IEEE Trans. Big Data*, vol. 32, no. 10, no. 1, pp. 77–91, 2024.
[53] X. Liu, "Incomplete multiple Kernel alignment maximization for clustering," *IEEE Trans. Pattern Anal. Mach. Intell.*, early access, Oct. 01, 2021, doi: 10.1109/TPAMI.2021.3116948.
[54] J. Wen et al., "A survey on incomplete multiview clustering," *IEEE Trans. Syst., Man, Cybern. Syst.*, vol. 53, no. 2, pp. 1136–1149, Feb. 2023.
[55] J. Wen, Z. Zhang, Y. Xu, B. Zhang, L. Fei, and H. Liu, "Unified embedding alignment with missing views inferring for incomplete multi-view clustering," in *Proc. AAAI Conf. Artif. Intell.*, 2019, pp. 5393–5400.
[56] X. Fang, Y. Hu, P. Zhou, and D. O. Wu, "Unbalanced incomplete multi-view clustering via the scheme of view evolution: Weak views are meat; strong views do eat," *IEEE Trans. Emerg. Topics Comput. Intell.*, vol. 6, no. 4, pp. 913–927, Aug. 2022.
[57] X. Liu et al., "Efficient and effective regularized incomplete multi-view clustering," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 43, no. 8, pp. 2634–2646, Aug. 2021.
[58] S.-G. Fang, D. Huang, X.-S. Cai, C.-D. Wang, C. He, and Y. Tang, "Efficient multi-view clustering via unified and discrete bipartite graph learning," *IEEE Trans. Neural Netw. Learn. Syst.*, early access, Apr. 03, 2023, doi: 10.1109/TNNLS.2023.3261460.
[59] G.-Y. Zhang, D. Huang, and C.-D. Wang, "Facilitated low-rank multi-view subspace clustering," *Knowl.-Based Syst.*, vol. 260, 2023, Art. no. 110141.
[60] Y. Li, P. Hu, Z. Liu, D. Peng, J. T. Zhou, and X. Peng, "Contrastive clustering," in *Proc. AAAI Conf. Artif. Intell.*, 2021, pp. 8547–8555.
[61] Y. Liu, K. Wang, L. Liu, H. Lan, and L. Lin, "TCGL: Temporal contrastive graph for self-supervised video representation learning," *IEEE Trans. Image Process.*, vol. 31, pp. 1978–1993, 2022.
[62] J. Liu, X. Liu, Y. Yang, Q. Liao, and Y. Xia, "Contrastive multi-view Kernel learning," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 45, no. 8, pp. 9552–9566, Aug. 2023.

Guang-Yu Zhang received the B.S. degree from Chang'an University, Xi'an, China, and the M.Sc. and Ph.D. degrees from Sun Yat-sen University, Guangzhou, China, in 2014 and 2016, respectively. He joined South China Agricultural University, Guangzhou, in 2021, where he is currently a Lecturer with the Department of Computer Science. His research interests include multi-view learning and applications of evolutionary computation in data mining.

Dong Huang (Member, IEEE) received the B.S. degree in computer science from the South China University of Technology, Guangzhou, China, in 2009, and the M.Sc. and Ph.D. degrees in computer science from Sun Yat-sen University, Guangzhou, in 2011 and 2015, respectively. In 2015, he joined South China Agricultural University, Guangzhou, where he is currently an Associate Professor and the Deputy Head with the Department of Computer Science. From 2017 to 2018, he was a visiting Fellow with the School of Computer Science and Engineering, Nanyang Technological University, Singapore. He has authored or coauthored more than 80 papers in refereed journals and conferences, including IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING, IEEE TRANSACTIONS ON CYBERNETICS, IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS, *ACM Transactions on Knowledge Discovery from Data*, *Pattern Recognition*, KDD, AAAI, and ICDM. His research interests include data mining and machine learning, and more specifically focus on ensemble clustering, multi-view clustering, and large-scale clustering. He was the recipient of the 2020 ACM Guangzhou Rising Star Award.

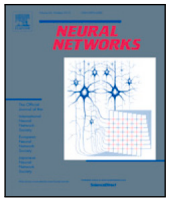
Chang-Dong Wang (Senior Member, IEEE) received the Ph.D. degree in computer science from Sun Yat-sen University, Guangzhou, China, in 2013. From January 2012 to November 2012, he was a visiting Student with the University of Illinois at Chicago, Chicago, IL, USA. In 2013, he joined Sun Yat-sen University, where he is currently an Associate Professor with the School of Data and Computer Science. His research interests include machine learning and data mining. He has authored or coauthored more than 70 scientific papers in international journals and conferences such as IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING, IEEE TRANSACTIONS ON CYBERNETICS, IEEE TRANSACTIONS ON NEURAL NETWORKS AND LEARNING SYSTEMS, *ACM Transactions on Knowledge Discovery from Data*, IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS, IEEE TRANSACTIONS ON INDUSTRIAL INFORMATICS, IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS PART C, KDD, AAAI, IJCAI, CVPR, ICDM, CIKM, and SDM. His ICDM 2010 paper was the recipient of the Honorable Mention for Best Research Paper Awards. He was also the recipient of the 2012 Microsoft Research Fellowship Nomination Award. and 2015 Chinese Association for Artificial Intelligence (CAAI) Outstanding Dissertation. He is an Associate Editor for *Journal of Artificial Intelligence Research*.

Neural Networks 179 (2024) 106529

Contents lists available at ScienceDirect

Neural Networks

journal homepage: www.elsevier.com/locate/neunet



Full Length Article

Tensorized Incomplete Multi-view Kernel Subspace Clustering

Guang-Yu Zhang [a], Dong Huang [a,*], Chang-Dong Wang [b,c]

[a] College of Mathematics and Informatics, South China Agricultural University, China
[b] School of Computer Science and Engineering, Sun Yat-sen University, China
[c] Guangdong Key Laboratory of Information Security Technology, China

ARTICLE INFO

Keywords:
Multi-view incomplete clustering
Kernelized model
Tensor subspace clustering
Unified framework

ABSTRACT

Recently considerable advances have been achieved in the incomplete multi-view clustering (IMC) research. However, the current IMC works are often faced with three challenging issues. First, they mostly lack the ability to recover the nonlinear subspace structures in the multiple kernel spaces. Second, they usually neglect the high-order relationship in multiple representations. Third, they often have two or even more hyper-parameters and may not be practical for some real-world applications. To tackle these issues, we present a Tensorized Incomplete Multi-view Kernel Subspace Clustering (TIMKSC) approach. Specifically, by incorporating the kernel learning technique into an incomplete subspace clustering framework, our approach can robustly explore the latent subspace structure hidden in multiple views. Furthermore, we impute the incomplete kernel matrices and learn the low-rank tensor representations in a mutual enhancement manner. Notably, our approach can discover the underlying relationship among the observed and missing samples while capturing the high-order correlation to assist subspace clustering. To solve the proposed optimization model, we design a three-step algorithm to efficiently minimize the unified objective function, which only involves one hyper-parameter that requires tuning. Experiments on various benchmark datasets demonstrate the superiority of our approach. The source code and datasets are available at: https://www.researchgate.net/publication/381828300_TIMKSC_20240629.

1. Introduction

The widely spread of information technology has brought a large amount of multi-view data. In practice, the multi-view data may suffer from missing views for different reasons, such as sensor errors and machine anomaly, which leads to the incomplete multi-view data. For example, in image clustering, some images are only represented by either visual or textual features while some other images may share both two feature sets. In biomedical informatics, some patients may choose both blood test and magnetic resonance imaging for disease diagnosis, but some other patients possibly just take one of these two tests. In recent years, the frequent occurrence of incomplete multi-view data gives great interests in the incomplete multi-view clustering research, which aims to achieve satisfactory clustering performance by integrating the complementary knowledge from multiple complete or incomplete views.

In the literature, there contains three categories of techniques to address the incomplete multi-view clustering problem, including the matrix factorization-based approaches, the graph-based approaches, and the kernel-based approaches. Among these techniques, the matrix factorization-based approaches have gained increasing attention due to their well-defined mathematical frameworks. In this category of approaches, the core idea is to build a consensus representation by capturing the complementary information across complete and incomplete views. To this end, many efforts have been made in developing different matrix factorization-based approaches. Li, Jiang, and Zhou (2014) designed a partial matrix factorization-based approach, where a consensus representation is learned for paired samples, and two individual representations are obtained for unpaired samples simultaneously. Along this line, Zhao, Liu, and Fu (2016) further exploited the global neighborhood structure from learned latent low-dimensional subspace for the matrix factorization-based formulation. In spite of this, early matrix factorization-based approaches usually require that at least one view is complete, which restricts their feasibility for complex IMC cases. To address this issue, Hu and Chen (2019a) presented a doubly aligned matrix factorization-based approach, by introducing a Laplacian graph regularizer via nonnegative representation learning. By means of a one-pass strategy, Hu and Chen (2019b) further proposed another work that directly computes the clustering labels from the shared representation. However, many of the existing matrix factorization approaches neglect the manifold structure hidden in various views,

* Corresponding author.
E-mail addresses: guangyuzhg@foxmail.com (G.-Y. Zhang), huangdonghere@gmail.com (D. Huang), changdongwang@hotmail.com (C.-D. Wang).

https://doi.org/10.1016/j.neunet.2024.106529
Received 30 March 2024; Received in revised form 29 June 2024; Accepted 7 July 2024
Available online 9 July 2024

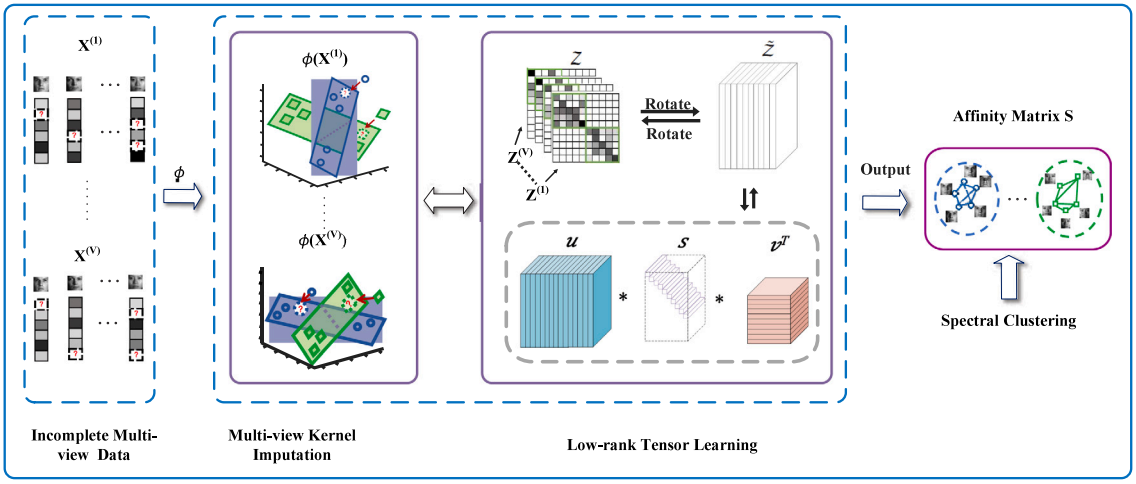


Fig. 1. The flowchart of the proposed approach.

lacking the ability to characterize the geometry relationship among the observed and missing samples.

Besides the matrix factorization-based approaches, the graph-based approaches are another important category in incomplete multi-view clustering. Particularly, the main goal of graph-based approaches is to learn a shared spectral embedding based on the well-known spectral graph theory. For example, Wen, Xu, and Liu (2018) presented a pioneering work that introduces the graph learning technique to automatically obtain a consistent spectral embedding across different views. Meanwhile, Wen, Zhang, Zhang, Fei, and Wang (2020) developed a novel approach by flexibly combining the graph learning and matrix factorization techniques. Following this line, a joint graph-based work is presented in Li, Wan, and He (2021), where the graph constructing technique and the partition fusion technique is adopted for generalized performance. Although the success that has been achieved, there are still two drawbacks to the prior graph-based works. On the one hand, they mostly tend to fill the missing views with zeros values or K-nearest neighbors (K-NN), which separates the graph recovering and graph clustering into two individual steps. On the other hand, many of them have more than two hyper-parameters or regularizers, which may suffer from the multi-tuning issue in some complex realistic scenarios.

Apart from the graph-based approaches, other researchers attempt to solve the incomplete multi-view clustering problem by employing kernel mapping techniques. Earliest works on this category were mainly derived from the traditional kernel k-means clustering. For instance, Liu, Zhu, Li, Wang, et al. (2019) developed a multiple kernel k-means approach with incomplete kernels. With the help of a one-stage strategy, this approach is able to conduct kernel imputation and kernel k-means clustering simultaneously. Moreover, Zhu et al. (2018) incorporated the local structure into an incomplete multiple kernel clustering framework and developed a Localized Incomplete kernel k-means clustering approach. Though these kernel-based approaches have achieved remarkable progress, yet many of them directly perform k-means clustering on the imputed kernel spaces, which may limit their subspace representability for handling high-dimensional tasks. Very recently, some kernel-based approaches like (Xia, Yang, Yang, & Li, 2023) attempt to solve the incomplete multi-view clustering problem via kernelized graph learning. While achieving some progress, these approaches cannot explore the high-order correlations among different views, restricting their ability to incorporate richer information for strengthening better clusterings.

In light of this, this paper presents a new incomplete multi-view clustering approach termed Tensorized Incomplete Multi-view Kernel subspace Clustering (TIMKSC). In comparison with previous works, the proposed approach aims to incorporate the kernel mapping technique and low-rank tensor factorization into the self-expressive subspace clustering formulation, upon which the missing kernels[1] are iteratively filled by exploiting the intrinsic subspace structure within each incomplete views, while the high-order correlations can be captured effectively to depict the global clustering characteristic across various views. As shown in Fig. 1, these two processes are seamlessly connected to promote the incomplete multi-view clustering performance. Experiments are conducted on eight incomplete multi-view benchmarks. The corresponding results confirm the significant advantage of our TIMKSC approach. To summarize, the contributions of this paper are threefold:

- This paper proposes to solve the incomplete multi-view subspace clustering problem from the nonlinear and tensorized perspectives, among which a iterative algorithm is provided with facilitated strategy.
- This paper bridges the gap between the multi-view kernel imputation and the tensor subspace clustering, so that these two processes can benefit each other in a self-expressive subspace clustering framework.
- This paper shows that the joint use of the low-rank tensor factorization and the kernel learning techniques in incomplete subspace clustering framework can significantly enhance the clustering performance.

The remainder of this paper is arranged as follows. Section 2 gives a brief review of the related works. Section 3 is dedicated to describe the proposed approach TIMKSC. Section 4 presents the experimental results on a number of incomplete multi-view benchmarks in order to evaluate the performance of TIMKSC. Section 5 provides the conclusion of this paper.

2. Related works

In this section, we review the related works on multi-view clustering (especially multi-view subspace clustering) and incomplete multi-view clustering, with special emphasis on their recent developments.

The multi-view clustering (MVC) has been widely applied to handle many scientific problems, among which the subspace-based approaches plays an central role. As an early attempt, Abavisani and Patel (2018) extended the single-view subspace clustering model to a multi-view formulation by enhancing the low-rank and sparse property across

[1] Multi-view kernel imputation means the dynamical imputation strategy for multiple view-specific kernels during the optimization, where each kernel corresponds to one view. Please see Section 3.3.3 for more details.

diverse views. Cai, Huang, Zhang, and Wang (2023) jointly modeled the multi-view consistency and inconsistency for robust subspace clustering. Zhang, Hu, Fu, Zhu, and Cao (2017) presented a novel approach that performs low-rank subspace clustering in the shared latent space. Chen, Huang, Wang, Huang and and Lai (2021) devised a relaxed multi-view subspace approach by enforcing a rank constraint on the Laplacian graph, which simultaneously recovers the latent embedding space and global subspace representation in a unified framework. Further, Zhang, Zhou, He, Wang, and Huang (2019) developed a robust multi-view kernel subspace clustering approach, which directly generates the common affinity matrix under a one-step formulation. In addition, Chen, Wang, et al. (2021) presented another multi-view kernel subspace clustering approach, where the self-paced learning, the tensor representation learning, and the kernel mapping are seamlessly combined to enhance the clustering performance. Recently, Lu, Liu, Long, Chen, and Zhu (2023) designed a tensorized multi-view subspace clustering approach via Tucker-O-Minus Decomposition. In terms of the large-scale problem, Kang, Lin, Zhu, and Xu (2021) designed a matrix factorization-based subspace clustering approach with structured graph learning. Wang et al. (2021) further developed a fast multi-view subspace clustering approach which can be guided by a set of consensus anchors of different views. In addition to the aforementioned approaches, some researches seek to handle the MVC problems with nonlinear projection strategy, which inspires the emergence of the spectral-based approaches and kernel-based approaches. In order to address the sub-optimal issue in conventional two-step scheme, Tang et al. (2022) proposed an novel unified framework for multi-view spectral clustering, among which the spectral embedding and k-means are integrated together to achieve the discrete clustering results. Moving forward, Wang, Tang, Wan, et al. (2023) further developed an efficient and effective approach for multi-view spectral clustering. This approach targets at form a latent partition representation by fusing the anchor representations of different views. In Wang, Tang, Zheng, et al. (2023), derived a fast multiple kernel clustering approach with low computational burden, which utilizes two approximated partition matrices to replace the original $n \times n$ original individual partition w.r.t. each base kernel. Inherited from the state-of-art approach in Liu (2022), Liu (2023) presented an elegant framework for multiple kernel clustering, where the hyperparameter-free model can be efficient solved by minimization–minimization–maximization optimization algorithm. These traditional multi-view clustering approaches require that the completeness of multi-view data. However, it is inevitable to generate the incomplete data views in practice, which naturally leads to the incomplete multi-view clustering research.

The incomplete multi-view clustering (IMC) has been a hot research topic in pattern recognition and machine learning. During the past few years, many IMC approaches have been proposed based on the matrix factorization models. In Shao, He, and Yu (2015), proposed a two-step matrix factorization-based approach with an $l_{2,1}$-regularizer. This approach fills the missing samples with the average values of the observed samples, and then builds a low-dimensional subspace via the weighted nonnegative matrix factorization technique. Further, Shao, He, Lu, and Philip (2016) devised an online IMC approach, which utilizes the weighted matrix factorization model to characterize the latent consensus across multiple views. Liu, Teng, et al. (2021) introduced a consensus learning approach to address the IMC problem, which aims to efficiently explore the complementary information from the observed samples. In Wen, Yan, et al. (2020), devised a generalized matrix factorization-based model with flexible locality structure diffusion, which simultaneously captures the local structure and individual representation for strengthening clustering structure. Recently, Deng et al. (2023) designed a simple yet effective IMC approach via projective learning and matrix factorization. Based on the feature space recovery viewpoint, Long, Zhu, Comon, Ren, and Liu (2023) proposed a scalable matrix factorization-based model with low-rank tensor ring-based consistency learning. Despite the obtained progress, these matrix factorization approaches mostly fill the missing views with some static values in advance, which may lead to suboptimal results or even ill clusters. Furthermore, many of them neglect the similarity graph structure among multiple views, which limits their ability to discover the semantic information among the observed and missing views.

To explore the intrinsic graph structure of multiple views, many approaches have recently been developed. In Liang, Liu, Bai, Cao, and Wang (2022), conducted the incomplete multi-view graph partitioning by means of the local and global structural co-regularization. To solve the large-scale IMC problem, Yu, Liu, Lin, Wu, and Zhang (2022) designed a fast graph-based approach from a sample-level perspective. In Wang, Zong, Liu, Yang, and Zhou (2019), a perturbation-oriented incomplete graph-based approach was devoted to weight each views for seeking a consensus Laplacian matrix. Next, Zhuge et al. (2020) developed a joint scheme for the incomplete multi-view grouping, where both the view-specific representation learning and the discrete spectral clustering are seamlessly combined for robust clustering results. Following this line, Liang, Liu, et al. (2022) devised a novel graph fusion-based IMC approach. Different from the previous graph-based approaches, this approach utilizes the sample-level weights to alleviate the latent influence of the missing samples. Besides, Li, Wan, and He (2021) further designed a graph-based IMC approach with a joint graph learning and partitioning strategy, which formulates the hidden space partition and the shared graph learning into a joint optimization model. Liu, Li, Wu, et al. (2023) developed a self-guided IMC approach from the graph propagation perspective. Despite the significant progress, most of the existing graph-based IMC approaches fail to model the subspace structure hidden in multi-view data, restricting their ability for handling high-dimensional tasks. Moreover, many graph-based IMC approaches involve some complex optimization models with two or even more hyper-parameters or regularizers, which undermine their practicability for realistic applications.

Apart from the matrix factorization based approaches and the graph based approaches, another popular technique for incomplete multi-view clustering is the kernel learning based approaches. Lately, Liu (2021) developed a representative approach for incomplete multiple kernel clustering, where the incomplete kernel imputation and the multi-kernel alignment are integrated together for boosting the clustering performance. In addition, Li, Xia, et al. (2021) presented a robust kernel-based approach with the matrix-induced regularization, which is able to strengthen the diversity and complementarity of different views. In Xia et al. (2023), a graph-oriented kernelized approach for incomplete multi-view clustering was proposed, where the graph learning, kernel imputation and clustering analysis are optimized in a mutual reinforcement manner. Although these kernel-based approaches have shown promising clustering performance for the IMC task, yet few of them have considered the high-order correlations among multiple views.

More recently, the rapid emergence of the deep learning technique (Xu et al., 2022; Yang, Deng, Dang, & Tao, 2021) also draw considerable attention toward the deep learning-based IMC research (Jin, Wang, Dong, Liu, & Zhu, 2023; Lin et al., 2021; Liu, Wen, et al., 2023). For example, Wang, Chang, Fu, Wen, and Zhao (2022b) presented a novel IMC approach by means of the cross-view relation transfer. Yang et al. (2022) developed a deep IMC approach with a novel contrastive learning paradigm, which makes the first effort to overcome the partially view-unaligned problem and partially sample-missing problem simultaneously. After that, Xu et al. (2023) proposed an imputation-free IMC approach. Concretely, this approach utilizes an adaptive feature projection strategy to avoid the imputation process of missing views. In contrast to the other categories of IMC methods, the deep learning-based approaches can efficiently discover the intrinsic information hidden in the complete and incomplete views. Recently, Wang, Chang, Fu, Wen, and Zhao (2022a) proposed a robust deep IMC approach via a cross-view contrastive learning framework. Within an end-to-end framework, it can perform view-specific representation learning

as well as cluster-level data aligning for incomplete graph contrastive clustering. However, these approaches still suffer from the parameter fine-tuning issue. For this kind of clustering methods, the robustness of the learned representation is heavily dependent on the structure of neural network and the predefined hyper-parameter settings. Therefore, it remains an open problem how to design a simple yet effective approach for incomplete kernel multi-view subspace clustering.

For more comprehensive literatures about incomplete multi-view clustering, please refer to the survey in Wen et al. (2022).

3. Methodology

This section firstly introduces the main notations throughout this paper. After that, the background of multi-view subspace clustering (MSC) is reviewed. On this basis, this section describes our TIMKSC model in detail. Next, a three-step algorithm is presented to optimize our approach. Finally, this section analyzes the time complexity of our TIMKSC model.

3.1. Main notations

In this paper, the 3-mode tensors are deployed by calligraphy letters, matrices by upper-case letters, and vector by lower-case letters. Notably, we denote $\mathrm{P}_{i:}$, $\mathrm{P}_{:j}$ and P_{ij} as the ith row, jth column, and ij-th element of matrix P, respectively. Besides, we denote $\mathcal{P}(i,:,:)$, $\mathcal{P}(:,j,:)$ and $\mathcal{P}(:,:,k)$ as the ith horizontal slice, jth lateral slice and kth frontal slice of 3-mode tensor $\mathcal{P} \in \mathbb{R}^{n_1\times n_2\times n_3}$, respectively. The operator $\overline{\mathcal{P}} = \mathrm{fft}(\mathcal{P},[],3)$ stands for the Discrete Fourier transform (DFT) along the third dimension of tensor $\mathcal{P}$, while the inverse DFT of tensor $\mathcal{P}$ can be denoted by $\mathcal{P} = \mathrm{ifft}(\overline{\mathcal{P}},[],3)$, conversely. For briefly, matrix $\mathrm{P}^{(k)}$ is used to denote the frontal slice $\mathcal{P}(:,:,k)$.

Based on the above notations, five block-based operators and the basic definitions on tensor, can be seen in the papers (Kilmer & Martin, 2011; Zhang, Zhou, Wang, Huang, & He, 2021). In what follows, two important definitions of t-SVD are given in detail.

Definition 1 (*t-SVD Kilmer & Martin, 2011*)**.** Given a tensor $\mathcal{P} \in \mathbb{R}^{n_1\times n_2\times n_3}$, the corresponding t-SVD could be denoted as

$$\mathcal{P} = \mathcal{W} * \mathcal{O} * \mathcal{Q}^T, \tag{1}$$

in which $\mathcal{W} \in \mathbb{R}^{n_1\times n_1\times n_3}$ and $\mathcal{Q} \in \mathbb{R}^{n_2\times n_2\times n_3}$ are two orthogonal tensors, and $\mathcal{O} \in \mathbb{R}^{n_1\times n_2\times n_3}$ is a f-diagonal tensor.

Definition 2 (*t-SVD Based Tensor Nuclear Norm Kilmer & Martin, 2011*)**.** Given a tensor $\mathcal{P} \in \mathbb{R}^{n_1\times n_2\times n_3}$, the corresponding t-SVD based tensor nuclear norm (i.e., $\| \mathcal{P} \|_{\circledast}$) is defined by the sum of the singular values of all frontal slices, which are as follows:

$$\| \mathcal{P} \|_{\circledast} = \sum_{k=1}^{n_3} \| \mathcal{P}_f^{(k)} \|_* = \sum_{i=1}^{min(n_1,n_2)} \sum_{k=1}^{n_3} | O_f^{(k)}(i,i) | \tag{2}$$

in which $O_f^{(k)}(i,i)$ is computed by the t-SVD of frontal slices of $\mathcal{P}_f$, i.e., $\mathcal{P}_f^{(k)} = \mathcal{W}_f^{(k)}\mathcal{O}_f^{(k)}\mathcal{Q}_f^{(k)T}$. Especially, Fig. 2 shows the t-SVD sample of a tensor $\mathcal{P} \in \mathbb{R}^{n_1\times n_2\times n_3}$.

3.2. Our TIMKSC model

Multi-view subspace clustering (MSC) plays a critical role in various scientific fields. For the present MSC studies, the common objective relies on capturing the low-dimensional subspaces embedded in multiple views. Especially, the popular assumption (also called the self-expressive property), which assumes that every data object could be linearly represented by the other objects, is essential to the theory and practical researches. Formally, we denote the input data matrix as $\mathrm{X} = [\mathrm{X}^{(1)^T}, \ldots, \mathrm{X}^{(V)^T}]^T$, where $\mathrm{X}^{(v)} = [\mathrm{X}^{(v)}_{:1}, \ldots, \mathrm{X}^{(v)}_{:n}] \in \mathbb{R}^{d_v\times n}$ represents the vth view data sub-matrix consisting of d_v features. If we only consider the specificity of each view, the self-expression formulation in multi-view scenario could be formulated as:

$$\mathrm{X}^{(v)} = \mathrm{X}^{(v)}\mathrm{Z}^{(v)} + \mathrm{E}^{(v)}, \forall v, \tag{3}$$

Here $\mathrm{Z}^{(v)} \in \mathbb{R}^{n\times n}$, $\mathrm{E}^{(v)} \in \mathbb{R}^{d_v\times n}$ denote the vth view subspace representation and vth view error matrix, respectively. Based on this, we can directly give the general model for MSC as follows:

$$\begin{aligned} &\min_{\{\mathrm{Z}^{(v)},\mathrm{E}^{(v)}\}_{v=1}^V} \sum_{v=1}^V \|\mathrm{E}^{(v)}\|_p + \lambda \sum_{v=1}^V \|\mathrm{Z}^{(v)}\|_q \\ &\text{s.t. } \mathrm{X}^{(v)} = \mathrm{X}^{(v)}\mathrm{Z}^{(v)} + \mathrm{E}^{(v)}, \forall v, \end{aligned} \tag{4}$$

Here $\|\cdot\|_p$ and $\|\cdot\|_q$ denote the specific matrix norms on variables $\mathrm{E}^{(v)}$ and $\mathrm{Z}^{(v)}$, respectively. If we utilize the Frobenius norm to characterize the reconstruction error term and substitute the constraint into objective function, then the above formulation could be formulated by:

$$\min_{\{\mathrm{Z}^{(v)}\}_{v=1}^V} \sum_{v=1}^V \|\mathrm{X}^{(v)} - \mathrm{X}^{(v)}\mathrm{Z}^{(v)}\|_F^2 + \lambda \sum_{v=1}^V \|\mathrm{Z}^{(v)}\|_q \tag{5}$$

Notice that, the practical multi-view data (especially for nonlinear multi-view scenario) are often coupled with complex structure. Therefore, a flexible solution is to integrate the kernel learning (Gönen & Alpaydın, 2011; Liu, Zhu, Li, Wang, et al., 2019) into the multi-view subspace clustering model. The brief introduction of kernel learning (Lin et al., 2021; Liu, Zhu, Li, Wang, et al., 2019) is given as follows. Let us denote $\phi(\cdot)$ as the kernel function that maps the data objects into the Hilbert space. Then the vth view data sub-matrix $\mathrm{X}^{(v)}$ could be mapped into $\phi(\mathrm{X}^{(v)}) = [\phi(\mathrm{X}^{(v)}_{:1}), \ldots, \phi(\mathrm{X}^{(v)}_{:n})]$ by a implicit but known kernel. On this basis, we denote $\mathrm{K}^{(v)} \in \mathbb{R}^{n\times n}$ as the vth view kernel matrix, and its elements represent the nonlinear similarity between data objects in the kernel space. Accordingly, the kernel matrix in the vth could be calculated by:

$$\mathrm{K}^{(v)} = \langle \phi(\mathrm{X}^{(v)}), \phi(\mathrm{X}^{(v)}) \rangle. \tag{6}$$

By calculating the kernel matrices for different views, we can obtain a general model for multi-view kernel subspace clustering (MKSC), which is expressed as follows:

$$\min_{\{\mathrm{Z}^{(v)}\}_{v=1}^V} \sum_{v=1}^V \|\phi(\mathrm{X}^{(v)}) - \phi(\mathrm{X}^{(v)})\mathrm{Z}^{(v)}\|_F^2 + \lambda\|\mathrm{Z}^{(v)}\|_q \tag{7}$$

Furthermore, the reconstruction error term of the above model could be rewritten in the trace form:

$$\begin{aligned} &\sum_{v=1}^V \|\phi(\mathrm{X}^{(v)}) - \phi(\mathrm{X}^{(v)})\mathrm{Z}^{(v)}\|_F^2 \\ &= \sum_{v=1}^V \mathrm{Tr}\{\mathrm{K}^{(v)} - 2\mathrm{K}^{(v)}\mathrm{Z}^{(v)} + \mathrm{Z}^{(v)^T}\mathrm{K}^{(v)}\mathrm{Z}^{(v)}\} \\ &= \sum_{v=1}^V \mathrm{Tr}\{\mathrm{K}^{(v)}(\mathrm{I} - \mathrm{Z}^{(v)} - \mathrm{Z}^{(v)^T} + \mathrm{Z}^{(v)}\mathrm{Z}^{(v)^T})\} \end{aligned} \tag{8}$$

Next, we come into the incomplete multi-view clustering. Let p_v denote the index set for the vth view observed samples, while $\mathrm{K}_p^{(v)}$ denotes the vth view kernel sub-matrix that computed by the observed samples. Previous researches have shown that one-stage based approach can achieve more robust performance over the traditional two-stage based approaches. In order to enhance the effectiveness and scalability, we propose to unify kernel imputation as well as kernel subspace clustering into one-stage framework. To be specific, the corresponding model is formulated as follows:

$$\begin{aligned} &\min_{\{\mathrm{Z}^{(v)},\mathrm{K}^{(v)}\}_{v=1}^V} \sum_{v=1}^V \mathrm{Tr}\{\mathrm{K}^{(v)}(\mathrm{I} - \mathrm{Z}^{(v)} - \mathrm{Z}^{(v)^T} + \mathrm{Z}^{(v)}\mathrm{Z}^{(v)^T})\} + \lambda \sum_{v=1}^V \|\mathrm{Z}^{(v)}\|_q \\ &\text{s.t. } \mathrm{K}^{(v)}(p_v,p_v) = \mathrm{K}_p^{(v)}, \mathrm{K}^{(v)} \succeq \mathbf{0}, \forall v, \end{aligned} \tag{9}$$

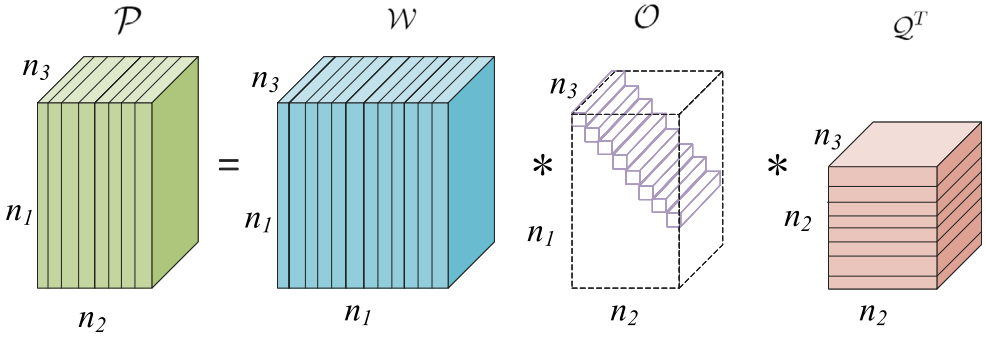


Fig. 2. The t-SVD sample of a tensor $\mathcal{P} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$.

in which $\mathbf{0} \in \mathbb{R}^{n \times n}$ represents the zero matrix. The constraint $\mathrm{K}^{(v)}(p_v, p_v) = \mathrm{K}_p^{(v)}$ is used to ensure that $\mathrm{K}^{(v)}$ maintains the observed entries during the optimization. To further enhance the robustness of this model, a potential solution relies to further exploit the high-order correlation from different views. Inspired by the low-rank tensor learning in multi-view clustering (Wu, Lin, & Zha, 2019; Xie, Tao, et al., 2018; Zhang et al., 2021), we incorporate the t-SVD-based tensor norm into our model (9), which stacks the subspace representations $\{\mathrm{Z}^v\}_{v=1}^V$ into a 3-mode tensor $\mathcal{Z}$ and then rotates its size to $n * V * n$, as shown in Fig. 3. By this means, the multi-view high-order correlations (i.e., the high-order complementary information) can be effectively explored by computing the row-wise (sample-specific) and the column-wise (view-specific) frontal tensor slices (Xie, Tao, et al., 2018; Xie, Zhang, Qu, Dai & and Tao, 2018). Under the incomplete subspace clustering scheme, this design can provide a deeper insight into the multi-view subspace representations. Thus, the final objective function of the proposed approach can be formulated as

$$\begin{aligned}
&\min_{\{\mathrm{Z}^{(v)},\mathrm{K}^{(v)}\}_{v=1}^V} \sum_{v=1}^{V} \mathrm{Tr}\{\mathrm{K}^{(v)}(\mathrm{I} - \mathrm{Z}^{(v)} - \mathrm{Z}^{(v)^T} + \mathrm{Z}^{(v)}\mathrm{Z}^{(v)^T})\} + \lambda\|\mathcal{Z}\|_{\circledast} \\
&\text{s.t. } \mathrm{K}^{(v)}(p_v, p_v) = \mathrm{K}_p^{(v)}, \mathrm{K}^{(v)} \succeq \mathbf{0}, \forall v, \\
&\quad \mathcal{Z} = \boldsymbol{\Phi}(\mathrm{Z}^{(1)}, \mathrm{Z}^{(2)}, \ldots, \mathrm{Z}^{(V)}).
\end{aligned} \tag{10}$$

Here λ is a positive balancing parameter. Compared to the current IMC approaches, our approach can jointly impute the multi-view kernel matrices while learning a series of low-rank tensor representations in a one-stage framework. By this means, the learning processes of kernel imputations and tensor subspace clustering are beneficial to each other in a mutual promotion manner. Moreover, the objective function in Eq. (10) only involves one balancing parameter, which efficiently reduces the burden of extra parameter tuning. Additionally, by inheriting the advantage of kernel learning and low-rank tensor factorization techniques, our approach has a great potential to achieve promising clustering performance. This can be validated by the comprehensive experiments on practical multi-view benchmark datasets (Please refer to Section 4 for detailed information).

3.3. The optimization algorithm

As shown in Eq. (10), the optimization problem could not to be solved directly. Nevertheless, we can use an ADMM-based algorithm to efficiently solve this problem. First of all, by introducing an auxiliary variable $\mathcal{C}$, the optimization problem is reformulated as the following equivalence problem:

$$\begin{aligned}
&\mathcal{L}\left(\{\mathrm{Z}^{(v)}\}_{v=1}^V, \mathcal{C}, \{\mathrm{K}^{(v)}\}_{v=1}^V\right) \\
&= \sum_{v=1}^{V} \mathrm{Tr}\{\mathrm{K}^{(v)}(\mathrm{I} - \mathrm{Z}^{(v)} - \mathrm{Z}^{(v)^T} + \mathrm{Z}^{(v)}\mathrm{Z}^{(v)^T})\} + \lambda\|\mathcal{C}\|_{\circledast} \\
&\quad + \langle \mathcal{Y}, \mathcal{Z} - \mathcal{C}\rangle + \frac{\mu}{2}\|\mathcal{Z} - \mathcal{C}\|_F^2, \\
&\text{s.t. } \mathrm{K}^{(v)}(p_v, p_v) = \mathrm{K}_p^{(v)}, \mathrm{K}^{(v)} \succeq \mathbf{0}, \forall v, \\
&\quad \mathcal{C} = \boldsymbol{\Phi}(\mathrm{C}^{(1)}, \mathrm{C}^{(2)}, \ldots, \mathrm{C}^{(V)}).
\end{aligned} \tag{11}$$

in which tensor $\mathcal{Y}_2$ is a Lagrange multiplier, while μ is a penalty parameter. Besides, $\langle \cdot, \cdot \rangle$ means the Frobenius inner product between two given matrices. For efficiently solving the problem in Eq. (11), an iterative algorithm is applied to update the variables $\{\mathrm{Z}^{(v)}\}_{v=1}^V$, $\mathcal{C}$ and $\{\mathrm{K}^{(v)}\}_{v=1}$ by turns. In what follows, the three update rules are introduced one by one.

3.3.1. The rule for updating $\{\mathrm{Z}^{(v)}\}_{v=1}^V$

While fixing the other variables, the optimization subproblem w.r.t. variable $\mathrm{Z}^{(v)}$ can be solved by (Note that $\boldsymbol{\Phi}_{(v)}^{-1}(\mathcal{Z}) = \mathrm{Z}^{(v)}$ and $\boldsymbol{\Phi}_{(v)}^{-1}(\mathcal{C}) = \mathrm{C}^{(v)}$):

$$\begin{aligned}
&\min_{\mathrm{Z}^{(v)}} \mathrm{Tr}(\mathrm{Z}^{(v)^T}\mathrm{K}^{(v)}\mathrm{Z}^{(v)}) - 2\mathrm{Tr}(\mathrm{K}^{(v)}\mathrm{Z}^{(v)}) \\
&+ \langle \mathrm{Y}^{(v)}, \mathrm{Z}^{(v)} - \mathrm{C}^{(v)}\rangle + \frac{\mu}{2}\|\mathrm{Z}^{(v)} - \mathrm{C}^{(v)}\|_F^2
\end{aligned} \tag{12}$$

After setting the derivative of this subproblem to be zero, we can obtain the following closed-form solution:

$$\mathrm{Z}^{(v)} = (\mathrm{K}^{(v)} + \frac{\mu}{2}\mathrm{I}_n)^{-1}(\mathrm{K}^{(v)} + \frac{\mu}{2}\mathrm{C}^{(v)} - \frac{1}{2}\mathrm{Y}^{(v)}) \tag{13}$$

3.3.2. The rule for updating $\mathcal{C}$

While fixing the other variables, the optimization subproblem w.r.t. variable $\mathcal{C}$ can be solved by:

$$\min_{\mathcal{C}} \lambda\|\mathcal{C}\|_{\circledast} + \langle \mathcal{Y}, \mathcal{Z} - \mathcal{C}\rangle + \frac{\mu}{2}\|\mathcal{Z} - \mathcal{C}\|_F^2 \tag{14}$$

Further, this problem can be reduced into the following form:

$$\min_{\mathcal{C}} \lambda\|\mathcal{C}\|_{\circledast} + \frac{\mu}{2}\|\mathcal{C} - \mathcal{Q}\|_F^2 \tag{15}$$

Here $\mathcal{Q} = \mathcal{Z} + \frac{\mathcal{Y}}{\mu}$. Following closely the recent research (Xie, Zhang, et al., 2018), we can solve the problem in Eq. (15) by tensor tubal-shrinkage operator (Hu, Tao, Zhang, Xie, & Yang, 2016):

$$\mathcal{D} = \mathcal{C}_{n_3\pi}(\mathcal{Q}) = \mathcal{U} * \mathcal{C}_{n_3\pi}(\mathcal{O}) * \mathcal{V}^T, \tag{16}$$

where $\pi = \mu/\lambda$ and $\mathcal{Q} = \mathcal{U} * \mathcal{O} * \mathcal{V}^T$. In addition, $\mathcal{C}_{n_3}(\mathcal{Q}) = \mathcal{Q} * \mathcal{W}$, herein, tensor $\mathcal{W} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$ is f-diagonal, and $\mathcal{W}_f(i,i,j) = (1 - \frac{n_3\pi}{\mathcal{O}_f^{(j)}(i;i)})_+$ means the diagonal element of $\mathcal{W}$ in the Fourier domain.

3.3.3. The rule for updating $\{\mathrm{K}^{(v)}\}_{v=1}^V$

While fixing the other variables, the optimization subproblem w.r.t. variable $\mathrm{K}^{(v)}$ can be solved by:

$$\begin{aligned}
&\min_{\mathrm{K}^{(v)}} \mathrm{Tr}(\mathrm{K}^{(v)}\mathrm{H}^{(v)}) \\
&\text{s.t. } \mathrm{K}^{(v)}(p_v, p_v) = \mathrm{K}_p^{(v)}, \mathrm{K}^{(v)} \succeq \mathbf{0}
\end{aligned} \tag{17}$$

where $\mathrm{H}^{(v)} = (\mathrm{I} - \mathrm{Z}^{(v)} - \mathrm{Z}^{(v)^T} + \mathrm{Z}^{(v)}\mathrm{Z}^{(v)^T})$. This is a standard optimization problem for kernel imputation. First of all, we can decompose the

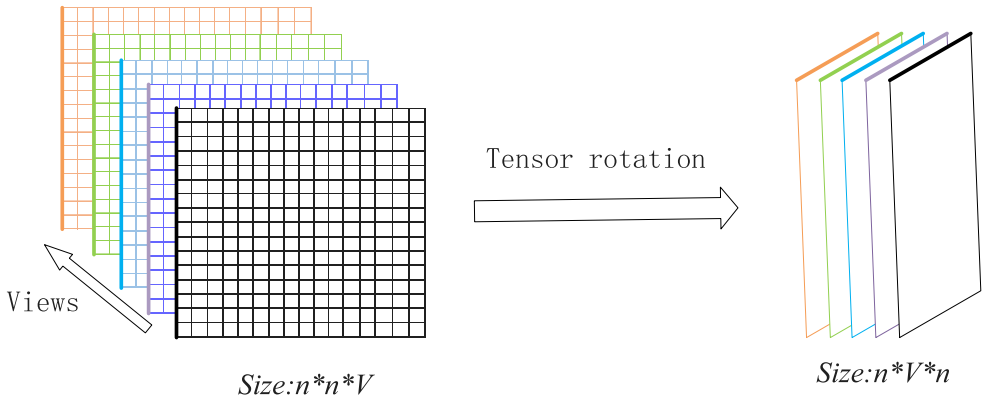


Fig. 3. Example of stacking and rotation operation.

kernel matrix $K^{(v)}$ into the form of $K^{(v)} = A^{(v)}A^{(v)T}$. Meanwhile, matrix $A^{(v)}$ could be further split into two parts, i.e., $A^{(v)} = [A_c{}^{(v)}; A_m{}^{(v)}]$, where $A_c{}^{(v)}$ is the observed part and $A_m{}^{(v)}$ is the missing part. Since matrix $H^{(v)}$ is symmetric, we can adopt the matrix blocking strategy to rewrite the above problem (see Eq. (17)) in the following form:

$$\begin{aligned} &\min_{A^{(v)}} \mathrm{Tr}(H^{(v)}A^{(v)}A^{(v)T}) \\ \Longleftrightarrow &\min_{A^{(v)}} \mathrm{Tr}(A^{(v)T}H^{(v)}A^{(v)}) \\ \Longleftrightarrow &\min_{A_m{}^{(v)}} \mathrm{Tr}\left(\begin{bmatrix} A_c{}^{(v)} \\ A_m{}^{(v)} \end{bmatrix}^T \begin{bmatrix} H_{cc}{}^{(v)} & H_{cm}{}^{(v)} \\ H_{cm}{}^{(v)T} & H_{mm}{}^{(v)} \end{bmatrix} \begin{bmatrix} A_c{}^{(v)} \\ A_m{}^{(v)} \end{bmatrix}\right) \end{aligned} \tag{18}$$

where $H^{(v)}$ is blocked as following forms:

$$H^{(v)} = \begin{bmatrix} H_{cc}{}^{(v)} & H_{cm}{}^{(v)} \\ H_{cm}{}^{(v)T} & H_{mm}{}^{(v)} \end{bmatrix} \tag{19}$$

Taking the derivative of this subproblem w.r.t. variable $A_m{}^{(v)}$, we could get the optimal solution as follows:

$$A_m{}^{(v)} = -H_{mm}{}^{(v)-1}H_{cm}{}^{(v)T}A_c{}^{(v)} \tag{20}$$

After some calculation and deduction, the optimal solution for variable $K^{(v)}$ is given by:

$$K^{(v)} = \begin{bmatrix} K_c{}^{(v)} & -K_c{}^{(v)}G^{(v)} \\ -G^{(v)T}K_c{}^{(v)} & G^{(v)T}K_c{}^{(v)}G^{(v)} \end{bmatrix} \tag{21}$$

where $G^{(v)} = H_{cm}{}^{(v)}\{H_{mm}{}^{(v)-1}\}^T$.

In summary, the aforementioned three steps are iteratively updated until the convergence conditions are met or the iteration number reaches the predefined value. For clarity, the detailed optimization process of TIMKSC is listed in Algorithm 1.

3.4. Time complexity analysis

As can be seen, the optimization problem of TIMKSC approach is divided into four alternative subproblems. First, the time complexity of updating $Z^{(v)}$ is $O((n_c{}^3) + 3n_c{}^2n_m)$. Second, according to the research in Xie, Tao, et al. (2018), the time complexity of updating C is $O(Vn^2\log(n))$. Third, the time complexity of updating $K^{(v)}$ is $O((n_m)^3 + n_m(n_c)^2 + 2n_c{}^2n_m)$. Fourth, the time complexity of updating multiplier Y is O(1). Suppose Algorithm 1 runs T times to convergence, since the condition $n_m{}^{(v)} + n_m{}^{(v)} = n$ is satisfied, the final time complexity of optimization algorithm is $OT(Vn^3 + Vn^2\log(n) + \sum_{v=1}^{V}\{(n_m{}^{(v)})^3 + n_m{}^{(v)}(n_c{}^{(v)})^2\}) = O(TVn^3)$. Here n denotes the number of data objects, n_c represents the number of observed sample, n_m represents the number of missing sample, V represents the number of view.

Algorithm 1 TIMKSC

Input: Incomplete multi-view kernels $\mathcal{K} = \{K^{(1)}, \cdots, K^{(V)}\}$, the maximum iterations it_{max}, and balancing parameter $\lambda > 0$.
Parameter Setup: Set $\mu = 10^{-2}$, $\eta = 1.5$, $\mu_{max} = 10^6$ and $\varepsilon = 10^{-4}$.

Initialization: Initialize $C^{(v)} = I$, $Y^{(v)} = \mathbf{0}, \forall v$. Initialize $K^{(v)}, \forall v$ by zero filling approach. Set $it = 1$.
repeat
  Obtain $\{Z^{(v)}\}_{v=1}^V$ by Eq. (13).
  Obtain C by Eq. (16).
  Obtain $\{K^{(v)}\}_{v=1}^V$ by Eq. (21).
  Obtain the multipliers $\{Y^{(v)}\}_{v=1}^V$ by $Y^{(v)} = Y^{(v)} + \mu(Z^{(v)} - C^{(v)})$.
  Obtain μ by $\mu = \min(\eta\mu, \mu_{\max})$.
  $it = it + 1$.
until The convergence conditions in Eq. (22) to Eq. (24) are met or $it > it_{max}$.
Construct the overall affinity matrix via $M = \sum_{v=1}^{V}(|C^{(v)}| + |C^{(v)T}|)/2$.
Run Ncut algorithm on the overall affinity matrix M, and obtain the clustering label for all data objects.
Output: The clustering label for all data objects.

4. Experiments

In this section, we conduct experiments to compare the proposed approach with several baseline approaches. Note that all the experiments are performed on a PC with an i5-6600 CPU and 64 GB of RAM.

4.1. Experimental materials and metrics

To verify the robustness of our TIMKSC approach, we carry out the substantial experiments on eight popular practical benchmark datasets. It is worth that these datasets are collected with different characteristics, which have been commonly adopted for experimental purpose. Specially, they are object datasets (i.e., ALOI-50, NH-4660 and COIL20 Liang, Huang, Wang, & Yu, 2022), handwritten digit datasets (i.e., UCI-3view Liang & Wang, 2019), news report datasets (i.e, Reuters (Zhang et al., 2019) and BBCSport Huang, Chao, & Wang, 2019), face image dataset (i.e., Yale Liu, Liu, et al., 2021). For clarity, the summarization of eight benchmarks is listed in Table 1.

In our experiments, we follow the existing researches to generate the incomplete multi-view datasets. For each incomplete multi-view dataset, parameter β named the missing rate is introduced, which controls the percentage of samples removed in the partial views. Obviously, the clustering performance in comparison may be affected by the parameter β. To prove this point, we adjust the parameter β within the range, i.e., $[0.1, \ldots, 0.5]$ with step 0.1.

Table 1
The details of eight benchmarks used in the experiments.

Dataset	#Sample	#View	#Class	Dimension
ALOI-50	5400	4	50	Similarity (77), Haralick (13), HSV (64), RGB (125)
NH-4660	4660	3	5	Intensity (2000), LBP (3304), Gabor (6750)
UCI-3view	2000	3	10	PIX (240), FOU (76), MOR (6)
COIL20	1440	3	20	Intensity (1024), LBP (3304), Gabor (6750)
Reuters	1200	5	6	English (2000), German (2000), French (2000), Spanish (2000), Italian (2000)
BBCSports	544	2	5	View1 (3183), View2 (3203)
ORL	400	3	40	LBP (3304), Gabor (6750), Intensity (4096)
Yale	165	3	15	LBP (3304), Gabor (6750), Intensity (4096)

Table 2
Kernel constructions and Parameter Settings of TIMKSC on the benchmark datasets.

	ALOI-50	NH-4660	UCI-3view	COIL20	Reuters	BBCSports	ORL	Yale
View 1	Gaussian	Linear	Gaussian	Gaussian	Gaussian	Linear	Linear	Linear
View 2	Gaussian	Linear	Linear	Gaussian	Gaussian	Linear	Linear	Linear
View 3	Gaussian	Linear	Linear	Gaussian	Gaussian	–	Linear	Gaussian
View 4	Gaussian	-	–	–	Gaussian	–	–	–
View 5	-	-	–	–	Linear	–	–	–
λ	100	1000	100	100	100	10	10	1000

At the same time, two metrics are used to evaluate the quality of clustering results, namely, Normalized Mutual Information (NMI) (Zhang, Wang, Huang, Zheng, & Zhou, 2018), and Accuracy (ACC) (Huang, Kang, Tsang & Xu, 2019). For these metrics, larger values indicate better clustering performance.

4.2. Comparison with other IMC approaches

This subsection empirically investigates our TIMKSC approach by comparing it with several Incomplete Multi-view clustering (IMC) approaches. From different perspectives, the selected IMC approaches for comparison can be categorized into two-stage based approaches and one-stage based approaches. On the one side, two multi-view kernel clustering (MKKM) with traditional imputation approaches, including MKKM with Zero Filling (MKKM-ZF) (Liu, 2021) and MKKM with Mean Filling (MKKM-MF) (Liu, 2021) are employed in the experiments. In additional, two recently proposed subspace-based approaches with zero-filling imputation strategy, are also added to our experiments. They are Multi-view MERA Subspace Clustering (Long, Zhu, Chen, et al., 2023) with Zero Filling (MERA-ZF) and Simple and Scalable Multi-View Tensor Clustering (Long, Wang, Ren, Liu, & Zhu, 2024) with Zero Filling (S^2MVTC-ZF). Besides, two-stage based approaches, i.e., Best Single View (BSV) (Wen et al., 2022) and Concatenating all views (Concat) (Wen et al., 2022) are also selected as baselines for comparison. The above two approaches firstly fill the missing samples with mean filling algorithm, and then apply k-means algorithm to achieve the final clustering results. On the other side, seven popular one-stage based approaches, namely, Unified Embedding Alignment Framework (UEAF) (Wen et al., 2019), Incomplete Multi-view Spectral Clustering with Adaptive Graph Learning (IMC-AGL) (Wen et al., 2018), Unbalanced Incomplete Multi-view Clustering (UIMC) (Fang, Hu, Zhou, & Wu, 2021), Multiple Kernel k-means with Incomplete Kernels (IK-MKKM) (Liu, Zhu, Li, Wang, et al., 2019), Effective and Efficient Multi-view Clustering (EE-IMVC) (Liu et al., 2020; Liu, Zhu, Li, Tang, et al., 2019) are used to investigate their effectiveness. The detail information of these IMC approaches can be found in the original papers.

For ensuring the comprehensiveness and fairness of experiments, the following experimental setting should be strictly observed:

- Based on the guidelines of original papers, the source code of all IMC approaches are directly downloaded. At the same time, the optimal parameter combinations of them are chosen within the range of $[10^{-4}, 10^{-3}, \ldots, 10^{3}, 10^{4}]$.
- For the baseline MERA-ZF, the optimal parameter r is chosen within the range of $[2, \ldots, 16]$ with step 2, while the optimal parameter λ is chosen with the range of $[10^{-4}, 10^{-3}, \ldots, 10^{3}, 10^{4}]$.
- For the baseline S^2MVTC-ZF, the optimal parameter r is chosen within the range of $[2, \ldots, 16]$ with step 2, while the optimal combination of other two parameters are chosen with the range of $[10^{-4}, 10^{-3}, \ldots, 10^{3}, 10^{4}]$.
- For our approach TIMKSC, the optimal parameter λ is also chosen within the same range, i.e., $[10^{-4}, 10^{-3}, \ldots, 10^{3}, 10^{4}]$.
- For the proposed approach, two types of kernels (i.e., the linear kernel and the Gaussian kernel) are used in the experiments. The kernel constructions and the parameter setting of our approach are reported in Table 2. As for Gaussian kernel, the parameter θ is computed by the average of Euclidean distance among the data samples.
- For each tested dataset, we repeat every IMC approach for 20 times, and meanwhile report the best clustering results by cross-validation strategy.
- Under the same computing environment, we repeat each comparison approach for 20 times with randomly incomplete pattern, and report their best clustering performance via cross-validation strategy.
- Apart from these described parameters, the nearest neighbor parameter K in MKKM+KNN is set to 9, as suggested by the corresponding paper.

Tables 3 and 4 report the clustering results on eight benchmark datasets, with the variation of missing ratio in $[0.1, \ldots, 0.5]$. Note that the aggregated results (i.e., NMI and Accuracy) as well as average rank are also recorded to enhance the comprehensiveness and completeness. In summary, the experimental results suggest the following three important findings.

- By comparing the one-stage based approaches with traditional two-stage based approaches, we find that the one-stage based approaches always achieve more promising results than the two-stage based approaches. In terms of Aggregated NMI, one-stage approach IK-MKKM exceeds the best two-stage approaches (i.e., MKKM-KNN and MKKM-MF) by 12.7%, 10.8%, 9.5% and 6.6% on four datasets (e.g., COIL20, BBCSports, ORL and Yale), respectively. This finding indicates that the unified optimization on imputation and clustering is more suitable for solving IMC problems, by comparing with the traditional two-stage based approaches.
- For the two-stage IMC cases, it can be observed that the performance of kernel-based approaches (i.e., MKKM-ZF, MKKM-MF

Table 3
The NMI (%) scores over 20 runs by twelve IMC approaches. On each dataset, the best scores are highlighted in [brackets], while the second best one are in **bold**.

Dataset	Missing rate	MKKM-ZF	MKKM-MF	BSV	Concat	UEAF	IMC-AGL	UIMC	IK-MKKM	EE-IMVC	MERA-ZF	S^2MVTC-ZF	TIMKSC
ALOI-50	0.1	68.5	67.9	58.2	47.1	>24 h	>24 h	58.4	73.8	82.6	>24 h	**87.2**	[94.9]
	0.2	62.3	60.8	49.1	44.2	>24 h	>24 h	52.3	69.6	80.2	>24 h	**82.9**	[94.8]
	0.3	55.8	56.0	47.3	39.1	>24 h	>24 h	46.8	66.4	76.1	>24 h	**86.0**	[94.4]
	0.4	49.3	50.2	39.6	36.0	>24 h	>24 h	42.9	60.9	70.7	>24 h	**84.3**	[94.6]
	0.5	43.6	45.2	32.6	35.7	>24 h	>24 h	36.4	55.2	63.9	>24 h	**83.6**	[94.6]
	Agg NMI	55.9	56.0	45.4	40.4	>24 h	>24 h	47.4	65.2	74.7	>24 h	**84.8**	[94.7]
NH-4660	0.1	67.7	60.4	57.0	57.9	>24 h	>24 h	76.3	75.2	71.9	>24 h	**81.3**	[90.8]
	0.2	64.4	57.5	52.4	54.2	>24 h	>24 h	76.4	63.5	65.6	>24 h	**82.5**	[90.4]
	0.3	52.8	51.7	44.3	61.3	>24 h	>24 h	72.8	62.0	61.5	>24 h	**76.7**	[90.6]
	0.4	45.6	44.8	37.0	53.3	>24 h	>24 h	66.4	41.1	56.9	>24 h	**85.8**	[91.2]
	0.5	37.5	37.8	30.6	32.6	>24 h	>24 h	58.4	36.7	50.6	>24 h	**71.5**	[88.5]
	Agg NMI	53.6	50.4	44.3	51.9	>24 h	>24 h	70.1	55.7	61.3	>24 h	**79.6**	[90.3]
UCI-3views	0.1	42.3	43.1	40.6	65.5	63.9	84.9	85.3	45.9	65.6	[99.6]	95.3	[99.6]
	0.2	38.6	38.3	35.2	57.7	65.0	80.7	81.4	42.4	59.4	[99.5]	94.3	[99.5]
	0.3	34.0	33.8	30.0	43.0	57.0	74.4	74.5	38.2	53.2	[99.4]	91.3	**98.9**
	0.4	29.3	29.1	25.7	37.5	46.6	66.3	67.8	35.3	45.8	[99.3]	84.7	**98.5**
	0.5	24.3	24.0	21.7	30.5	35.7	57.0	58.1	30.9	38.0	[99.0]	87.3	**97.4**
	Agg NMI	33.7	33.7	30.6	46.8	53.6	72.7	73.4	38.5	52.4	[99.4]	90.6	**98.8**
COIL20	0.1	70.6	79.2	63.1	68.3	77.8	89.3	70.1	81.1	75.3	[98.5]	87.2	**92.0**
	0.2	64.1	69.8	56.4	49.0	77.6	87.5	68.6	79.7	69.5	[98.7]	87.5	**92.5**
	0.3	55.5	62.0	51.0	45.2	76.2	84.6	66.4	77.8	61.9	[98.0]	85.6	**94.2**
	0.4	49.5	54.6	41.6	37.9	73.1	79.2	63.6	73.6	53.2	[96.3]	79.6	**93.8**
	0.5	42.0	45.9	36.5	28.8	69.3	71.3	59.6	69.3	42.8	[93.2]	81.6	**95.4**
	Agg NMI	56.3	62.3	49.7	45.8	74.8	82.4	65.7	76.3	60.5	[96.9]	84.3	**93.6**
Reuters	0.1	2.9	29.0	1.7	0.4	29.7	14.3	18.8	29.2	26.7	**86.5**	60.3	[95.1]
	0.2	1.4	28.6	0.8	0.4	32.8	14.4	17.8	29.1	26.2	**81.4**	57.5	[96.4]
	0.3	1.0	28.0	16.2	0.4	30.4	14.1	16.7	28.9	26.3	**85.3**	46.7	[95.0]
	0.4	2.7	26.3	0.4	0.6	31.0	14.5	17.1	28.3	24.8	**66.7**	46.9	[94.1]
	0.5	2.4	20.6	2.7	0.4	28.1	12.6	15.7	27.1	22.0	**78.4**	52.4	[95.5]
	Agg NMI	2.1	26.5	4.3	0.5	30.4	14.0	17.2	28.5	25.2	**79.7**	52.8	[95.2]
BBCSports	0.1	44.9	55.7	45.0	0.8	75.9	82.5	84.7	60.8	74.5	**87.6**	NA	[98.0]
	0.2	31.7	51.1	0.7	3.5	75.6	74.0	77.7	58.4	68.2	**83.7**	NA	[99.3]
	0.3	21.8	48.3	0.9	1.2	69.7	68.7	68.6	56.6	57.3	**80.3**	NA	[100.0]
	0.4	11.9	34.6	0.7	1.5	63.4	57.3	57.7	51.8	47.3	**81.6**	NA	[95.8]
	0.5	6.6	29.7	0.7	0.8	53.6	46.2	45.9	45.9	37.0	**83.6**	NA	[95.9]
	Agg NMI	23.4	43.9	9.6	1.5	67.7	65.8	66.9	54.7	56.9	**83.4**	NA	[97.8]
ORL	0.1	78.8	83.8	63.1	54.8	73.8	**88.6**	84.4	85.7	86.8	[99.2]	NA	[99.2]
	0.2	74.7	78.0	42.7	30.6	72.0	87.8	83.3	84.9	85.8	[99.2]	NA	**98.9**
	0.3	69.4	71.4	42.9	25.8	71.3	85.3	81.1	82.0	83.6	**98.5**	NA	[98.7]
	0.4	63.0	64.7	37.2	23.2	65.3	80.6	76.4	78.4	79.7	**96.9**	NA	[97.3]
	0.5	57.2	57.7	34.4	29.0	60.4	73.1	69.7	71.8	73.2	**93.0**	NA	[96.5]
	Agg NMI	68.6	71.1	44.1	32.7	68.5	83.1	79.0	80.6	81.8	**97.4**	NA	[98.1]
Yale	0.1	53.5	64.6	38.9	40.8	57.8	67.2	68.3	65.8	64.1	[97.7]	NA	**88.5**
	0.2	48.8	60.9	51.6	34.0	57.3	65.9	66.0	65.8	63.4	[98.5]	NA	**85.3**
	0.3	46.7	55.7	44.5	23.5	56.9	64.4	63.0	62.3	60.9	[94.6]	NA	**80.4**
	0.4	42.2	48.4	22.1	28.6	59.6	59.7	60.6	57.1	56.7	[94.8]	NA	**81.4**
	0.5	37.2	40.9	27.4	18.8	51.2	55.0	53.8	55.7	53.2	[91.8]	NA	**72.8**
	Agg NMI	45.7	54.1	36.9	29.1	56.6	62.4	62.3	61.3	59.7	[95.5]	NA	**81.4**

Note that "Agg NMI" indicates the aggregated NMI results, ">24 h" means the baseline exceeds more than 24 hours, and "NA" means the baseline is not applicable to this dataset.

and MKKM-KNN) are more reliable over that of the k-means approaches (i.e., BSV and Concat). This is probably because the kernel-based approaches have capability to exploit the non-linear structure hidden in multiple views, which demonstrates the effectiveness of kernel learning for handling IMC tasks.

- For the one-stage IMC cases, we could see that the graph-based approaches (i.e., UEAF and IMSC-AGL) have shown consistent (or even better) performance over the recently proposed approach EE-IMVC. This demonstrates the effectiveness of graph learning for boosting the clustering performance.
- It is noteworthy that two latest MSC approach with zero-filling strategy (i.e., MERA-ZF and S^2MVTC), have gained remarkable performance against the other baselines. For example, in terms of Aggregated NMI and Aggregated Accuracy, MERA-ZF approach obtains 99.4% and 99.6% on UCI-3views dataset, which significantly outperforms the other two-stage competitors by large margins.
- It turns out that our approach has obtained very competitive results on eight benchmark datasets, with missing rate β varied in the range of $[0.1, \ldots, 0.5]$. Surprisingly, the proposed approach achieves nearly perfect performance on UCI-3viewsa and BBCSports datasets. For the commonly recognized challenging datasets (i.e., Reuters and Yale), our approach has made remarkable (or even breakthrough) improvements compared to the state-of-the-art IMC approaches. Beyond the expectation, our approach outperforms the second-best approach with large margins on Reuters dataset, which obtains an increase of 19.4% and 12.1% improvements over MERA-ZF approach by Aggregated NMI and Aggregated Accuracy, respectively.

In contrast to the comparison approaches, the proposed approach can efficiently capture the high-order correlation among different views. In addition, our TIMKSC approach seeks to flexibly unify multi-view kernel imputation as well as tensor subspace clustering into unified framework, and this may contribute to the advantageous clustering performance of our TIMKSC approach.

Table 4
The Accuracy (%) scores over 20 runs by twelve IMC approaches. On each dataset, the best scores are highlighted in [brackets], while the best one are in **bold**.

Dataset	Missing rate	MKKM-ZF	MKKM-MF	BSV	Concat	UEAF	IMC-AGL	UIMC	IK-MKKM	EE-IMVC	MERA-ZF	S^2MVTC-ZF	TIMKSC
ALOI-50	0.1	56.6	57.4	44.2	32.1	>24 h	>24 h	44.7	66.5	**75.9**	>24 h	53.6	[91.9]
	0.2	52.0	50.2	35.7	31.2	>24 h	>24 h	42.8	61.8	**74.5**	>24 h	51.7	[91.8]
	0.3	43.7	43.7	36.8	29.5	>24 h	>24 h	37.1	57.9	**70.8**	>24 h	52.3	[90.4]
	0.4	36.7	38.2	30.1	27.0	>24 h	>24 h	31.5	51.6	**66.0**	>24 h	46.4	[91.6]
	0.5	31.0	32.3	24.6	24.4	>24 h	>24 h	24.8	46.8	**59.5**	>24 h	46.0	[92.7]
	Agg Acc	44.0	44.4	34.3	28.8	>24 h	>24 h	36.2	56.9	**69.3**	>24 h	50.0	[91.7]
NH-4660	0.1	71.8	69.7	72.9	57.8	>24 h	>24 h	83.6	84.7	86.4	>24 h	**92.5**	[96.8]
	0.2	68.7	70.6	71.3	61.3	>24 h	>24 h	83.4	69.7	83.0	>24 h	**84.4**	[96.5]
	0.3	67.8	64.6	66.0	67.7	>24 h	>24 h	82.5	71.5	81.2	>24 h	**84.6**	[96.5]
	0.4	63.4	57.6	59.0	64.4	>24 h	>24 h	78.6	57.1	**77.8**	>24 h	74.6	[96.4]
	0.5	57.8	50.4	53.4	49.1	>24 h	>24 h	72.0	52.2	70.6	>24 h	**80.6**	[95.9]
	Agg Acc	65.9	62.6	64.5	60.1	>24 h	>24 h	80.0	67.0	79.8	>24 h	**83.3**	[96.4]
UCI-3views	0.1	44.0	45.2	38.6	71.1	70.9	86.5	89.1	46.7	76.7	[99.8]	90.4	[99.8]
	0.2	40.8	41.6	35.1	64.7	78.1	84.9	87.6	45.4	73.0	[99.7]	89.3	[99.7]
	0.3	36.9	37.8	30.8	48.7	69.9	81.7	82.8	41.7	67.5	[99.7]	80.9	**99.6**
	0.4	33.0	34.0	27.7	45.5	56.5	76.0	79.3	39.1	59.8	[99.6]	78.6	**99.4**
	0.5	29.2	28.9	26.1	40.9	46.0	68.9	69.8	34.6	51.9	[99.2]	80.5	**98.9**
	Agg Acc	36.8	37.5	31.7	54.2	64.3	79.6	81.7	41.5	65.8	[99.6]	83.9	**99.5**
COIL20	0.1	63.6	70.8	49.5	55.5	66.8	77.1	54.2	72.4	66.2	[95.2]	67.5	**85.2**
	0.2	55.8	61.4	47.1	36.9	68.9	76.2	51.7	71.7	62.4	[95.6]	71.9	**88.3**
	0.3	46.2	52.7	42.7	37.8	68.0	76.3	50.6	69.6	57.6	[93.0]	65.6	**88.7**
	0.4	40.3	45.8	36.5	32.6	66.3	72.9	50.4	65.1	51.3	**88.9**	58.7	[89.8]
	0.5	34.1	37.4	31.6	26.1	63.5	66.5	47.6	63.1	43.2	**86.3**	62.9	[95.3]
	Agg Acc	48.0	53.6	41.5	37.8	66.7	73.8	50.9	68.4	56.2	[91.8]	65.3	**89.5**
Reuters	0.1	23.5	51.2	18.3	16.9	52.7	35.8	38.9	51.0	46.6	**93.5**	60.5	[98.1]
	0.2	21.7	51.3	17.4	16.9	52.0	36.1	37.4	51.3	46.4	**87.6**	57.5	[98.7]
	0.3	20.5	50.9	34.9	16.9	48.5	35.4	36.8	51.3	47.3	**94.2**	47.8	[98.1]
	0.4	21.9	49.4	17.1	17.2	51.3	37.2	38.6	50.6	45.8	**78.3**	44.8	[97.6]
	0.5	22.2	42.2	19.2	16.9	48.0	35.7	38.3	49.8	43.8	**83.9**	52.1	[98.3]
	Agg Acc	22.0	49.0	21.4	17.0	50.5	36.0	38.0	50.8	46.0	**87.5**	52.5	[98.1]
BBCSports	0.1	64.0	71.9	68.4	35.5	87.4	93.8	95.0	76.6	89.3	**94.4**	NA	[99.4]
	0.2	52.8	65.7	35.7	34.0	91.2	87.8	91.1	73.8	86.7	**93.6**	NA	[99.8]
	0.3	44.6	62.5	35.8	35.8	88.1	87.2	85.3	71.8	78.1	**90.3**	NA	[100.0]
	0.4	36.2	52.8	35.7	35.8	77.6	78.0	78.7	67.6	70.0	**91.7**	NA	[98.7]
	0.5	31.5	46.5	35.7	35.5	69.0	68.1	66.0	63.1	60.4	**93.1**	NA	[98.7]
	Agg Acc	45.8	59.9	42.2	35.3	82.6	83.0	83.2	70.6	76.9	**92.6**	NA	[99.3]
ORL	0.1	63.9	69.5	51.0	38.9	58.1	76.1	68.7	72.1	75.8	[96.6]	NA	**96.4**
	0.2	59.6	63.2	31.7	26.5	56.0	75.8	68.5	72.1	74.4	[96.4]	NA	**95.8**
	0.3	52.7	55.7	34.5	20.7	56.8	73.6	67.2	67.4	71.8	**93.8**	NA	[95.8]
	0.4	44.3	47.7	29.5	16.5	51.4	68.3	62.5	64.9	66.8	**91.2**	NA	[94.3]
	0.5	38.3	40.3	29.5	21.5	46.0	59.9	56.0	58.0	59.9	**83.7**	NA	[93.1]
	Agg Acc	51.8	55.3	35.2	24.8	53.7	70.7	64.6	66.9	69.7	**92.3**	NA	[95.1]
Yale	0.1	49.3	62.2	31.6	31.9	53.5	65.9	67.3	62.5	62.0	[98.2]	NA	**86.5**
	0.2	44.6	57.5	48.5	30.3	51.5	64.6	65.1	62.9	61.4	[98.8]	NA	**87.2**
	0.3	42.7	51.4	41.8	26.7	51.8	63.1	61.8	59.2	58.5	[94.2]	NA	**78.4**
	0.4	37.5	42.9	21.2	27.3	51.5	58.7	59.1	53.6	53.9	[91.2]	NA	**82.6**
	0.5	32.8	37.5	26.1	19.4	46.1	53.2	52.0	51.8	49.9	[90.4]	NA	**79.2**
	Agg Acc	41.4	50.3	33.8	27.1	50.9	61.1	61.1	58.0	57.2	[94.6]	NA	**82.8**

Note that "Agg Acc" indicates the aggregated Accuracy results, ">24 h" means the baseline exceeds more than 24 hours, and "NA" means the baseline is not applicable to this dataset.

4.3. Visualization results

In this subsection, we carry out experiments to visually compare our approach with UEAF approach, where the missing rate β is varied in the range of $[01, 0.3, 0.5]$. Following the existing research in Zhang, Fu, et al. (2018), we use the t-SNE algorithm (Maaten & Hinton, 2008) to plot the visualization results on UCI-3views and BBCSports dataset. From the Figs. 4 and 5, we can clearly observe that the proposed approach clusters different data samples well under different setting of missing rate. Besides, it can be observed that TIMKSC approach obtains less overlap between different categories, in comparison with the competitor (i.e., the separated clusters achieved by TIMKSC are much clearer over those achieved by UEAF).

Further, we provide the intermediate visual comparison between our approach and the UEAF approach, where the number of iterations varies from 10 to 30. The visualization results on the UCI-3views and BBCsports datasets are illustrated in Fig. 6 and Fig. 7, respectively. It can be observed that the cluster structure of TIKMSC becomes more and more compact as the iteration number varied from 10 to 30. Notably, we can also see that the proposed approach has better revealed the block-diagonal structures over the competitor UEAF approach. The obtained visualization results are well consistent with the clustering results in Tables 3 and 4, which demonstrate the effectiveness of our approach in discovering the intrinsic cluster structures.

4.4. Parameter sensitivity analysis

This subsection practically investigates the influence of balancing parameter λ in our approach. Particularly, the balancing parameter λ is adjusted within the range of $[10^{-4}, 10^{-3}, \ldots, 10^{3}, 10^{4}]$, while the missing rate β is set as 0.1. The performance over three metrics (i.e., NMI and Accuracy) on eight benchmark datasets is shown in Fig. 8. From this figure, we can observe that our approach yields relatively stable performance on UCI-3views and ORL datasets, within a wide range of

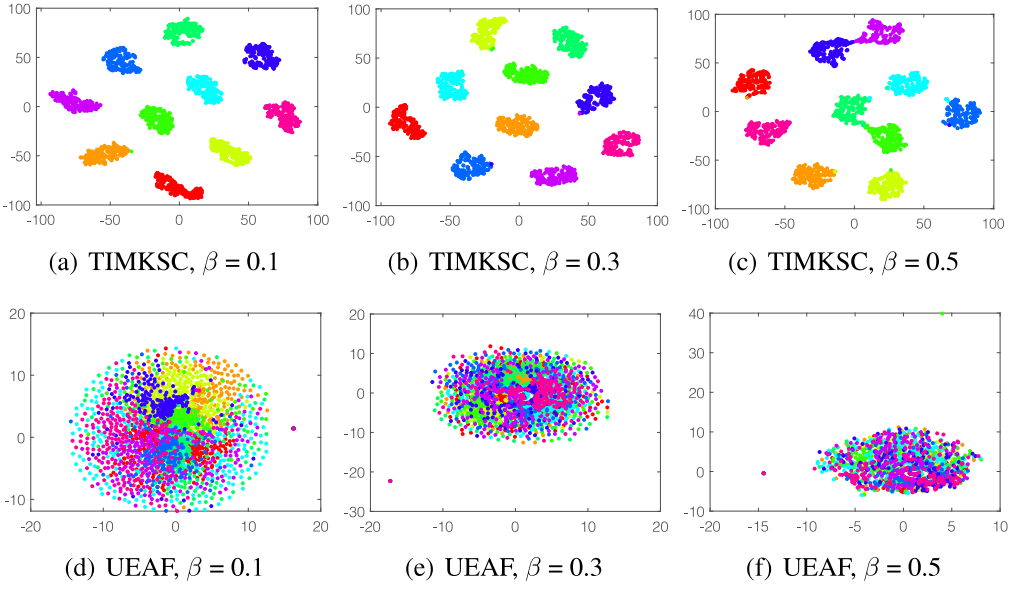


Fig. 4. The t-SNE visualization of two comparison approaches on UCI-3views dataset.

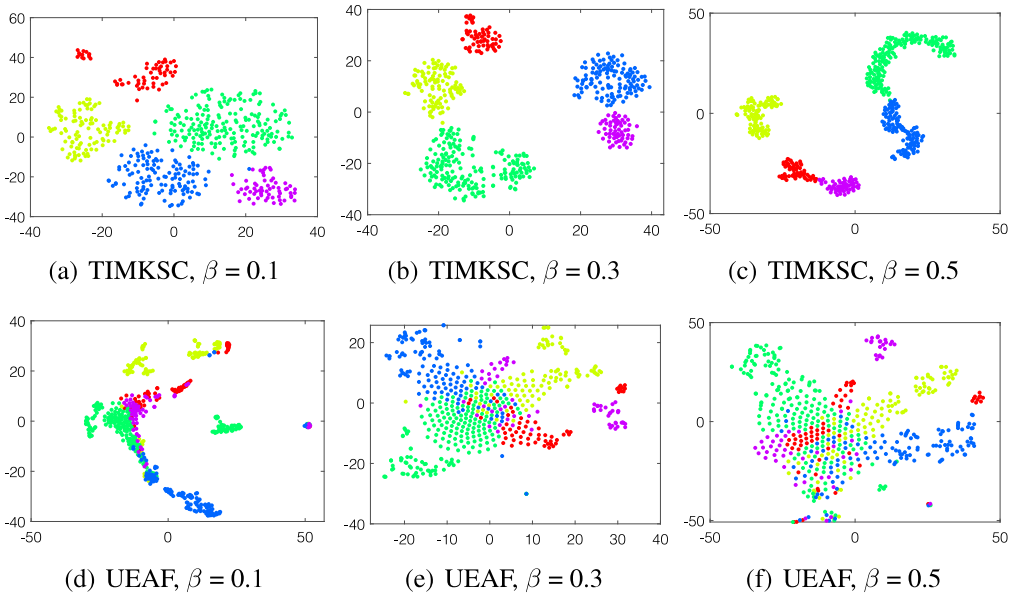


Fig. 5. The t-SNE visualization of two comparison approaches on BBCSports dataset.

parameter λ. When it comes to Reuters and BBCSports datasets, we find that our approach generates relatively competitive results when the parameter λ is set as moderate values (e.g., 10^1 or 10^2). Finally, it can be also seen that, as the value of parameter λ increases, the clustering performance of our approach will gradually increase to a peak value, and then maintains it up with slight degeneration on UCI-3views, ORL and Yale datasets. Based on the above analysis, users of our approach are suggested to adjust the parameter λ within the moderate range ($[10^1, 10^3]$) for real-world applications.

4.5. Convergence and computational time analysis

This subsection analyzes the convergence property of our approach and presents the convergence behavior curves on benchmark datasets. In Algorithm 1, the convergence conditions are determined by three error terms, namely the match error (ME), the iteration error for variable $\mathcal{Z}$ ($\mathrm{IE}_{\mathcal{Z}}$) and the iteration error for variable C (IE_C). To be more specific, they can be defined by:

$$\mathrm{ME} = \sum_{v=1}^{V} \|\mathrm{Z}^{(v)} - \mathrm{C}^{(v)}\|_\infty \tag{22}$$

and

$$\mathrm{IE}_{\mathcal{Z}} = \|\mathcal{Z}_{current} - \mathcal{Z}_{pre}\|_\infty \tag{23}$$

and

$$\mathrm{IE}_C = \|C_{current} - C_{pre}\|_\infty \tag{24}$$

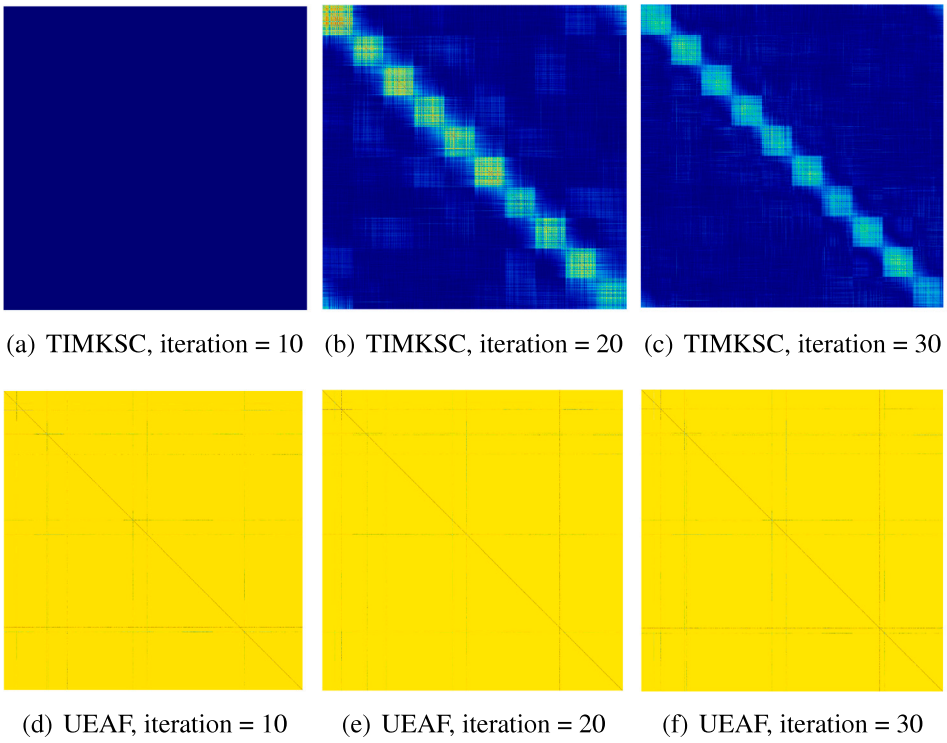


Fig. 6. The visualization of affinity matrices w.r.t. two comparison approaches on UCI-3views dataset.

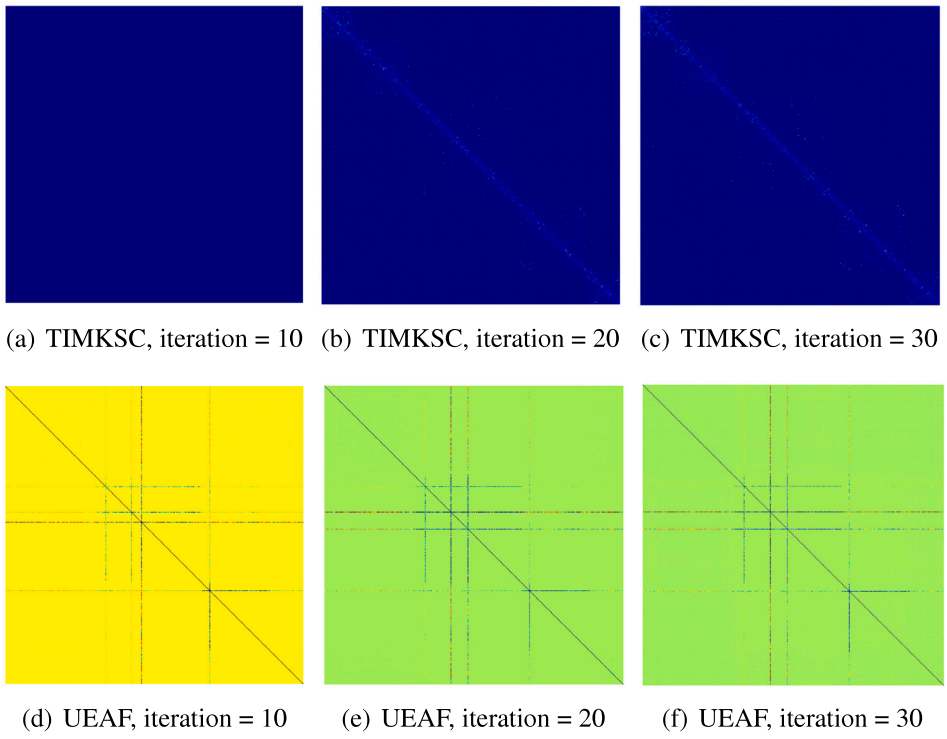


Fig. 7. The visualization of affinity matrices w.r.t. two comparison approaches on BBCSports dataset.

According to the curves displayed in Fig. 9, we can see that the proposed approach has a relatively fast convergence speed. Usually, the three convergence conditions will concurrently meet between 30 to 40 iterations.

Moving forward, we further compare the computational efficiency of different IMC approaches on eight benchmark datasets. According to the experimental results shown in Table 5, we have the following observations. First, it can be seen that the one-stage based approaches usually require more time than the two-stage based approaches. But it is noteworthy that the superiority of the one-stage based approaches in clustering performance compensates for the increase in the computational complexity. Second, we can see that the IK-MKKM, EE-IMVC

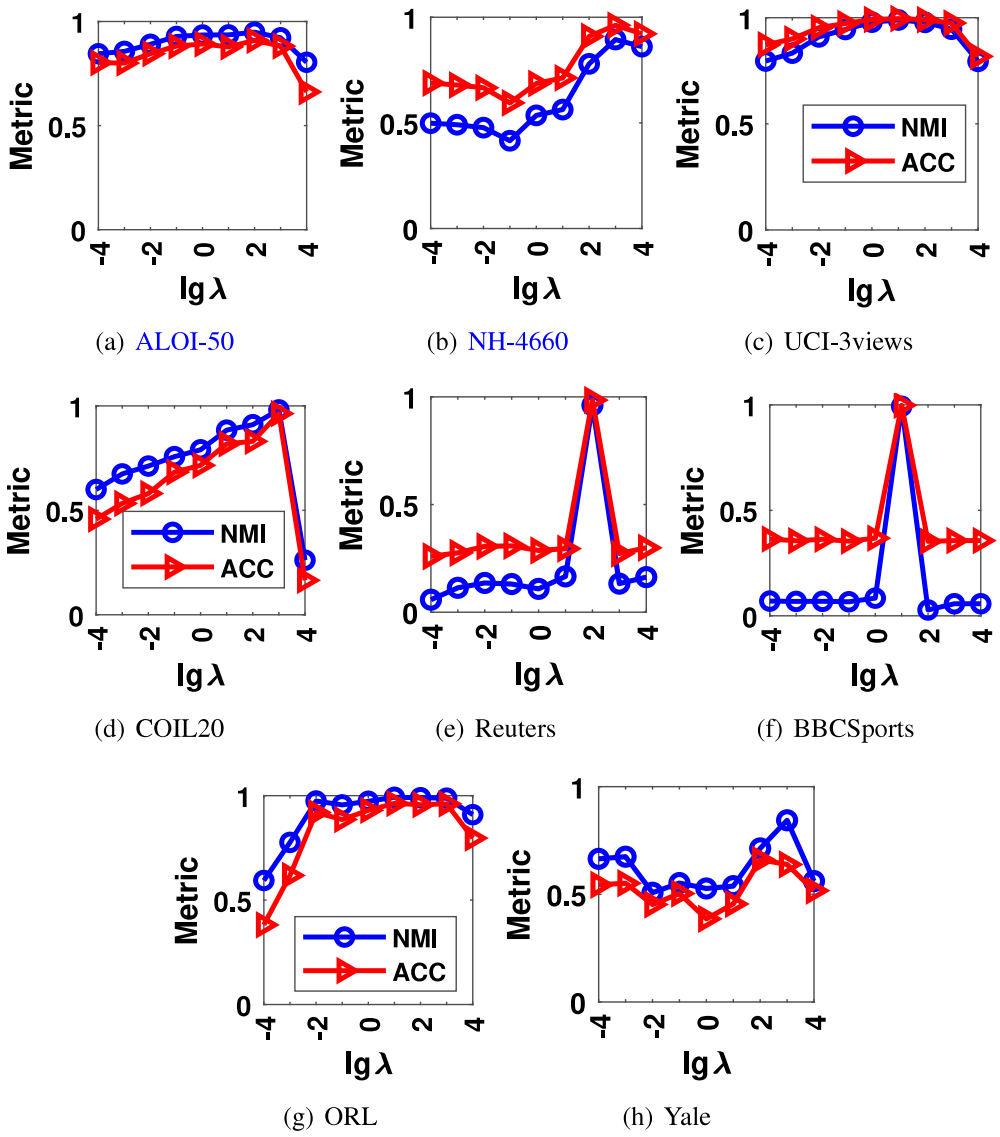


Fig. 8. Analysis on the influence of parameter λ when missing rate is 0.1.

and S^{2}MVTC approaches take less time than the other one-stage based approaches, probably due to their simplified objective models with some fast optimization algorithms. Third, our approach is more efficient than the UEAF, IMC-AGL and UIMC approaches, indicating the computational efficiency in comparison with the mainstream matrix factorization-based approaches. Particularly, the time complexity of the twelve baseline IMC approaches is reported in Table 6. It is noteworthy that the time complexity of our proposed algorithm is still very high, which may be not suitable for large-scale tasks. In the future, we expect to design a more efficient approach by borrowing the theory of bipartite graph formulation (Fang et al., 2023; Long et al., 2024).

4.6. Ablation study

In this subsection, we conduct the ablation study by comparing the proposed approach with a special baseline. The brief introduction of this baseline is as follows:

- **Tensorized Incomplete Multi-view Subspace Clustering (TIMSC) :** TIMSC is a special case of our approach. If we use the linear kernels to construct all complete views, then the proposed TIMKSC approach will degenerates into the TIMSC approach.

According to the parameter setting in Tables 2 and, we use the UCI-3views, Reuters and Yale datasets for ablation study. As can be observed in Figs. 10 and 11, the TIMKSC approach exhibits overall advantages than the linear variant on most datasets. Especially, it has achieved very powerful performance on Reuters dataset, which improves the baseline TIMSC approach by a large margin. This indicates the necessary contribution of kernel learning technique in the proposed approach. Therefore, in some specific tasks, a flexible way is to integrate two components, the kernelized subspace learning and low-rank tensor factorization for better clustering performance.

5. Conclusion

This paper develops a novel approach for incomplete multi-view subspace clustering, which inherits the advantages of kernel learning and low-rank tensor factorization techniques. For the purpose of handling practical datasets with complex structures, our approach is

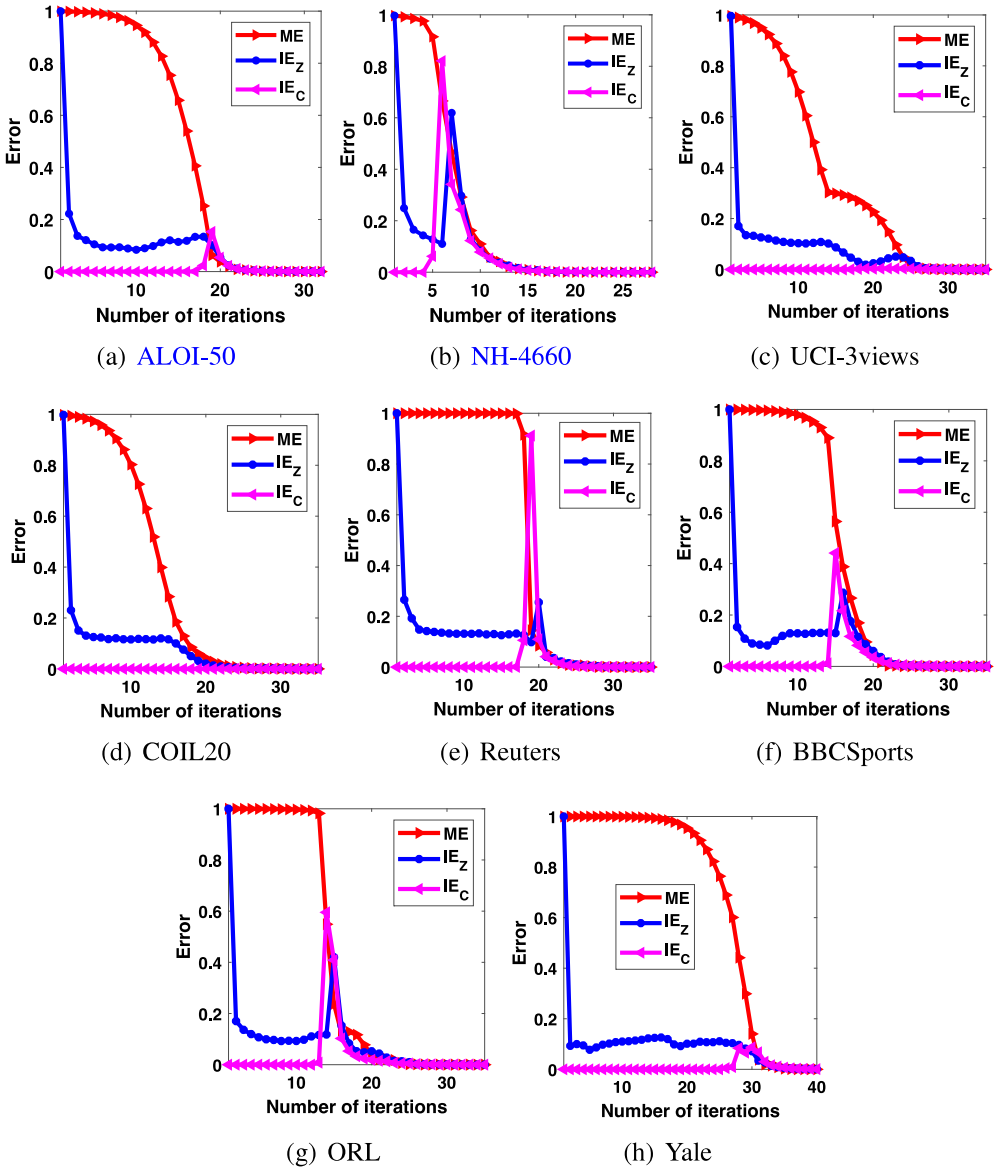


Fig. 9. Convergence analysis on eight benchmark datasets when missing rate is 0.1.

Table 5
The average running time over different missing rates by twelve IMC approaches. On each dataset, we run every IMC approach 20 times with missing rate varying from 0.1 to 0.5.

Datasets	MKKM-ZF	MKKM-MF	BSV	Concat	UEAF	IMC-AGL	UIMC	IK-MKKM	EE-IMVC	MERA-ZF	S^2MVTC-ZF	TIMKSC
ALOI-50	97.36	89.47	1.27	1.52	>24 h	>24 h	321.32	1592.63	29.42	>24 h	0.72	1172.46
NH-4660	16.31	16.81	6.75	6.87	>24 h	>24 h	265.60	730.67	6.34	>24 h	0.08	462.35
UCI-3views	1.08	1.14	2.05	0.95	0.82	16.63	571.31	144.38	19.76	2536.70	1.17	54.67
COIL20	0.63	0.68	1.10	1.12	1.32	42.48	112.2	194.32	9.25	573.27	1.09	28.03
Reuters	0.39	0.46	0.93	0.51	0.62	9.96	24.96	106.64	2.75	570.39	0.56	30.59
BBCSports	0.07	0.08	0.09	1.24	1.26	6.54	28.23	34.74	0.34	97.18	NA	2.72
ORL	0.12	0.13	0.16	0.37	0.46	24.71	24.46	134.49	1.01	50.41	NA	2.32
Yale	0.03	0.03	0.04	0.13	0.15	15.80	6.16	124.51	0.17	9.15	NA	0.85

Note that ">24 h" means the baseline exceeds more than 24 hours, and "NA" means the baseline is not applicable to this dataset.

Table 6
Time complexity of the twelve IMC approaches.

	Method	Overall time complexity
Two-stage based approaches	MKKM-ZF	$O(Vn_m + Tnk)$
	MKKM-MF	$O(Vn_m + Tnk)$
	BSV	$O(Vn_m + TVnk)$
	Concat	$O(Vn_m + Tnk)$
	MERA-ZF	$O(Vn_m + TVn^3)$
	S^2MVTC-ZF	$O(Vn_m + \max(TVp^3, TVnpr))$
One-stage based approaches	UEAF	$O(T(Vn^3 + \sum_{v=1}^{V} d_v k^2))$
	IMC-AGL	$O(TVn^3)$
	UIMC	$O(TVn^3)$
	IK-MKKM	$O(Tnk + \sum_{v=1}^{V}\{(n_m)^3 + n_m(n_c)^2\})$
	EE-IMVC	$O(Tnk^2 + TV(k^3 + n_m)k^2)$
	TIMKSC	$O(TVn^3)$

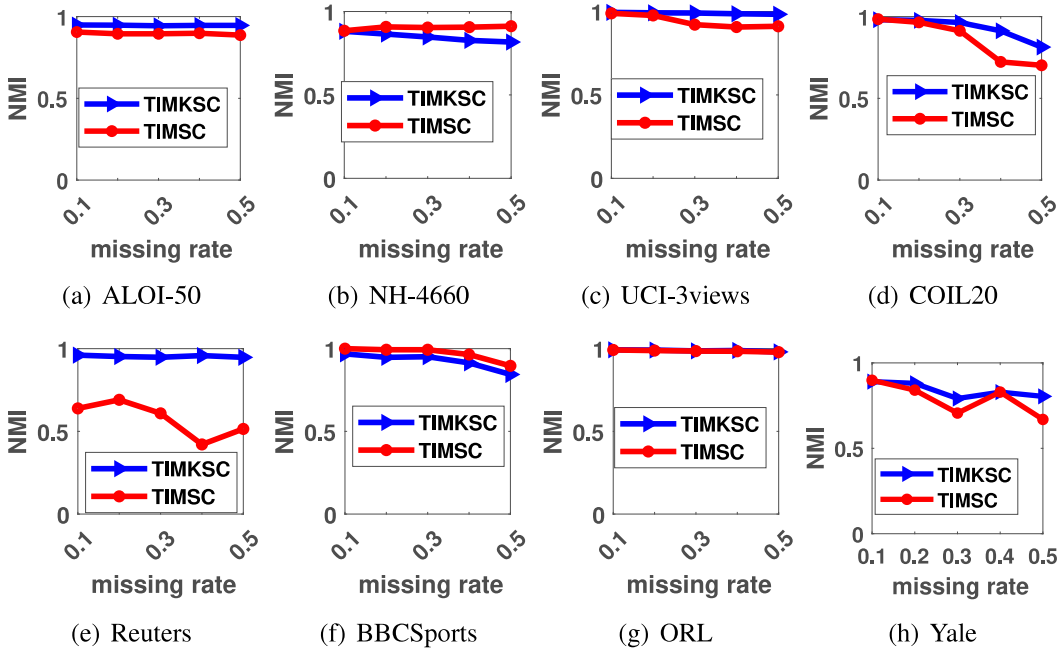


Fig. 10. Ablation study on three benchmark datasets w.r.t. NMI metric.

designed to discover the nonlinear subspace structure embedded in incomplete multi-view data. Notably, the proposed approach is capable of imputing multi-view kernel matrices and learning the low-rank tensor representations based on a unified framework. The benefit of this general framework involves generating the robust imputed kernels, as well as capturing the high-order correlations across different views for effective subspace clustering. Moreover, the optimization model of our approach only involves one hyper-parameter that can be efficiently solved in an efficient manner. Finally, empirical results on eight benchmark datasets clearly demonstrate the significant superiority of our approach beyond the state-of-art baselines. In terms of the future work, we expect to enhance the robustness of our TIMKSC approach via effective deep learning framework (He, Sun, Yan, Li, & Fu, 2022; Liu, Li, & Lin, 2023; Wang, Wang, et al., 2023).

CRediT authorship contribution statement

Guang-Yu Zhang: Writing – original draft, Methodology, Formal analysis, Data curation, Conceptualization. **Dong Huang:** Writing – review & editing, Resources, Project administration, Methodology, Funding acquisition. **Chang-Dong Wang:** Writing – review & editing, Validation, Funding acquisition.

Declaration of competing interest

The manuscript, or part of it, has neither been published nor is currently under consideration for publication by any other journal. All authors have read and approved the content. No disclosure, ethical or legal conflicts and no competing interests exist in this study. For the manuscript, there is not any potential financial or substantive conflict. We hope that it is in accordance with your publishing regulations and under consideration for publication once got approved.

Data availability

No data was used for the research described in the article.

Acknowledgments

This work was supported by the NSFC (62206099, 61976097 and 61876193) and Science and Technology Program of Guangzhou (2024A04J4451).

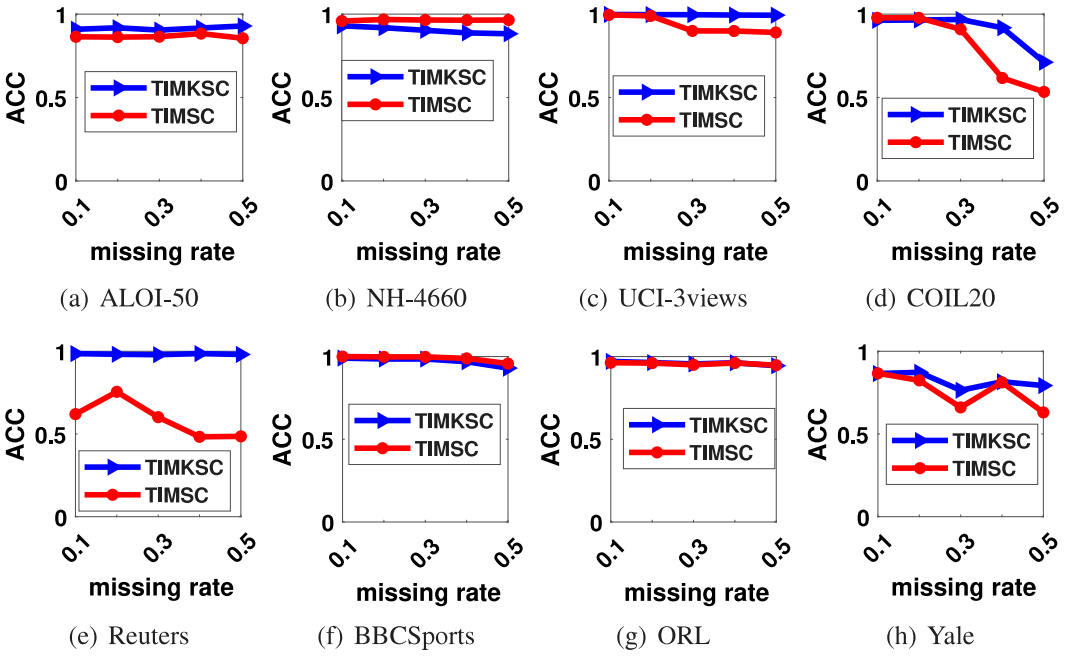


Fig. 11. Ablation study on three benchmark datasets w.r.t. ACC metric.

References

Abavisani, M., & Patel, V. M. (2018). Multimodal sparse and low-rank subspace clustering. *Information Fusion*, *39*, 168–177.

Cai, X., Huang, D., Zhang, G.-Y., & Wang, C.-D. (2023). Seeking commonness and inconsistencies: A jointly smoothed approach to multi-view subspace clustering. *Information Fusion*, *91*, 364–375.

Chen, M. S., Huang, L., Wang, C. D., Huang, D., & Lai, J. H. (2021). Relaxed multi-view clustering in latent embedding space. *Information Fusion*, *68*, 8–21.

Chen, Y., Wang, S., Xiao, X., Liu, Y., Hua, Z., & Zhou, Y. (2021). Self-paced enhanced low-rank tensor kernelized multi-view subspace clustering. *IEEE Transactions on Multimedia*, *24*, 4054–4066.

Deng, S., Wen, J., Liu, C., Yan, K., Xu, G., & Xu, Y. (2023). Projective incomplete multi-view clustering. *IEEE Transactions on Neural Networks and Learning Systems*.

Fang, X., Hu, Y., Zhou, P., & Wu, D. O. (2021). Unbalanced incomplete multi-view clustering via the scheme of view evolution: Weak views are meat; strong views do eat. *IEEE Transactions on Emerging Topics in Computational Intelligence*.

Fang, S. G., Huang, D., Cai, X. S., Wang, C. D., He, C., & Tang, Y. (2023). Efficient multi-view clustering via unified and discrete bipartite graph learning. *IEEE Transactions on Neural Networks and Learning Systems*.

Gönen, M., & Alpaydın, E. (2011). Multiple kernel learning algorithms. *Journal of Machine Learning Research*, *12*, 2211–2268.

He, Q., Sun, X., Yan, Z., Li, B., & Fu, K. (2022). Multi-object tracking in satellite videos with graph-based multitask modeling. *IEEE Transactions on Geoscience and Remote Sensing*, *60*, 1–13.

Hu, M., & Chen, S. (2019a). Doubly aligned incomplete multi-view clustering. In *Proceedings of the international joint conference on artificial intelligence* (pp. 2262–2268).

Hu, M., & Chen, S. (2019b). One-pass incomplete multi-view clustering. In *Proceedings of the AAAI conference on artificial intelligence, vol. 33, no. 01* (pp. 3838–3845).

Hu, W., Tao, D., Zhang, W., Xie, Y., & Yang, Y. (2016). The twist tensor nuclear norm for video completion. *IEEE transactions on neural networks and learning systems*, *28*(12), 2961–2973.

Huang, L., Chao, H. Y., & Wang, C. D. (2019). Multi-view intact space clustering. *Pattern Recognition*, *86*, 344–353.

Huang, S., Kang, Z., Tsang, I. W., & Xu, Z. (2019). Auto-weighted multi-view clustering via kernelized graph learning. *Pattern Recognition*, *88*, 174–184.

Jin, J., Wang, S., Dong, Z., Liu, X., & Zhu, E. (2023). Deep incomplete multi-view clustering with cross-view partial sample and prototype alignment. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition* (pp. 11600–11609).

Kang, Z., Lin, Z., Zhu, X., & Xu, W. (2021). Structured graph learning for scalable subspace clustering: From single view to multiview. *IEEE Transactions on Cybernetics*, *52*(9), 8976–8986.

Kilmer, M. E., & Martin, C. D. (2011). Factorization strategies for third-order tensors. *Linear Algebra and its Applications*, *435*(3), 641–658.

Li, S. Y., Jiang, Y., & Zhou, Z. H. (2014). Partial multi-view clustering. In *Proceedings of the AAAI conference on artificial intelligence, vol. 28, no. 1*.

Li, L., Wan, Z., & He, H. (2021). Incomplete multi-view clustering with joint partition and graph learning. *IEEE Transactions on Knowledge and Data Engineering*.

Li, M., Xia, J., Xu, H., Liao, Q., Zhu, X., & Liu, X. (2021). Localized incomplete multiple kernel k-means with matrix-induced regularization. *IEEE Transactions on Cybernetics*.

Liang, Y. W., Huang, D., Wang, C. D., & Yu, P. S. (2022). Multi-view graph learning by joint modeling of consistency and inconsistency. *IEEE Transactions on Neural Networks and Learning Systems*.

Liang, J., Liu, X., Bai, L., Cao, F., & Wang, D. (2022). Incomplete multi-view clustering via local and global co-regularization. *Science China. Information Sciences*, *65*(5), Article 152105.

Liang, D. H., & Wang, C. D. (2019). Consistency meets inconsistency: A unified graph learning framework for multi-view clustering. In *Proceedings of the 19th IEEE international conference on data mining* (pp. 1204–1209).

Lin, Y., Gou, Y., Liu, Z., Li, B., Lv, J., & Peng, X. (2021). Completer: Incomplete multi-view clustering via contrastive prediction. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition* (pp. 11174–11183).

Liu, X. (2021). Incomplete multiple kernel alignment maximization for clustering. *IEEE Transactions on Pattern Analysis and Machine Intelligence*.

Liu, X. (2022). Simplemkkm: Simple multiple kernel k-means. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *45*(4), 5174–5186.

Liu, X. (2023). Hyperparameter-free localized simple multiple kernel K-means with global optimum. *IEEE Transactions on Pattern Analysis and Machine Intelligence*.

Liu, Y., Li, G., & Lin, L. (2023). Cross-modal causal relational reasoning for event-level visual question answering. *IEEE Transactions on Pattern Analysis and Machine Intelligence*.

Liu, X., Li, M., Tang, C., Xia, J., Xiong, J., Liu, L., et al. (2020). Efficient and effective regularized incomplete multi-view clustering. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *43*(8), 2634–2646.

Liu, C., Li, R., Wu, S., Che, H., Jiang, D., Yu, Z., et al. (2023). Self-guided partial graph propagation for incomplete multiview clustering. *IEEE Transactions on Neural Networks and Learning Systems*.

Liu, J., Liu, X., Zhang, Y., Zhang, P., Tu, W., Wang, S., et al. (2021). Self-representation subspace clustering for incomplete multi-view data. In *Proceedings of the 29th ACM international conference on multimedia* (pp. 2726–2734).

Liu, J., Teng, S., Fei, L., Zhang, W., Fang, X., Zhang, Z., et al. (2021). A novel consensus learning approach to incomplete multi-view clustering. *Pattern Recognition*, *115*, Article 107890.

Liu, C., Wen, J., Wu, Z., Luo, X., Huang, C., & Xu, Y. (2023). Information recovery-driven deep incomplete multiview clustering network. *IEEE Transactions on Neural Networks and Learning Systems*.

Liu, X., Zhu, X., Li, M., Tang, C., Zhu, E., Yin, J., et al. (2019). Efficient and effective incomplete multi-view clustering. In *Proceedings of the AAAI conference on artificial intelligence, vol. 33, no. 01* (pp. 4392–4399).

Liu, X., Zhu, X., Li, M., Wang, L., Zhu, E., Liu, T., et al. (2019). Multiple kernel k k-means with incomplete kernels. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *42*(5), 1191–1204.

Long, Z., Wang, Q., Ren, Y., Liu, Y., & Zhu, C. (2024). S2MVTC: a simple yet efficient scalable multi-view tensor clustering. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition* (pp. 26213–26222).

Long, Z., Zhu, C., Chen, J., Li, Z., Ren, Y., & Liu, Y. (2023). Multi-view MERA subspace clustering. *IEEE Transactions on Multimedia*.

Long, Z., Zhu, C., Comon, P., Ren, Y., & Liu, Y. (2023). Feature space recovery for efficient incomplete multi-view clustering. *IEEE Transactions on Knowledge and Data Engineering*.

Lu, Y., Liu, Y., Long, Z., Chen, Z., & Zhu, C. (2023). O-minus decomposition for multi-view tensor subspace clustering. *IEEE Transactions on Artificial Intelligence*.

Maaten, L. v. d., & Hinton, G. (2008). Visualizing data using t-SNE. *Journal of Machine Learning Research*, *9*(Nov), 2579–2605.

Shao, W., He, L., Lu, C. t., & Philip, S. Y. (2016). Online multi-view clustering with incomplete views. In *2016 IEEE international conference on big data* (pp. 1012–1017). IEEE.

Shao, W., He, L., & Yu, P. S. (2015). Multiple incomplete views clustering via weighted nonnegative matrix factorization with L2, 1 regularization. In *Joint European conference on machine learning and knowledge discovery in databases* (pp. 318–334). Springer.

Tang, C., Li, Z., Wang, J., Liu, X., Zhang, W., & Zhu, E. (2022). Unified one-step multi-view spectral clustering. *IEEE Transactions on Knowledge and Data Engineering*, *35*(6), 6449–6460.

Wang, Y., Chang, D., Fu, Z., Wen, J., & Zhao, Y. (2022a). Graph contrastive partial multi-view clustering. *IEEE Transactions on Multimedia*.

Wang, Y., Chang, D., Fu, Z., Wen, J., & Zhao, Y. (2022b). Incomplete multiview clustering via cross-view relation transfer. *IEEE Transactions on Circuits and Systems for Video Technology*, *33*(1), 367–378.

Wang, S., Liu, X., Zhu, X., Zhang, P., Zhang, Y., Gao, F., et al. (2021). Fast parameter-free multi-view subspace clustering with consensus anchor guidance. *IEEE Transactions on Image Processing*, *31*, 556–568.

Wang, J., Tang, C., Wan, Z., Zhang, W., Sun, K., & Zomaya, A. Y. (2023). Efficient and effective one-step multiview clustering. *IEEE Transactions on Neural Networks and Learning Systems*.

Wang, J., Tang, C., Zheng, X., Liu, X., Zhang, W., Zhu, E., et al. (2023). Fast approximated multiple kernel K-means. *IEEE Transactions on Knowledge and Data Engineering*.

Wang, B., Wang, Z., Sun, X., He, Q., Wang, H., & Fu, K. (2023). TDNet: A novel transductive learning framework with conditional metric embedding for few-shot remote sensing image scene classification. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*.

Wang, H., Zong, L., Liu, B., Yang, Y., & Zhou, W. (2019). Spectral perturbation meets incomplete multi-view data. In *Proceedings of the 28th international joint conference on artificial intelligence* (pp. 3677–3683).

Wen, J., Xu, Y., & Liu, H. (2018). Incomplete multiview spectral clustering with adaptive graph learning. *IEEE Transactions on Cybernetics*, *50*(4), 1418–1429.

Wen, J., Yan, K., Zhang, Z., Xu, Y., Wang, J., Fei, L., et al. (2020). Adaptive graph completion based incomplete multi-view clustering. *IEEE Transactions on Multimedia*, *23*, 2493–2504.

Wen, J., Zhang, Z., Fei, L., Zhang, B., Xu, Y., Zhang, Z., et al. (2022). A survey on incomplete multiview clustering. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*.

Wen, J., Zhang, Z., Xu, Y., Zhang, B., Fei, L., & Liu, H. (2019). Unified embedding alignment with missing views inferring for incomplete multi-view clustering. In *Proceedings of the AAAI conference on artificial intelligence, vol. 33, no. 01* (pp. 5393–5400).

Wen, J., Zhang, Z., Zhang, Z., Fei, L., & Wang, M. (2020). Generalized incomplete multiview clustering with flexible locality structure diffusion. *IEEE Transactions on Cybernetics*, *51*(1), 101–114.

Wu, J., Lin, Z., & Zha, H. (2019). Essential tensor learning for multi-view spectral clustering. *IEEE Transactions on Image Processing*, *28*(12), 5910–5922.

Xia, D., Yang, Y., Yang, S., & Li, T. (2023). Incomplete multi-view clustering via kernelized graph learning. *Information Sciences*, *625*, 1–19.

Xie, Y., Tao, D., Zhang, W., Liu, Y., Zhang, L., & Qu, Y. (2018). On unifying multi-view self-representations for clustering by tensor multi-rank minimization. *International Journal of Computer Vision*, *126*(11), 1157–1179.

Xie, Y., Zhang, W., Qu, Y., Dai, L., & Tao, D. (2018). Hyper-Laplacian regularized multilinear multiview self-representations for clustering and semisupervised learning. *IEEE Transactions on Cybernetics*, *50*(2), 572–586.

Xu, J., Li, C., Peng, L., Ren, Y., Shi, X., Shen, H. T., et al. (2023). Adaptive feature projection with distribution alignment for deep incomplete multi-view clustering. *IEEE Transactions on Image Processing*, *32*, 1354–1366.

Xu, J., Ren, Y., Tang, H., Yang, Z., Pan, L., Yang, Y., et al. (2022). Self-supervised discriminative feature learning for deep multi-view clustering. *IEEE Transactions on Knowledge and Data Engineering*, *35*(7), 7470–7482.

Yang, X., Deng, C., Dang, Z., & Tao, D. (2021). Deep multiview collaborative clustering. *IEEE Transactions on Neural Networks and Learning Systems*, *34*(1), 516–526.

Yang, M., Li, Y., Hu, P., Bai, J., Lv, J., & Peng, X. (2022). Robust multi-view clustering with incomplete information. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *45*(1), 1055–1069.

Yu, X., Liu, H., Lin, Y., Wu, Y., & Zhang, C. (2022). Auto-weighted sample-level fusion with anchors for incomplete multi-view clustering. *Pattern Recognition*, *130*, Article 108772.

Zhang, C., Fu, H., Hu, Q., Cao, X., Xie, Y., Tao, D., et al. (2018). Generalized latent multi-view subspace clustering. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *42*(1), 86–99.

Zhang, C., Hu, Q., Fu, H., Zhu, P., & Cao, X. (2017). Latent multi-view subspace clustering. In *Proceedings of the 30th IEEE conference on computer vision and pattern recognition* (pp. 4279–4287).

Zhang, G. Y., Wang, C. D., Huang, D., Zheng, W. S., & Zhou, Y. R. (2018). TW-Co-k-means: Two-level weighted collaborative k-means for multi-view clustering. *Knowledge-Based Systems*, *150*, 127–138.

Zhang, G. Y., Zhou, Y. R., He, X. Y., Wang, C. D., & Huang, D. (2019). One-step kernel multi-view subspace clustering. *Knowledge-Based Systems*, Article 105126.

Zhang, G. Y., Zhou, Y. R., Wang, C. D., Huang, D., & He, X. Y. (2021). Joint representation learning for multi-view subspace clustering. *Expert Systems with Applications*, *166*, Article 113913.

Zhao, H., Liu, H., & Fu, Y. (2016). Incomplete multi-modal visual data grouping.. In *Proceedings of the international joint conference on artificial intelligence* (pp. 2392–2398).

Zhu, X., Liu, X., Li, M., Zhu, E., Liu, L., Cai, Z., et al. (2018). Localized incomplete multiple kernel k-means. In *International joint conferences on artificial intelligence organization, Proceedings of the twenty-seventh international joint conference on artificial intelligence*.

Zhuge, W., Tao, H., Luo, T., Zeng, L.-L., Hou, C., & Yi, D. (2020). Joint representation learning and clustering: A framework for grouping partial multiview data. *IEEE Transactions on Knowledge and Data Engineering*, *34*(8), 3826–3840.

RESEARCH-ARTICLE

Large-Scale Tensorized Multi-View Kernel Subspace Clustering

GUANGYU ZHANG, South China Agricultural University, Guangzhou, Guangdong, China

DONG HUANG, South China Agricultural University, Guangzhou, Guangdong, China

CHANGDONG WANG, Sun Yat-Sen University, Guangzhou, Guangdong, China

Published: 23 July 2025
Online AM: 16 May 2025
Accepted: 16 April 2025
Revised: 03 February 2025
Received: 20 March 2023

Large-Scale Tensorized Multi-View Kernel Subspace Clustering

GUANG-YU ZHANG and DONG HUANG, College of Mathematics and Informatics, South China Agricultural University, Guangzhou, China
CHANG-DONG WANG, School of Computer Science and Engineering, Sun Yat-sen University, Guangzhou, China, Guangxi Key Laboratory of Digital Infrastructure, Guangxi Zhuang Autonomous Region Information Center, Nanning, China, Guangdong Key Laboratory of Big Data Analysis and Processing, Guangzhou, China, and Key Laboratory of Machine Intelligence and Advanced Computing, Ministry of Education, Guangzhou, China

The anchor-based multi-view subspace clustering (AMSC) has turned into a favorable tool for large-scale multi-view clustering. However, there still exist some limitations to the current AMSC approaches. First, they typically recover anchor graph structure in the original linear space, restricting their feasibility for nonlinear scenarios. Second, they usually overlook the potential benefits of jointly capturing the inter-view and intra-view information for enhancing the anchor representation learning. Third, these approaches mostly perform anchor-based subspace learning by a specific matrix norm, neglecting the latent high-order correlation across different views. To overcome these limitations, this article presents an efficient and effective approach termed Large-Scale Tensorized Multi-View Kernel Subspace Clustering (LTKMSC). Different from the existing AMSC approaches, our LTKMSC approach exploits both inter-view and intra-view awareness for anchor-based representation building. Concretely, the low-rank tensor learning is leveraged to capture the high-order correlation (i.e., the inter-view complementary information) among distinct views, upon which the $l_{1,2}$ norm is imposed to explore the intra-view anchor graph structure in each view. Moreover, the kernel learning technique is leveraged to explore the nonlinear anchor–sample relationships embedded in multiple views. With the unified objective function formulated, an efficient optimization algorithm that enjoys low computational complexity is further designed. Extensive experiments on a variety of multi-view datasets have confirmed the efficiency and effectiveness of our approach when compared with the other competitive approaches.

CCS Concepts: • **Computer systems organization** → **Embedded systems**; *Robotics*; • **Hardware** → *Redundancy*; • **Networks** → Network reliability;

This work was supported by the NSFC (62206099) and Science and Technology Program of Guangzhou (2024A04J4451), in part by the Open Project Program of Guangxi Key Laboratory of Digital Infrastructure (GXDIOP2024011) and Guangdong Basic and Applied Basic Research Foundation (2022B1515120059), and in part by the NSFC (62476102) and Natural Science Foundation of Guangdong Province (2025A1515010245).
Authors' Contact Information: Guang-Yu Zhang, College of Mathematics and Informatics, South China Agricultural University, Guangzhou, China; e-mail: guangyuzhg@foxmail.com; Dong Huang (corresponding author), College of Mathematics and Informatics, South China Agricultural University, Guangzhou, China; e-mail: huangdonghere@gmail.com; Chang-Dong Wang, School of Computer Science and Engineering, Sun Yat-sen University, Guangzhou, China, Guangxi Key Laboratory of Digital Infrastructure, Guangxi Zhuang Autonomous Region Information Center, Nanning, China, Guangdong Key Laboratory of Big Data Analysis and Processing, Guangzhou, China, and Key Laboratory of Machine Intelligence and Advanced Computing, Ministry of Education, Guangzhou, China; e-mail: changdongwang@hotmail.com.

ACM 2157-6912/2025/7-ART85
https://doi.org/10.1145/3735644

Additional Key Words and Phrases: Data clustering, Multi-view clustering, Large-scale clustering, Tensorized kernel subspace clustering, Inter-view and intra-view awareness

ACM Reference format:
Guang-Yu Zhang, Dong Huang, and Chang-Dong Wang. 2025. Large-Scale Tensorized Multi-View Kernel Subspace Clustering. *ACM Trans. Intell. Syst. Technol.* 16, 4, Article 85 (July 2025), 21 pages.
https://doi.org/10.1145/3735644

1 Introduction

The rapid development of multimedia technology gives rise to the enormous emergence of multi-view data in scientific applications. Practically, multi-view data could be obtained from diverse data sources (or extracted by multiple feature collectors), whereas every source (or views) consists of the particular statistic property of the same data object. As an instance, the same news about a certain event is reported in various languages, like Chinese, English, and German. The same scene image could be depicted by multiple visual features, such LBP, GIST, and HOG. A video surveillance system usually monitors the same scene by a series of cameras that are positioned in different locations. As the multi-view data conveys diverse types of features, how to efficiently and effectively employ the hidden knowledge from heterogeneous views has become an open yet critical problem. To address this problem, considerable attention has been paid to the multi-view clustering research, which aims to boost the performance by fusing the common and heterogeneous knowledge of multiple views.

In the past decade, there has been an explosion toward the study of multi-view clustering. In terms of the heterogeneous perspective, the present multi-view clustering approaches may roughly be split into five categories, namely, the k-means based approaches [13, 50, 52], the kernel-based approaches, the matrix factorization-based approaches, the graph-based approaches, and the subspace-based approaches. Specifically, benefiting from the theoretical guarantees and some scalable optimization algorithms, the graph-based approaches have aroused widespread research interest in multi-view clustering. By deriving from individual view to multiple views, the objective of graph-based approaches is to discover the latent similarity structure (or spectral embedding) for robust graph partitioning [24, 35, 46]. Beyond the aforementioned approaches, the subspace-based approaches have been demonstrated as a promising technique for the clustering analysis of high-dimensional data. A variety of subspace-based approaches have been developed [18, 27, 47], among which the key is to pursue a robust subspace representation contained in heterogeneous views, in order to build a meaningful similarity matrix for the post-processing algorithm.

Though the existing subspace-based approaches have made considerable progress, yet majority of them still encounter the expensive time overhead, which significantly restricts their applications in large-scale (or even some general-scale) scenarios. To breakthrough the computational bottleneck, the anchor-based multi-view subspace clustering (abbreviated as the anchor subspace-based approaches in the following) has been a popular topic in the present literature. Instead of learning an $n \times n$ subspace representation for each view, the anchor subspace-based approaches generate a set of anchor representations (with size of $p \times n$) to efficiently model the anchor–sample relationship from multiple views, where n is the number of data samples, and p is the number of anchors (also known as landmarks or representatives). During the last few years, massive efforts have been made on the growth of the anchor subspace-based approaches [16, 17, 34]. In contrast to the traditional subspace-based approaches, the current anchor subspace-based approaches have reduced the computational cost from $\mathrm{O}(n^3)$ to $\mathrm{O}(n)$ (when the condition $p \ll n$ holds). However, there are still three crucial questions that remain to be settled. First of all, how to enable the anchor

subspace clustering for nonlinear large-scale datasets? Second, how to jointly capture both the inter-view and intra-view awareness for robust multi-view clustering? Last but not least, how to efficiently exploit the higher-order correlation contained in multiple views, so as to model the latent anchor-sample relationship from a deeper perspective?

In light of these crucial questions, we present a novel multi-view clustering approach named **Large-Scale Tensorized Multi-View Kernel Subspace Clustering (LTKMSC)**, which enjoys both low complexity and highly competitive clustering performance. Going beyond the traditional practice, our approach harnesses the power of inter-view and intra-view awareness for robust anchor representation learning. Particularly, the t-SVD-based tensor decomposition is leveraged to capture the high-order correlation among different views, while an $l_{1,2}$-norm regularizer is introduced to model the discriminative structure hidden in each view. Further, the kernel learning is incorporated into the proposed approach, which improves the robustness of the anchor subspace learning when handling nonlinearly separable datasets. Experiments are carried on 11 multi-view datasets. The experimental results have confirmed the advantages of our proposed approach in both effectiveness and efficiency.

As a summary, we conclude the main contributions of this work as follows:

- To the best of our knowledge, this article for the first time bridges the gap between the anchor multi-view subspace clustering and the kernelized tensor subspace clustering.
- A unified objective function with both inter-view and intra-view awareness is formulated, upon which a highly efficient optimization algorithm is designed.
- A new anchor subspace-based approach termed LTKMSC is proposed. Extensive experiments demonstrate its superiority in clustering quality and scalability over the state-of-the-art.

The rest of this article is organized in what follows. Section 2 gives a brief review of the related works. Section 3 is dedicated to describe the proposed LTKMSC approach. Section 4 reports the experimental results on several multi-view datasets. Section 5 provides the conclusion of this article.

2 Related Work

This section briefly reviews the representative works on **multi-view subspace clustering (MSC)**, tensorized multi-view clustering, and large-scale multi-view clustering, which are highly relevant to our proposed LTKMSC approach.

MSC takes an essential part in a variety of scientific researches. Naturally, MSC researches were extended from single-view subspace learning. For example, in [9], Gao et al. proposed a pioneering work on MSC, which extends the classical subspace learning model to multi-view domain by enhancing the consistency embedded in diverse views. Following this work, Cao et al. [3] presented a diversity-induced subspace approach for multi-view clustering, whose key idea is to explore the complementary knowledge across multi-view subspace representations via the HSIC Criterion [11]. Additionally, Zhang et al. [48] proposed another subspace-based approach for multi-view clustering that learns the shared latent representation contained in various views. Along this line, Wang et al. [39] further developed a unified MSC approach, where the complementarity and consistency across multi-view representations are modeled in a general framework. In [14], Huang et al. derived a kernel subspace clustering approach which performs multi-view kernel clustering and graph learning mutually. Besides, Chen et al. [4] proposed a multi-view clustering in latent embedding space approach, which is capable of learning a latent space embedded in multiple views while discovering the global cluster structure simultaneously. Furthermore, Li et al. [19] developed a unified deep MSC approach, which takes advantages of two mutually beneficial modules, namely, the hierarchical self-representative layers and a series of multi-view backward encoding networks. Cai et al. [2] presented a smoothed subspace-based approach that aims to seek the commonness

and inconsistency across multiple views. Wang et al. [36] designed a structural multi-pathway network for deep subspace representation learning, where both low-level and high-level features are explicitly utilized to explore the latent structure across diverse views.

To explore the high-order information, the tensor decomposition technique has recently exhibited its promising capacity in community detection [15, 32, 42] and multi-view clustering [7, 40, 41]. For instance, Gorovits et al. [10] derived an overlapping community detection approach from a high-order perspective, which seeks to discover the community structure while exploring the community activity profile simultaneously. Besides, Al-Sayouri et al. [1] developed a simple yet effective approach for multi-view representation learning, which employs the canonical polyadic decomposition to capture the high-order correlation between two mutually connected views, namely the conventional adjacency view and the corresponding side information view. Moving forward, Sheikholeslami and Giannakis [33] proposed a top-to-bottom community detection approach with high-order exploitation, among which the egonet-tensor decomposition technique is utilized to characterize the underlying structural information hidden in overlapping communities. Furthermore, several attempts have been made to incorporate the multi-view clustering (especially MSC) with the tensor decomposition technique. For this purpose, Xie et al. [43] introduced a flexible constraint on the rotated tensor, and simultaneously designed a robust subspace-based approach termed t-SVD-MSC. By preserving the specific locality in each view, Xie et al. [44] introduced the hyper-Laplacian regularizer into the model of t-SVD-MSC and designed two novel approaches for clustering and semi-supervised learning tasks, namely, HLR-M^{2}VS and semi-HLR-M^{2}VS, respectively. By means of the tensor graph learning for different views, Chen et al. [6] designed a tensor MSC approach termed LRRG. Li et al. [22] further derived a unified multi-view graph clustering approach, which formulates the spectral embedding and low-rank tensor decomposition into a one-step learning scheme. Moving forward, Long et al. [26] developed a MERA-oriented tensorized MSC approach, where the multi-view tensor representation is factorized into one core factor and a series of orthogonal factors so as to reveal the high-order correlation among the multiple self-presentations. Although these conventional approaches have achieved remarkable progress in the literature, most of them still suffer from expensive computational burden, which limits their capability in solving large-scale (or even some general-scale) clustering problems.

To breakthrough the efficiency bottleneck, several anchor-based multi-view clustering approaches have been developed via different strategies. For instance, Li et al. [20] proposed a parameter-free and scalable approach that can efficiently fuse the complementary among multiple bipartite graphs. Yang et al. [45] utilized an embedded anchor graph to reduce the computational complexity and enhance the clustering robustness. Based on the bipartite graph learning technique, Kang et al. [17] presented a linear-time large-scale approach for MSC termed LMVSC. Sun et al. [34] made an early attempt to incorporate the unified anchor learning with scalable multi-view subspace learning. Wang et al. [38] proposed a fast approach for scalable MSC, where the latent anchor graph is optimized without any hyper-parameter tuning. Liu et al. [25] developed a fast subspace-based approach that builds the common anchor graph from a view-independent perspective. More recently, Fang et al. [8] devised an efficient multi-view subspace approach with unified and discrete bipartite graph learning, which models the bipartite graph construction and the discrete cluster structure learning into a joint objective function. Wang et al. [37] proposed a flexible and scalable subspace-based approach with adaptive anchor fusion strategy, where both the anchor graph construction and the graph alignment are jointly considered for boosting the clustering quality. Despite the remarkable progress, we observe that their performances can be further improved by three aspects of considerations, namely, (i) how to preserve the nonlinear subspace structure in the kernel space, (ii) how to enhance the high-order complementarity among heterogeneous views, and

(iii) how to harness both inter-view and intra-view awareness to enhance the large-scale subspace clustering quality.

3 Proposed LTKMSC Approach

This section describes our proposed LTKMSC approach in detail. Specifically, the primary notations throughout this article is provided in Section 3.1. Then the objective formulation of the proposed approach is introduced in Section 3.2. After that, an efficient optimization algorithm is presented in Section 3.3. Finally, the time complexity and space complexity are analyzed in Section 3.4.

3.1 Main Notations

For convenience, this article utilizes the calligraphy letters for 3-mode tensors (e.g., $\mathcal{P}$) and uppercase letters for matrices (e.g., P). Particularly, the ith row, jth column, and ijth element of the matrix P are deployed by $\mathrm{P}_{i:}$, $\mathrm{P}_{:j}$, and P_{ij}, respectively. In addition, the ith horizontal slice, jth lateral slice, and kth frontal slice of the 3-mode tensor $\mathcal{P} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$ are deployed by $\mathcal{P}(i,:,:)$, $\mathcal{P}(:,j,:)$, and $\mathcal{P}(:,:,k)$, respectively. The operator $\overline{\mathcal{P}} = \mathrm{fft}(\mathcal{P}, [], 3)$ stands for the DFT transform towards the third dimension of tensor $\mathcal{P}$, while the inverse DFT of tensor $\mathcal{P}$ can be denoted by $\mathcal{P} = \mathrm{ifft}(\overline{\mathcal{P}}, [], 3)$. For convenience, the matrix $\mathrm{P}^{(k)}$ is used to denote the frontal slice $\mathcal{P}(:,:,k)$.

Based on the above notations, five block-based operators and the basic definitions that are relevant to t-SVD are given in detail. First of all, we can have the following block circulant matrix for any given tensor $\mathcal{P} \in \mathcal{R}^{n_1 n_3 \times n_2 n_3}$:

$$\mathrm{bcirc}(\mathcal{P}) = \begin{bmatrix} \mathrm{P}^{(1)} & \mathrm{P}^{(n_3)} & \cdots & \mathrm{P}^{(2)} \\ \mathrm{P}^{(2)} & \mathrm{P}^{(1)} & \cdots & \mathrm{P}^{(3)} \\ \vdots & \ddots & \ddots & \vdots \\ \mathrm{P}^{(n_3)} & \mathrm{P}^{(n_3-1)} & \cdots & \mathrm{P}^{(1)} \end{bmatrix}.$$

The *unfold* and *fold* operations of tensor $\mathcal{P}$ could be given in what follows:

$$\mathrm{unfold}(\mathcal{P}) = \begin{bmatrix} \mathrm{P}^{(1)} \\ \mathrm{P}^{(2)} \\ \vdots \\ \mathrm{P}^{(n_3)} \end{bmatrix}, \quad \mathrm{fold}(\mathrm{unfold}(\mathcal{P})) = \mathcal{P}.$$

Additionally, the *bdiag* and *bdfold* operations of tensor $\mathcal{P}$ are given in the following:

$$\mathrm{bdiag}(\mathcal{P}) = \begin{bmatrix} \mathrm{P}^{(1)} & & & \\ & \mathrm{P}^{(2)} & & \\ & & \ddots & \\ & & & \mathrm{P}^{(n_3)} \end{bmatrix}, \tag{1}$$

$$\mathrm{bdfold}(\mathrm{bdaig}(\mathcal{P})) = \mathcal{P}. \tag{2}$$

Definition 1 (t-Product). Given any two tensors $\mathcal{P} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$ and $\mathcal{Q} \in \mathbb{R}^{n_2 \times n_4 \times n_3}$, the corresponding t-product $\mathcal{C} = \mathcal{P} * \mathcal{Q}$ is a tensor of size $n_1 \times n_4 \times n_3$:

$$\mathcal{C} = \mathcal{P} * \mathcal{Q} = \mathrm{fold}(\mathrm{bcirc}(\mathcal{P}) \cdot \mathrm{unfold}(\mathcal{Q})). \tag{3}$$

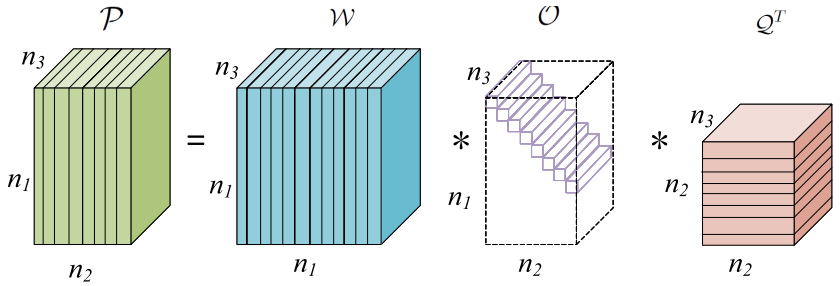


Fig. 1. The t-SVD sample of a tensor $\mathcal{P} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$.

Definition 2 (t-SVD). Given a tensor $\mathcal{P} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$, the corresponding t-SVD could be denoted as:

$$\mathcal{P} = \mathcal{W} * \mathcal{O} * \mathcal{Q}^T, \tag{4}$$

in which $\mathcal{W} \in \mathbb{R}^{n_1 \times n_1 \times n_3}$ and $\mathcal{Q} \in \mathbb{R}^{n_2 \times n_2 \times n_3}$ are two orthogonal tensors, and $\mathcal{O} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$ is a f-diagonal tensor.

Definition 3 (t-SVD-Based Tensor Nuclear Norm). Given a tensor $\mathcal{P} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$, the corresponding t-SVD-based tensor nuclear norm (i.e., $\| \mathcal{P} \|_{\circledast}$) is defined by the sum of the singular values of all frontal slices, which are as follows:

$$\| \mathcal{P} \|_{\circledast} = \sum_{k=1}^{n_3} \| \mathcal{P}_f^{(k)} \|_* = \sum_{i=1}^{min(n_1,n_2)} \sum_{k=1}^{n_3} | O_f^{(k)}(i,i) |, \tag{5}$$

in which $O_f^{(k)}(i,i)$ is computed by the t-SVD of frontal slices of $\mathcal{P}_f$, i.e., $\mathcal{P}_f^{(k)} = \mathcal{W}_f^{(k)} \mathcal{O}_f^{(k)} \mathcal{Q}_f^{(k)T}$. Especially, Figure 1 shows the t-SVD sample of a tensor $\mathcal{P} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$.

3.2 The Objective Formulation

Before describing our objective formulation, we delve into the anchor multi-view subspace clustering (AMSC) in advance. Formally, let $\mathrm{X}_d = [\mathrm{X}^{(1)}; \ldots; \mathrm{X}^{(V)}]$ be a multi-view dataset consisting of V views, where $\mathrm{X}^{(v)} \in \mathbb{R}^{d_v \times n}$ is the data sub-matrix of the vth view (with n data samples and a dimension of d_v). In conventional MSC studies, all the n data samples are used as the dictionary, and it is assumed that data samples lie in a union of low-dimension subspaces embedded in original feature spaces. Based on this, the data samples of the vth view could be described by following self-expressive formulation:

$$\mathrm{X}^{(v)} \approx \mathrm{X}^{(v)}(\mathrm{Z}^{(v)}), \forall v, \tag{6}$$

in which $\mathrm{Z}^{(v)} \in \mathbb{R}^{n \times n}$ is the subspace representation of the vth view. Then, the basic model for MSC could be formulated in the following form:

$$\min_{\{\mathrm{Z}^{(v)}\}_{v=1}^{V}} \sum_{v=1}^{V} \left\| \mathrm{X}^{(v)} - \mathrm{X}^{(v)}(\mathrm{Z}^{(v)}) \right\|_F^2 + \sum_{v=1}^{V} \Omega(\mathrm{Z}^{(v)}), \tag{7}$$

where $\Omega(.)$ is a specific regularization term for $\{\mathrm{Z}^{(v)}\}_{v=1}^{V}$. As shown in Equation (7), the traditional MSC approaches have the capability to capture the global structural information within each individual view. Yet the time-consuming issue (w.r.t. the $\mathrm{O}(n^3)$) complexity still restricts them from large-scale (or even some general-scale) clustering tasks.

In light of this, several endeavors have been devoted to the context of anchor multi-view subspace clustering. Instead of pursing the subspace representation (with $n \times n$ size) for every view, the anchor subspace-based approaches focus on generating a series of bipartite graphs (also known as anchor graphs) with a smaller size (i.e., $p \times n$). Specifically, the key of the anchor subspace-based approaches is to characterize the local structural relationship between the data samples and the selected anchors. That is, every data sample could be linearly expressed by a small set of anchors, which can be formulated in the following:

$$\mathrm{X}^{(v)} \approx \mathrm{A}^{(v)}(\mathrm{B}^{(v)}), \forall v, \tag{8}$$

in which $\mathrm{A}^{(v)}$ stands for the anchor set of the vth view, while $\mathrm{B}^{(v)}$ stands for the anchor representation in the same view. Notice that the anchor set within each view is usually generated by k-means or by random sampling. Based on the above equation, we can obtain the following basic model for anchor multi-view subspace clustering:

$$\min_{\{\mathrm{B}^{(v)}\}_{v=1}^{V}} \sum_{v=1}^{V} \left\| \mathrm{X}^{(v)} - \mathrm{A}^{(v)}(\mathrm{B}^{(v)}) \right\|_F^2 + \sum_{v=1}^{V} \Theta(\mathrm{B}^{(v)}), \tag{9}$$

where $\Theta(.)$ is a specific regularization term for $\{\mathrm{B}^{(v)}\}_{v=1}^{V}$. This basic model (i.e., Equation (9)) inspires a union of promising works in anchor subspace-based approaches, which can efficiently accelerate the computation by utilizing the specific structure of bipartite graph. In spite of this, we argue that their performance can be further enhanced by capturing the inherent properties from deeper viewpoint.

Next, we proceed to formulate the objective function of our proposed approach. Suppose we have a nonlinear mapping $\phi(.)$ that is guided by a specific kernel function. Then the whole data samples could be linearly characterized by the selected anchors in the kernel space, that is,

$$\phi(\mathrm{X}^{(v)}) \approx \phi(\mathrm{A}^{(v)})(\mathrm{B}^{(v)}), \forall v. \tag{10}$$

Thereby, the kernel version of the basic model in Equation (9) can be given as follows:

$$\min_{\{\mathrm{B}^{(v)}\}_{v=1}^{V}} \sum_{v=1}^{V} \left\| \phi(\mathrm{X}^{(v)}) - \phi(\mathrm{A}^{(v)})(\mathrm{B}^{(v)}) \right\|_F^2 + \sum_{v=1}^{V} \Omega(\mathrm{B}^{(v)}). \tag{11}$$

The aforementioned model has the advantage in handling more complex clustering problem, i.e., large-scale multi-view data with nonlinear distribution. Further, we impose the t-SVD-based tensor norm on the anchor representations, aiming at exploring the high-order correlation hidden in different bipartite graphs (with each bipartite graph corresponding to one view). Mathematically, this presentation leads to the optimization problem by:

$$\begin{aligned} \min_{\{\mathrm{B}^{(v)}\}_{v=1}^{V}} & \sum_{v=1}^{V} \left\| \phi(\mathrm{X}^{(v)}) - \phi(\mathrm{A}^{(v)})(\mathrm{B}^{(v)}) \right\|_F^2 + \lambda_1 \|\mathcal{B}\|_{\circledast} \\ \text{s.t.} \quad & \mathcal{B} = \Psi(\mathrm{B}^{(1)}, \mathrm{B}^{(2)}, \cdots, \mathrm{B}^{(V)}), \end{aligned} \tag{12}$$

where $\Psi(\cdot)$ is an operator that stacks the multi-view anchor representations $\{\mathrm{B}^{(v)}\}_{v=1}^{V}$ into a 3-mode tensor $\mathcal{B}$, and then rotates its size to $n \times V \times p$, as shown in Figure 2. By introducing the low-rank tensor norm $\|\mathcal{B}\|_{\circledast}$, this model is capable of providing a deeper insight into multi-view anchor representations, which can characterize the low-rank spatial information embedded in heterogeneous views from a high-order perspective.

In the context of graph learning, it has been proven that the performance can be significantly promoted by removing the unnecessary links of fully constructed graphs. Therefore, we adopt the $l_{1,2}$ sparse norm on the anchor representations to preserve the reliable structural information

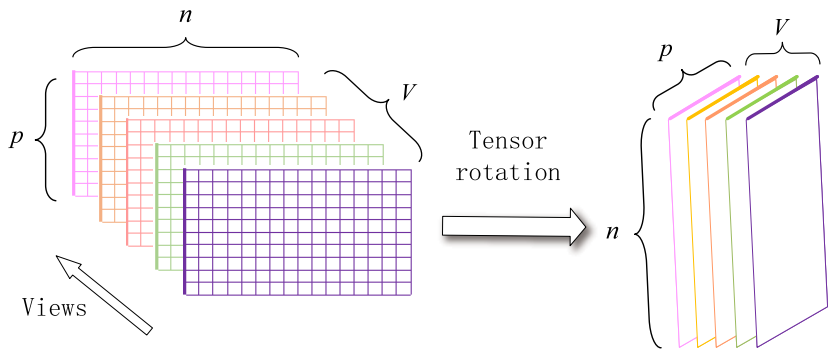


Fig. 2. Example of rotation operation.

within each view. Finally, based on the above considerations, we could formulate the objective function of our LTKMSC approach by:

$$\begin{aligned} \min_{\{\mathrm{B}^{(v)}\}_{v=1}^{V}} & \sum_{v=1}^{V} \left\| \phi(\mathrm{X}^{(v)}) - \phi(\mathrm{A}^{(v)})(\mathrm{B}^{(v)}) \right\|_F^2 + \lambda_1 \|\mathcal{B}\|_{\circledast} \\ & + \lambda_2 \sum_{v=1}^{V} \|\mathrm{B}^{(v)}\|_{1,2}^2 \\ \text{s.t.} \quad & \mathcal{B} = \Psi(\mathrm{B}^{(1)}, \mathrm{B}^{(2)}, \cdots, \mathrm{B}^{(V)}), \end{aligned} \tag{13}$$

where λ_1 and λ_2 are two positive hyper-parameters. Besides, the $l_{1,2}$-norm of $\mathrm{B}^{(v)}$ can be expressed as $\|\mathrm{B}^{(v)}\|_{1,2}^2 = \sum_{i=1}^{p}(\sum_{j=1}^{n} |\mathrm{B}_{ij}{}^{(v)}|)$.

For clarity, we give three remarks as follows.

Remark 1. There are three terms in the overall objective function (13). Specifically, the first term (i.e., *the kernelized reconstruction term*) characterizes the latent relationships among data samples and selected anchors in the kernel space. The second term concerns with *inter-view awareness*, which helps to exploit the high-order complementarity among heterogeneous views. The third term concerns with *intra-view awareness*, whose aims to discover the discriminative structural information embedded within each view. If we incorporate the kernelized reconstruction term with the intra-view awareness term, the following objective model can be achieved:

$$\min_{\{\mathrm{B}^{(v)}\}_{v=1}^{V}} \sum_{v=1}^{V} \left\| \phi(\mathrm{X}^{(v)}) - \phi(\mathrm{A}^{(v)})(\mathrm{B}^{(v)}) \right\|_F^2 + \lambda_2 \sum_{v=1}^{V} \|\mathrm{B}^{(v)}\|_{1,2}^2. \tag{14}$$

This model captures the group effect (i.e., the underlying anchor–sample relationship) within each individual view, such that the view-specific particularity (i.e., the intra-view awareness) can be effectively preserved. In order to further exploit the high-order correlation across multiple views, the inter-view awareness should also be considered. As described in objective function (12), the view-specific anchor representations can be formed into a unified 3-order tensor. Then we employ the t-SVD norm on this 3-order tensor, aiming to encourage different anchor representations to convey the inter-view semantic consistency (i.e., similar low-rank clustering structure). Furthermore, a unified kernerlized subspace clustering framework is presented in objective function (13) to jointly characterize both intra-view and inter-view awareness for enhanced clustering performance.

Remark 2. By using the kernel trick, minimizing the reconstruction term of Equation (13) can be reduced into the following trace form:

$$\begin{aligned}
&\min_{\{\mathrm{B}^{(v)}\}_{v=1}^{V}} \sum_{v=1}^{V} \left\| \phi(\mathrm{X}^{(v)}) - \phi(\mathrm{A}^{(v)})(\mathrm{B}^{(v)}) \right\|_F^2 \\
\Leftrightarrow &\min_{\{\mathrm{B}^{(v)}\}_{v=1}^{V}} \sum_{v=1}^{V} \mathrm{Tr}\{\mathrm{K_X}^{(v)} - 2\mathrm{K_{AX}}^{(v)}\mathrm{B}^{(v)} + \mathrm{B}^{(v)^T}\mathrm{K_A}^{(v)}\mathrm{B}^{(v)}\} \\
\Leftrightarrow &\min_{\{\mathrm{B}^{(v)}\}_{v=1}^{V}} \sum_{v=1}^{V} \mathrm{Tr}\{-2\mathrm{K_{AX}}^{(v)}\mathrm{B}^{(v)} + \mathrm{B}^{(v)^T}\mathrm{K_A}^{(v)}\mathrm{B}^{(v)}\}.
\end{aligned} \tag{15}$$

Here, $\mathrm{K_X}^{(v)}$ stands for the kernel matrix of the vth view, which could be calculated as $\mathrm{K_X}^{(v)} = \phi(\mathrm{X}^{(v)})^T\phi(\mathrm{X}^{(v)})$. Similarity, two useful definitions that relevant to $\mathrm{K_{AX}}^{(v)}$ and $\mathrm{K_A}^{(v)}$ are given as follows:

Definition 4. Given the data samples $\mathrm{X}^{(v)} \in \mathbb{R}^{d_v \times n}$ and selected anchors $\mathrm{A}^{(v)} \in \mathbb{R}^{d_v \times p}$ of the vth view, the corresponding kernel matrix $\mathrm{K_{AX}}^{(v)} \in \mathbb{R}^{p \times n}$ is denoted as:

$$\mathrm{K_{AX}}^{(v)} = \phi(\mathrm{A}^{(v)})^T\phi(\mathrm{X}^{(v)}). \tag{16}$$

i.e.,

$$\mathrm{K_{AX}}^{(v)} = \begin{bmatrix} \langle \phi(\mathrm{A}_{:,1}{}^{(v)}), \phi(\mathrm{X}_{:,1}{}^{(v)}) \rangle & \cdots & \langle \phi(\mathrm{A}_{:,1}{}^{(v)}), \phi(\mathrm{X}_{:,n}{}^{(v)}) \rangle \\ \langle \phi(\mathrm{A}_{:,2}{}^{(v)}), \phi(\mathrm{X}_{:,1}{}^{(v)}) \rangle & \cdots & \langle \phi(\mathrm{A}_{:,2}{}^{(v)}), \phi(\mathrm{X}_{:,n}{}^{(v)}) \rangle \\ \vdots & \ddots & \vdots \\ \langle \phi(\mathrm{A}_{:,p}{}^{(v)}), \phi(\mathrm{X}_{:,1}{}^{(v)}) \rangle & \cdots & \langle \phi(\mathrm{A}_{:,p}{}^{(v)}), \phi(\mathrm{X}_{:,n}{}^{(v)}) \rangle \end{bmatrix}.$$

Definition 5. Given the selected anchors $\mathrm{A}^{(v)} \in \mathbb{R}^{d_v \times p}$ of the vth view, the corresponding kernel matrix $\mathrm{K_A}^{(v)} \in \mathbb{R}^{p \times p}$ is denoted as:

$$\mathrm{K_A}^{(v)} = \phi(\mathrm{A}^{(v)})^T\phi(\mathrm{A}^{(v)}). \tag{17}$$

i.e.,

$$\mathrm{K_A}^{(v)} = \begin{bmatrix} \langle \phi(\mathrm{A}_{:,1}{}^{(v)}), \phi(\mathrm{A}_{:,1}{}^{(v)}) \rangle & \cdots & \langle \phi(\mathrm{A}_{:,1}{}^{(v)}), \phi(\mathrm{A}_{:,p}{}^{(v)}) \rangle \\ \langle \phi(\mathrm{A}_{:,2}{}^{(v)}), \phi(\mathrm{A}_{:,1}{}^{(v)}) \rangle & \cdots & \langle \phi(\mathrm{A}_{:,2}{}^{(v)}), \phi(\mathrm{A}_{:,p}{}^{(v)}) \rangle \\ \vdots & \ddots & \vdots \\ \langle \phi(\mathrm{A}_{:,p}{}^{(v)}), \phi(\mathrm{A}_{:,1}{}^{(v)}) \rangle & \cdots & \langle \phi(\mathrm{A}_{:,p}{}^{(v)}), \phi(\mathrm{A}_{:,p}{}^{(v)}) \rangle \end{bmatrix}.$$

Subsequently, a remark is introduced in what follows.

Remark 3. If we set $\lambda_1 \to \infty$, then the objective formulation of Equation (13) is reduced as follows:

$$\begin{aligned}
&\min_{\mathcal{B}} \|\mathcal{B}\|_{\circledast} \\
&\text{s.t.} \quad \mathcal{B} = \Psi(\mathrm{B}^{(1)}, \mathrm{B}^{(2)}, \cdots, \mathrm{B}^{(V)}).
\end{aligned} \tag{18}$$

Given a series of anchor representations (or anchor graphs), if the condition $\lambda_1 \to \infty$ hold, the aforementioned problem is equivalent to the tensor learning problem w.r.t. multiple bipartite graphs.

3.3 Optimization Algorithm

Under the ADMM framework, a fast optimization algorithm is presented in this subsection. Firstly, by combining Equations (13) and (18), the objective formulation is rewritten in the following form:

$$\begin{aligned}\min_{\{\mathrm{B}^{(v)}\}_{v=1}^{V}} & \sum_{v=1}^{V} \mathrm{Tr}\{-2\mathrm{K_{AX}}^{(v)}\mathrm{B}^{(v)} + \mathrm{B}^{(v)^T}\mathrm{K_A}^{(v)}\mathrm{B}^{(v)}\} + \lambda_1\|\mathcal{B}\|_{\circledast} \\ & + \lambda_2 \sum_{v=1}^{V} \|\mathrm{B}^{(v)}\|_{1,2}^2 \\ \text{s.t.} \quad & \mathcal{B} = \Psi(\mathrm{B}^{(1)}, \mathrm{B}^{(2)}, \cdots, \mathrm{B}^{(V)}).\end{aligned} \tag{19}$$

After introducing two kinds of auxiliary variables $\{\mathrm{C}^{(v)}\}_{v=1}^{V}$ and $\{\mathrm{P}^{(v)}\}_{v=1}^{V}$, the augmented Lagrangian function should be considered in what follows:

$$\begin{aligned}& \mathcal{L}\left(\mathcal{C}, \{\mathrm{B}^{(v)}\}_{v=1}^{V}, \{\mathrm{P}^{(v)}\}_{v=1}^{V}\right) \\ =& \sum_{v=1}^{V} \mathrm{Tr}\{-2\mathrm{K_{AX}}^{(v)}\mathrm{B}^{(v)} + \mathrm{B}^{(v)^T}\mathrm{K_A}^{(v)}\mathrm{B}^{(v)}\} + \lambda_1\|\mathcal{C}\|_{\circledast} \\ & + \lambda_2 \sum_{v=1}^{V} \|\mathrm{P}^{(v)}\|_{1,2}^2 + (\mathcal{Y}_1, \mathcal{B} - \mathcal{C}) + \frac{\mu}{2}\|\mathcal{B} - \mathcal{C}\|_F^2 \\ & + \sum_{v=1}^{V}\left(\left\langle \mathrm{Y}_2^{(v)}, \mathrm{B}^{(v)} - \mathrm{P}^{(v)}\right\rangle + \frac{\rho}{2}\left\|\mathrm{B}^{(v)} - \mathrm{P}^{(v)}\right\|_F^2\right) \\ \text{s.t.} \quad & \mathcal{B} = \Phi(\mathrm{B}^{(1)}, \mathrm{B}^{(2)}, \cdots, \mathrm{B}^{(V)}), \\ & \mathcal{C} = \Phi(\mathrm{C}^{(1)}, \mathrm{C}^{(2)}, \cdots, \mathrm{C}^{(V)}),\end{aligned} \tag{20}$$

in which matrices tensor $\mathcal{Y}_1$ and $\{\mathrm{Y}_2^{(v)}\}_{v=1}^{V}$ are two kinds of Lagrange multipliers, while ρ and μ are two penalty parameters. Consequently, an iterative algorithm is applied to solve the subproblems w.r.t. $\mathcal{C}$, $\{\mathrm{B}^{(v)}\}_{v=1}^{V}$, $\{\mathrm{P}^{(v)}\}_{v=1}^{V}$, $\mathcal{Y}_1$, and $\{\mathrm{Y_2}^{(v)}\}_{v=1}^{V}$ by turns.

3.3.1 Solving the $\mathcal{C}$-Subproblem. Fix variables $\{\mathrm{B}^{(v)}\}_{v=1}^{V}$ and $\{\mathrm{P}^{(v)}\}_{v=1}^{V}$, the subproblem of Equation (20) w.r.t. $\mathcal{C}$ becomes the following form:

$$\min_{\mathcal{C}} \ \lambda_1||\mathcal{C}||_{\circledast} + \frac{\mu}{2}||\mathcal{C} - \mathcal{Q}||_F^2. \tag{21}$$

Here, $\mathcal{Q} = \mathcal{B} + \frac{\mathcal{Y}_1}{\mu}$. Inspired closely by the research in [44], we solve this subproblem by tensor tubal-shrinkage operator [12]:

$$\mathcal{D} = \mathcal{C}_{n_3\pi}(\mathcal{Q}) = \mathcal{U} * \mathcal{C}_{n_3\pi}(\mathcal{O}) * \mathcal{V}^T. \tag{22}$$

Here, $\pi = \mu/\lambda_1$ and $\mathcal{Q} = \mathcal{U} * \mathcal{O} * \mathcal{V}^T$. In addition, $\mathcal{C}_{n_3}(\mathcal{Q}) = \mathcal{Q} * \mathcal{W}$, herein, tensor $\mathcal{W}$ is f-diagonal, and $\mathcal{W}_f(i,i,j) = (1 - \frac{n_3\pi}{\mathcal{O}_f^{(j)}(i;i)})_+$ means the diagonal element of $\mathcal{W}$ in the Fourier domain.

3.3.2 Solving the $\mathrm{B}^{(v)}$-Subproblem. Fix variables $\mathcal{C}$ and $\{\mathrm{P}^{(v)}\}_{v=1}^{V}$, the subproblem of Equation (20) w.r.t. $\mathrm{B}^{(v)}$ becomes the following form (Notice that the conditions $\Psi_{(v)}^{-1}(\mathcal{B}) = \mathrm{B}^{(v)}$

and $\Psi_{(v)}^{-1}(\mathcal{P}) = \mathrm{P}^{(v)}$):

$$\begin{aligned}\min_{\mathrm{B}^{(v)}} \mathrm{Tr}\{\mathrm{B}^{(v)^T}(\mathrm{K_A}^{(v)} + \frac{\mu+\rho}{2})\mathrm{B}^{(v)}\} \\ + \mathrm{Tr}\{(-2\mathrm{K_{AX}}^{(v)} + \mathrm{M}^{(v)})\mathrm{B}^{(v)}\}.\end{aligned} \tag{23}$$

Here, $\mathrm{M}^{(v)} = \mathrm{Y_1}^{(v)^T} - \mu\mathrm{C}^{(v)^T} + \mathrm{Y_2}^{(v)^T} - \rho\mathrm{P}^{(v)^T}$. Thus, the optimal solution for $\mathrm{B}^{(v)}$ is given by:

$$\mathrm{B}^{(v)} = (\mathrm{K_A}^{(v)} + (\frac{\mu+\rho}{2})\mathrm{I})^{-1}(\mathrm{K_{AX}}^{(v)} + (\mathrm{M}^{(v)})^T). \tag{24}$$

3.3.3 Solving the $\mathrm{P}^{(v)}$-Subproblem. Fix variables C and $\{\mathrm{B}^{(v)}\}_{v=1}^{V}$, the subproblem of Equation (20) w.r.t. $\mathrm{P}^{(v)}$ becomes the following form:

$$\min_{\mathrm{P}^{(v)}} \lambda_2||\mathrm{P}^{(v)}||_{1,2}^2 + \frac{\rho}{2}||\mathrm{P}^{(v)} - (\mathrm{B}^{(v)} + \frac{\mathrm{Y_2}^{(v)}}{\rho})||_F^2. \tag{25}$$

Following closely the relevant research, this subproblem can be efficiently solved by [28].

Also, two kinds of multipliers $\{\mathrm{Y}_1^{(v)}\}_{v=1}^{V}$ and $\mathcal{Y}_2$ can be obtained by using the following equations:

$$\begin{cases}\mathcal{Y}_1 = \mathcal{Y}_1 + \mu(\mathcal{B} - C) \\ \mathrm{Y_2}^{(v)} = \mathrm{Y_2}^{(v)} + \rho(\mathrm{B}^{(v)} - \mathrm{P}^{(v)})\end{cases}. \tag{26}$$

In a nutshell, the above four steps are iteratively updated until the convergence conditions in Equation (28) are met. For clarity, the overall process of our LTKMSC algorithm is given in Algorithm 1. When the iterative optimization is finished, the output of $\{\mathrm{C}^{(v)}\}_{v=1}^{V}$ can be formulated into a unified representation $\overline{\mathrm{C}}$, that is,

$$\overline{\mathrm{C}} = \sqrt{V}\begin{bmatrix}\mathrm{C}^{(1)} \\ \mathrm{C}^{(2)} \\ \vdots \\ \mathrm{C}^{(V)}\end{bmatrix}. \tag{27}$$

Then the **singular value decomposition (SVD)** is conducted on the unified representation $\overline{\mathrm{C}}$, denoted as $\overline{\mathrm{C}} = \mathrm{SVD}(\mathrm{U_C}\Lambda_\mathrm{C}\mathrm{W_C}^T)$. Following that, we can run the k-means clustering on the obtained matrix $\mathrm{W_C}$ and thus generate the final clustering result.

3.4 Computational and Space Complexity Analysis

In the first place, this section analyzes the time complexity and space complexity of our LTKMSC approach, as described in Section 3.4.1 and Section 3.4.2, respectively.

3.4.1 Time Complexity. The time complexity of Algorithm 1 mainly involves two stage, i.e., initialization process and optimization process. In the first stage, the initialization of $\{\mathrm{A}^{(v)}\}_{v=1}^{V}$ and $\{\mathrm{B}^{(v)}\}_{v=1}^{V}$ require $\mathrm{O}(Vnp\sum_{v=1}^{V} d_v)$ and $\mathrm{O}(Vnp\log(p))$ time cost, with V, n, p, d_v stand for the number of views, data samples, anchors, and dimensions of the vth view, respectively. Besides, the calculation of $\{\mathrm{K_{AX}}^{(v)}\}_{v=1}^{V}$ and $\{\mathrm{K_A}^{(v)}\}_{v=1}^{V}$ need $\mathrm{O}(np\sum_{v=1}^{V} d_v)$ and $\mathrm{O}(p^2\sum_{v=1}^{V} d_v)$ time cost, respectively. In the second stage, the update for C, $\{\mathrm{B}^{(v)}\}_{v=1}^{V}$, and $\{\mathrm{P}^{(v)}\}_{v=1}^{V}$ require $\mathrm{O}(Vnp\log(Vn) + V^2np)$, $\mathrm{O}(Vp^3 + Vnp^2)$, and $\mathrm{O}(Vnp + np^2)$ time cost, respectively. Suppose Algorithm 1 runs T times to convergence, since the condition $p, T, V \ll n, d_v$ is satisfied, the final time complexity of optimization algorithm lies in $\mathrm{O}(n\log n + n\sum_{v=1}^{V} d_v)$, which enjoys highly efficient computational speed.

Algorithm 1: LTKMSC

Input: Dataset with V views $\{\mathrm{X}^{(v)}\}_{v=1}^{V}$, the maximum iterations it_{max}, number of selected anchors p and two balancing hyper-parameter $\lambda_1, \lambda_2 > 0$.

Parameter Setup: Set $\mu = 10^2$, $\rho = 1$, $\eta_1 = \eta_2 = 1.3$, $\mu_{max} = \rho_{max} = 10^6$ and $\varepsilon = 10^{-4}$.

Initialization: Initialize $\mathrm{D}^{(v)} = \mathbf{0}, \forall v$, $\mathrm{P}^{(v)} = \mathbf{0}, \forall v$, $\mathrm{Y}_1^{(v)} = \mathbf{0}, \forall v$ and $\mathcal{Y}_2 = \mathbf{0}$. Adopt the same method in [41] to initialize $\mathrm{A}^{(v)}, \forall v$ and $\mathrm{B}^{(v)}, \forall v$. Calculate the kernels $\mathrm{K_{AX}}^{(v)}, \forall v$ and $\mathrm{K_A}^{(v)}, \forall v$ by Eq. (16) and Eq. (17), respectively. Set $it = 1$.

repeat

  Obtain $\mathcal{C}$ by Eq. (22).

  Obtain $\{\mathrm{B}^{(v)}\}_{v=1}^{V}$ by Eq. (24).

  Obtain $\{\mathrm{P}^{(v)}\}_{v=1}^{V}$ by the process in research [28].

  Obtain $\mathcal{Y}_1$ and $\{\mathrm{Y_2}^{(v)}\}_{v=1}^{V}$ by Eq. (26).

  Update μ by using $\mu = \min(\eta_1\mu, \mu_{\max})$.

  Update ρ by using $\rho = \min(\eta_2\rho, \rho_{\max})$.

  $it = \ it + 1$.

until The following convergence conditions are met or $it > it_{max}$.

$$\|\mathcal{B} - \mathcal{C}\|_F < \varepsilon \text{ and } \|\mathcal{B} - \mathcal{P}\|_F < \varepsilon \tag{28}$$

Output: $\{\mathrm{C}^{(v)}\}_{v=1}^{V}$.

3.4.2 Space Complexity. In the first stage, the calculation of $\{\mathrm{K_{AX}}^{(v)}\}_{v=1}^{V}$ and $\{\mathrm{K_A}^{(v)}\}_{v=1}^{V}$ require $\mathrm{O}(Vnp)$ and $\mathrm{O}(Vp^2)$ space cost, respectively. Additionally, the initialization of $\{\mathrm{A}^{(v)}\}_{v=1}^{V}$ and $\{\mathrm{B}^{(v)}\}_{v=1}^{V}$ require $\mathrm{O}(p\sum_{v=1}^{V} d_v)$ and $\mathrm{O}(Vnp)$ space cost, respectively. In the second stage, the update for $\mathcal{C}$, $\{\mathrm{B}^{(v)}\}_{v=1}^{V}$, $\{\mathrm{P}^{(v)}\}_{v=1}^{V}$ $\mathcal{Y}_1$, and $\{\mathrm{Y_2}^{(v)}\}_{v=1}^{V}$ require the same space cost of $\mathrm{O}(Vnp)$. Thereby, the overall space complexity of optimization algorithm lies in $\mathrm{O}(n + \sum_{v=1}^{V} d_v)$, whose storage is linear in data samples or dimensions of all views.

4 Experiments

This section evaluates our LTKMSC approach against several state-of-the-art approaches. Notice that all experiments are carried out on a PC with an i5-6600 CPU and 64 GB of RAM.

4.1 Experimental Settings

To testify the effectiveness and efficiency of our LTKMSC approach, we employ 11 multi-view datasets in our experiments, including 6 general-scale datasets (with $n < 9{,}000$) and 5 large-scale datasets (with $n \geq 9{,}000$). It is worth that these datasets are collected with different characteristics, which have been commonly adopted for experimental purpose.

On the one hand, the six general-scale datasets involves *ORL* [4], *Notting-Hill* [27], *BBCSport* [23], *Caltech101-7* [30], *UCI Digits* [21], and *NH-4660* [3]. On the other hand, the five large-scale datasets involves *Caltech101-all* [49], *ALOI-100* [49], *Reuters-lee* [24], *NUSWIDEOBJ* [24], and *AwA* [29]. The summarization of these multi-view datasets is listed in Table 1.

To empirically evaluate the clustering quality of different approaches, we employ two widely used metrics for evaluation. They are **Normalized Mutual Information (NMI)** [16] and Accuracy [31, 51]. On each dataset, we conduct each clustering approach 20 times and report its average score.

Table 1. Multi-View Datasets Used in the Experiments

Dataset		#Sample	#View	#Class	Views and Dimensions
General	*ORL*	400	3	40	LBP(3,304), Intensity(4,096), Gabor(6,750)
	BBCSport	544	2	5	View 1(3,183), View 2(3,203)
	Notting-Hill	550	3	5	Intensity(2,000), LBP(3,304), Gabor(6,750)
	Caltech101-7	1,474	3	7	GIST(512), LBP(928), HOG(1,984)
	UCI Digits	2,000	3	10	Pix(240), Fou(76), MOR(6)
	NH-4660	4,660	3	5	Intensity(2,000), LBP(3,304), Gabor(6,750)
Large	*Caltech101-all*	9,144	6	102	WM(40), Gabor(48), CENTRIST(254), GIST(512), LBP(928), HOG(1,984)
	ALOI-100	10,800	4	100	Haralick(13), HSV(64), Similarity(77), RGB(125)
	Reuters-lee	18,758	5	6	English(21,513), French(24,892), German(34,251), Spanish(15,506), Italian(11,547)
	NUSWIDEOBJ	30,000	5	31	CH(64), CM(255), CORR(144), EDH(73), WT(128)
	AWA	30,076	6	50	CQ(2,688), LSS(2,000), PHOG(252), SIFT(2,000), RGSIFT(2,000), SURF(2,000)

4.2 Comparison against Other Approaches

This subsection empirically investigates our LTKMSC approach by comparing it against various multi-view clustering baselines, they are:

— *DiMSC* [3]: Diversity-induced MSC.
— *LMSC* [48]: Latent MSC.
— *SwMC* [30]: Self-weighted multi-view clustering with multiple graph learning.
— *MLAN* [29]: Multi-view graph clustering with adaptive neighbor learning.
— *MVSC* [21] : Large-scale multi-view spectral clustering by means of bipartite graph.
— *SFMC* [20]: Scalable and parameter-free multi-view graph clustering.
— *LMVSC* [17]: Linear MSC for large-scale data.
— *SMVSC* [34]: Scalable MSC with unified anchors.
— *FPMVS* [38]: Fast parameter-free MSC with consensus anchor guidance.
— *OMSC* [5]: Efficient orthogonal MSC.

Based on the different perspectives, these 10 baseline approaches for comparison can be categorized into two classes, namely traditional approaches and scalable approaches. The traditional approaches include two graph-based approaches (e.g., SwMC and MLAN) and two subspace-based approaches (e.g., DiMSC and LMSC). The scalable approaches include two anchor graph-based approaches (i.e., MVSC and SFMC) and four anchor subspace-based approaches (e.g., LMVSC, SMVSC, FPMVS, and OMSC). The detailed information refer to them can be found in the original papers.

It is noteworthy that, the following experimental setting should be observed in our experiments:

— The code of comparison approaches are directly downloaded in public access. At the same time, their hyper-parameters will be varied within the range of $[10^{-5}, 10^{-4}, \cdots, 10^{4}, 10^{5}]$, except some specific range (or value) is suggested in the original paper.
— As for our approach LTKMSC, the hyper-parameters λ_1 and λ_2 are varied within the range of $[10^{-4}, 10^{-3}, \cdots, 10^{3}, 10^{4}]$, while the anchor number is varied within the range of $[100, \cdots, 500]$ with step 100. For small-scale datasets ($n < 1{,}000$), the anchor number is varied within the range of $[100, 150, 200, 250, 300]$. Notice that the corresponding kernel constructions and the default parameters of our proposed approach are given in Table 2.
— Under the same computing environment, we repeat each comparison approach for 20 times, and report their best clustering performance via cross-validation strategy.

Table 2. Kernel Constructions and Parameter Settings of LTKMSC on the Benchmark Datasets

Dataset	View 1	View 2	View 3	View 4	View 5	View 6	λ_1	λ_2	p
ORL	Linear	Linear	Linear	-	-	-	10^3	10^{-2}	300
BBCSport	Cosine	Cosine	-	-	-	-	10^3	10^{-3}	150
Notting-Hill	Cosine	Cosine	Cosine	-	-	-	10^3	10^{-3}	300
Caltech101-7	Cosine	Gaussian	Cosine	-	-	-	10^3	10^{-1}	300
UCI Digits	Cosine	Cosine	Cosine	-	-	-	10^3	10^{-3}	100
NH-4660	Cosine	Cosine	Cosine	-	-	-	10^3	10^{-2}	300
Caltech101-all	Cosine	Cosine	Cosine	Cosine	Cosine	Cosine	10^4	10^3	200
ALOI-100	Cosine	Cosine	Gaussian	Gaussian	-	-	10^4	10^{-3}	200
Reuters-lee	Gaussian	Gaussian	Cosine	Cosine	Cosine	-	10^4	10	200
NUSWIDEOBJ	Gaussian	Gaussian	Gaussian	Gaussian	Gaussian	-	10^4	10	100
AWA	Gaussian	Gaussian	Gaussian	Gaussian	Gaussian	Gaussian	10^4	10^{-1}	100

Table 3. The Average NMI (%) Scores over 20 Runs by Different Multi-View Clustering Approaches

Datasets	DiMSC	LMSC	SwMC	MLAN	MVSC	SFMC	LMVSC	SMVSC	FPMVS	OMSC	LTKMSC
ORL	90.78	[92.23]	83.31	78.58	84.69	80.64	76.42	75.26	74.30	78.38	**96.41**
Notting-Hill	78.22	81.05	82.04	64.69	77.09	[84.72]	72.59	82.49	71.35	83.32	**97.81**
BBCSport	68.95	75.18	46.93	51.37	69.29	5.15	[82.88]	14.83	10.37	12.44	**99.34**
Caltech101-7	50.92	48.74	52.97	[63.58]	54.46	53.78	43.53	47.79	47.32	51.69	**79.48**
UCI Digits	79.15	77.91	69.75	[93.44]	64.38	75.20	75.60	80.40	73.49	83.58	**99.59**
NH-p4660	78.55	67.37	8.77	55.50	53.98	[84.11]	68.06	66.64	67.15	72.15	**95.95**
Caltech101-all	N/A	N/A	15.92	N/A	23.48	33.42	[37.23]	37.02	35.91	36.54	**86.63**
ALOI-100	N/A	N/A	N/A	N/A	30.48	60.80	[74.95]	57.38	55.51	49.29	**92.84**
Reuters-lee	N/A	N/A	N/A	N/A	N/A	5.81	[44.46]	32.22	34.60	32.30	**95.68**
NUSWIDEOBJ	N/A	N/A	N/A	N/A	N/A	8.13	8.55	12.02	12.42	[12.65]	**21.54**
AwA	N/A	N/A	N/A	N/A	N/A	5.45	7.54	9.57	9.71	[10.22]	**93.80**
Average Score	N/A	N/A	N/A	N/A	N/A	45.20	[53.80]	46.87	44.74	47.51	**87.19**
Average Rank	6.18	6.73	8.00	8.09	7.91	5.45	5.36	5.73	6.91	[4.64]	**1.00**

On each dataset, the best score is highlighted in bold, while the second best one in [brackets].
N/A means the out-of-memory error.

The empirical results w.r.t. NMI and ACC are presented in Tables 3 and 4, respectively. Notice that if a multi-view clustering approach cannot be applied on a dataset (owing to the out-of-memory error), then the score of this approach on this dataset will be marked as "N/A." As the experimental results suggested, three aspects of important findings can be concluded as below.

—For the general-scale cases, it can be seen that the traditional approaches are overall inferior to that of the scalable approaches. Especially, traditional subspace-based approaches (e.g., DiMSC and LMSC) have achieved promising clustering results on five general-scale datasets except *Caltech101-7*, which demonstrates the effectiveness of subspace learning for handling multi-view clustering problems.
—For the large-scale cases, we can see that the traditional approaches are not computationally flexible for datasets with ten-thousand-scale. Notably, out of 11 multi-view clustering approaches, only the scalable approach and our LTKMSC approach have the capability to deal with all of the experimental datasets. This validates the necessity of bipartite (or anchor) graph learning for large-scale multi-view clustering.
—It turns out that the proposed approach has obtained very competitive performance on all the benchmark datasets. Especially, our approach achieves nearly perfect results on the

Table 4. The Average Accuracy (%) Scores over 20 Runs by Different Multi-View Clustering Approaches

Datasets	DiMSC	LMSC	SwMC	MLAN	MVSC	SFMC	LMVSC	SMVSC	FPMVS	OMSC	LTKMSC
ORL	80.90	[81.29]	70.75	68.40	72.36	63.50	57.00	56.25	56.00	61.50	**95.75**
Notting-Hill	83.45	90.45	87.82	67.82	81.20	84.73	80.73	91.45	70.55	[91.64]	**99.09**
BBCSport	86.53	85.47	62.87	67.35	77.67	34.19	[94.67]	32.35	40.63	35.66	**99.82**
Caltech101-7	57.14	57.47	64.79	[78.09]	63.68	63.98	45.05	46.68	54.55	69.00	**85.01**
UCI Digits	86.51	86.28	65.50	[97.02]	61.95	75.20	80.30	83.35	72.20	91.35	**99.85**
NH-p4660	84.27	74.63	33.88	59.21	62.93	[91.89]	73.76	71.61	71.31	87.15	**98.20**
Caltech101-all	N/A	N/A	16.09	N/A	20.08	16.52	11.66	26.65	29.48	[31.91]	**55.64**
ALOI-100	N/A	N/A	N/A	N/A	12.36	45.44	[56.61]	34.30	31.53	26.05	**83.85**
Reuters-lee	N/A	N/A	N/A	N/A	N/A	32.94	[59.46]	55.21	57.48	48.03	**95.61**
NUSWIDEOBJ	N/A	N/A	N/A	N/A	N/A	11.89	12.06	18.85	[19.18]	18.71	**22.31**
AwA	N/A	N/A	N/A	N/A	N/A	6.58	6.77	9.14	9.12	[9.19]	**86.21**
Average Score	N/A	N/A	N/A	N/A	N/A	47.90	[52.55]	47.80	46.55	51.84	**83.76**
Average Rank	6.09	6.36	7.82	8.09	8.00	5.91	6.00	5.91	6.55	[4.27]	**1.00**

On each dataset, the best score is highlighted in bold, while the second best one in [brackets].
N/A means the out-of-memory error.

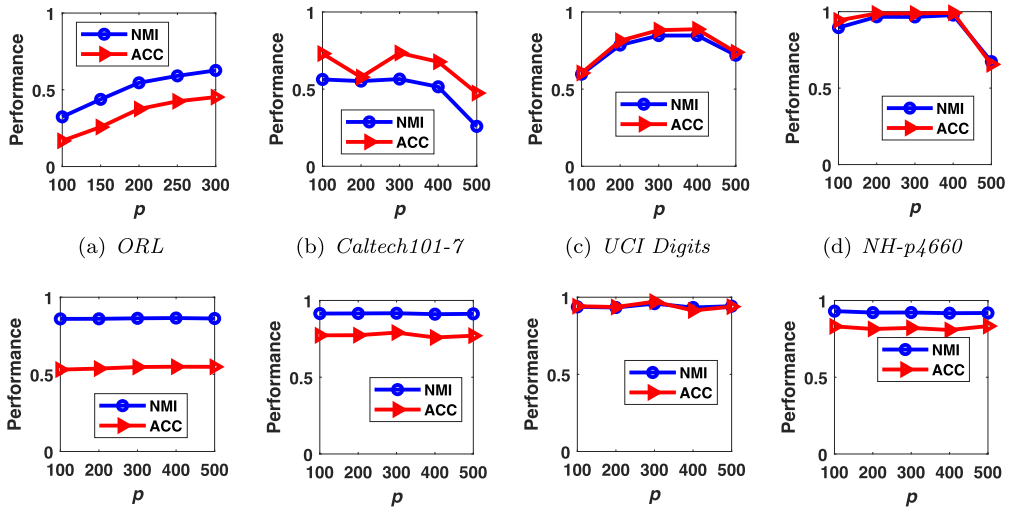
BBCSport and *UCI Digits* datasets, while gaining impeccable performance on the *Notting-Hill* and *NH-p4660* datasets. Beyond the expectation, our approach is remarkably superior to the second-best competitor on four large-scale datasets, namely *Caltech101-all*, *ALOI-100*, *Reuters-lee*, and *AwA* (with large margins over 20%). Take *AWA* dataset as an example, it obtains an increase of 75.57% and 83.51% improvements over the OMSC approach by NMI and Accuracy, respectively. Surprisingly, our proposed approach is ranked in the first position on all 11 datasets over NMI and Accuracy, as shown in the last line of Tables 3 and 4.

As a summary, these experimental results have confirmed the robust performance of our LTKMSC approach when compared with the other comparison approaches.

4.3 Parameter Sensitivity

This subsection practically investigates the influence of λ_1, λ_2, and p in our approach. The λ_1 and λ_2 are two balancing hyper-parameters that control the strength of the regularization term, whereas the parameter p represents the number of selected anchors. In the first palace, we test the parameter p whereas the hyper-parameters λ_1 and λ_2 are fixed as constraint value ($\lambda_1 = 10{,}000$ and $\lambda_2 = 0.001$). The performance over two metrics on eight benchmark datasets is shown in Figure 3. As depicted in this figure, adopting a larger value of p always generates favorable for general-scale datasets. Especially, as p goes from 200 to 300, the proposed approach shows competitive clustering performance on *UCI Digits* and *NH-p4660* datasets. When it comes to large-scale datasets, we find that our approach performs stable over a wide range of parameter p. This demonstrates that our approach is not very sensitive to the parameter p for large-scale clustering tasks.

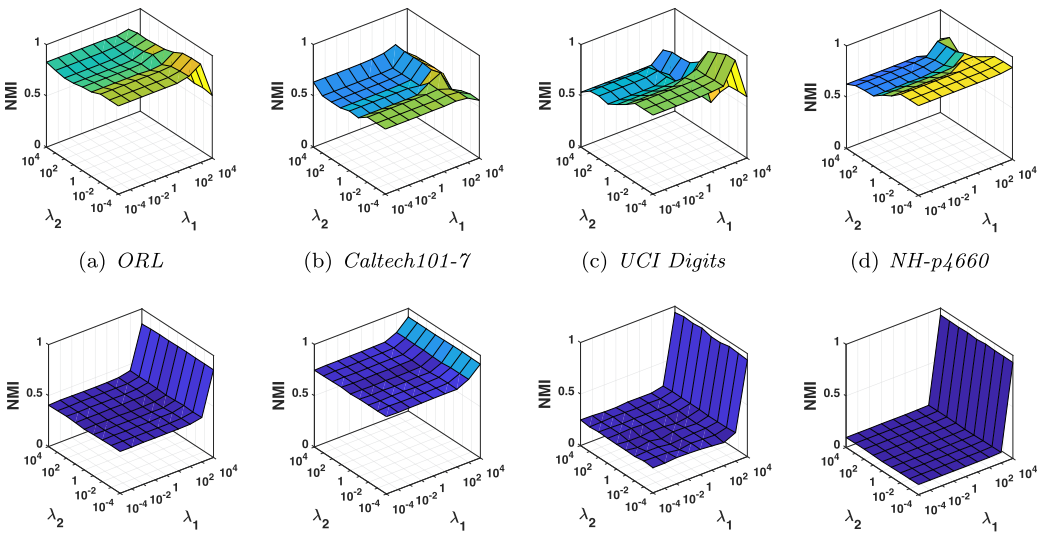
Moving forward, we further carry out experiments to study the hyper-parameters λ_1 and λ_2 in our LTKMSC approach. As previously, we study their influence while fixing the left parameter p as static value. Especially, the parameters λ_1 and λ_2 are varied in the same range of $[10^{-5}, 10^{-4}, \cdots, 10^{4}, 10^{5}]$. Note that, we only report the results over NMI in the following experiments since the Accuracy results have a similar trend. The corresponding results are shown in Figure 4. Based on the this, we make two principal observations. First of all, our approach achieves relatively stable performance on the general-scale datasets, with different combination of parameters λ_1 and λ_2. In addition,



(a) *ORL* (b) *Caltech101-7* (c) *UCI Digits* (d) *NH-p4660*

(e) *Caltech101-all* (f) *ALOI-100* (g) *Reuters-lee* (h) *AwA*

Fig. 3. The clustering performance of LTKMSC on the anchor number p while fixing the hyper-parameters λ_1 and λ_2 as static values.



(a) *ORL* (b) *Caltech101-7* (c) *UCI Digits* (d) *NH-p4660*

(e) *Caltech101-all* (f) *ALOI-100* (g) *Reuters-lee* (h) *AwA*

Fig. 4. Parameter study (in terms of NMI) on the parameters λ_1 and λ_2 while fixing the parameter k as static value.

when relatively large parameter λ_1 (i.e., $\lambda_1 \geq 10^3$) and relatively small parameter λ_2 (i.e., $\lambda_2 \leq 1$) is selected, our approach has capability to obtain fairly better clustering results in most cases. Consequently, we argue that a practical combination of these two parameters (λ_1 and λ_2) is crucial in our LTKMSC approach for various applications.

Table 5. The Average Running Time over 20 Runs by Different Multi-View Clustering Approaches

Datasets	DiMSC	LMSC	SwMC	MLAN	MVSC	SFMC	LMVSC	SMVSC	FPMVS	OMSC	LTKMSC
ORL	3.95	69.09	7.39	**0.47**	2.77	[1.00]	1.50	10.73	22.90	19.52	1.18
Notting-Hill	8.04	76.38	22.92	**0.79**	3.23	[1.05]	1.78	4.10	7.98	6.88	1.78
BBCSport	6.63	21.45	16.48	25.05	1.81	**0.46**	[0.75]	3.94	7.09	13.11	1.62
Caltech101-7	68.97	172.62	220.47	4.85	4.01	2.91	**2.45**	10.11	18.11	36.72	[2.58]
UCI Digits	116.93	344.20	521.25	12.35	3.78	[3.73]	4.42	14.01	21.99	22.70	**1.97**
NH-p4660	1,340.31	4,732.38	5,709.74	255.79	52.98	[5.74]	23.75	33.95	66.69	321.18	**3.50**
Caltech101-all	N/A	N/A	5,012.04	N/A	231.75	**21.61**	72.56	682.34	400.46	289.54	[61.06]
ALOI-100	N/A	N/A	N/A	N/A	205.71	[25.65]	25.46	124.94	241.03	192.33	**18.04**
Reuters-lee	N/A	N/A	N/A	N/A	N/A	95.58	[43.37]	7,373.98	1,205.08	19,379.73	**38.14**
NUSWIDEOBJ	N/A	N/A	N/A	N/A	N/A	**59.47**	91.99	203.15	386.47	309.96	[62.85]
AwA	N/A	N/A	N/A	N/A	N/A	**86.50**	629.28	481.79	964.13	772.75	[316.75]
Average Score	N/A	N/A	N/A	N/A	N/A	**27.61**	81.57	813.00	303.81	1,942.22	[46.32]
Average Rank	7.82	9.55	9.45	7.45	6.27	**1.91**	2.82	5.27	6.73	6.73	[2.00]

On each dataset, the lowest time cost is highlighted in bold, while the second lowest one in [brackets].
N/A means the out-of-memory error.

4.4 Computational Efficiency

This subsection analyzes the time efficiency of all comparison approaches on 11 multi-view datasets. The corresponding computational results are presented in Table 5. By observing, two aspects of finding can be concluded in what follows.

- —For the general-scale cases, traditional approaches require more time over the scalable approaches on the medium-scale datasets (with $1{,}000 < n < 9{,}000$). This validates that anchor graph learning has high efficiency for multi-view clustering.
- —For the large-scale cases, traditional approaches are not computationally suitable due to the high computational burden of $O(n^3)$. Moreover, we can observe that anchor graph-based approaches run more quickly than the anchor subspace-based approaches on most of the large-scale datasets.
- —It is worth that our proposed approach shows comparable or more reliable efficiency over the scalable approaches. Especially, it runs much faster over other anchor subspace-based approaches (e.g., SMVSC, FPMVS, and OMSC) on both general-scale and large-scale datasets.

In terms of the computational efficiency, our LTKMSC approach ranks at top three across the 11 multi-view clustering approaches. Consequently, we argue that it can achieve the best clustering performance on all benchmark datasets in comparison with the baseline approaches, while exhibiting highly competitive efficiency for both general-scale and large-scale datasets.

4.5 Sensitivity Analysis of Kernel Type

This subsection evaluates the impact of various kernel functions in the proposed approach. To this end, four baselines with several widely used kernels (i.e., Linear kernels, Cosine kernels, Gaussian kernels, and Polynomial kernels) are employed for sensitivity analysis. For clarity, these four baselines are abbreviated as *L-LTKMSC* (LTKMSC with Linear kernels), *C-LTKMSC* (LTKMSC with Cosine kernels), *G-LTKMSC* (LTKMSC with Gaussian kernels), and *P-LTKMSC* (LTKMSC with Polynomial kernels), respectively.

The comparison results w.r.t. NMI and Accuracy are presented in Tables 6 and 7, respectively. It can be observed that the Cosine-based and Gaussian-based baselines (i.e., C-LTKMSC and G-LTKMSC) perform substantially better than the baselines with Linear and Polynomial kernels (i.e., L-LTKMSC and P-LTKMSC). Especially, C-LTKMSC has obtained promising results on most of the

Table 6. The Clustering Performance (in Terms of NMI Metric) of Our LTKMSC Approach with Different Kernel Functions

Datasets	L-LTKMSC	C-LTKMSC	G-LTKMSC	P-LTKMSC
ORL	**97.94**	96.41	96.31	93.20
BBCSport	29.21	**99.34**	49.55	41.34
Notting-Hill	87.95	**97.81**	94.54	78.98
Caltech101-7	68.32	**75.56**	70.82	67.74
UCI Digits	84.26	**99.59**	99.25	79.77
NH-p4660	92.34	**95.95**	92.91	76.98
Caltech101-all	45.97	**86.63**	85.97	33.99
ALOI-100	12.58	91.70	**92.93**	12.68
Reuters-lee	20.43	**95.68**	86.91	21.02
NUSWIDEOBJ	12.04	19.16	**21.54**	9.22
AwA	8.05	93.46	**93.80**	7.90
Average Score	50.83	**86.48**	80.41	47.53

The best scores are highlighted in bold for each dataset.

Table 7. The Clustering Performance (in Terms of Accuracy) of Our LTKMSC Approach with Different Kernel Functions

Datasets	L-LTKMSC	C-LTKMSC	G-LTKMSC	P-LTKMSC
ORL	**97.50**	95.75	95.75	89.50
BBCSport	51.65	**99.82**	64.89	57.17
Notting-Hill	88.91	**99.09**	97.82	83.27
Caltech101-7	84.87	84.87	**85.62**	84.12
UCI Digits	83.3	**99.85**	99.70	80.35
NH-p4660	94.89	**98.20**	95.04	80.43
Caltech101-all	22.51	**55.64**	53.38	15.86
ALOI-100	4.89	79.27	**79.81**	4.95
Reuters-lee	41.53	**95.61**	92.12	36.91
NUSWIDEOBJ	13.09	20.02	**22.31**	11.07
AwA	7.73	84.05	**86.21**	6.95
Average Score	53.72	**82.92**	79.33	50.05

The best scores are highlighted in bold for each dataset.

benchmark datasets, which suggests that the Cosine kernel can be a suitable choice in a variety of situations.

5 Conclusion

This article develops a novel approach for large-scale MSC, which inherits the advantage of low-time consuming efficiency and highly competitive effectiveness. By comparing with other scalable approaches, our LTKMSC approaches jointly models the inter-view and intra-view awareness for anchor representation learning. By means of this mechanism, the proposed approach flexibly exploits the latent high-order consistency across various views, while robustly discovering the discriminative structure contained in individual view. Furthermore, our approach seeks to efficiently

recover the nonlinear structure embedded in kernel space. These advantageous points contribute to the superior performance of our proposed approach, which have been experimentally validated on multiple general-scale and large-scale datasets.

References

[1] Saba A. Al-Sayouri, Ekta Gujral, Danai Koutra, Evangelos E. Papalexakis, and Sarah S. Lam. 2018. t-PNE: tensor-based predictable node embeddings. In *Proceedings of the 2018 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)*. IEEE, 491–494.

[2] Xiaosha Cai, Dong Huang, Guang-Yu Zhang, and Chang-Dong Wang. 2023. Seeking commonness and inconsistencies: A jointly smoothed approach to multi-view subspace clustering. *Information Fusion* 91 (2023), 364–375.

[3] Xiaochun Cao, Changqing Zhang, Huazhu Fu, Si Liu, and Hua Zhang. 2015. Diversity-induced multi-view subspace clustering. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 586–594.

[4] Man-Sheng Chen, Ling Huang, Chang-Dong Wang, and Dong Huang. 2020. Multi-view clustering in latent embedding space. In *Proceedings of the 34th AAAI Conference on Artificial Intelligence*.

[5] Man-Sheng Chen, Chang-Dong Wang, Dong Huang, Jian-Huang Lai, and Philip S. Yu. 2022. Efficient orthogonal multi-view subspace clustering. In *Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, 127–135.

[6] Yongyong Chen, Xiaolin Xiao, Chong Peng, Guangming Lu, and Yicong Zhou. 2021. Low-rank tensor graph learning for multi-view subspace clustering. *IEEE Transactions on Circuits and Systems for Video Technology* 32, 1 (2021), 92–104.

[7] Yongyong Chen, Xiaolin Xiao, and Yicong Zhou. 2019. Jointly learning kernel representation tensor and affinity matrix for multi-view clustering. *IEEE Transactions on Multimedia 22* (2019), 1985–1997.

[8] Si-Guo Fang, Dong Huang, Xiao-Sha Cai, Chang-Dong Wang, Chaobo He, and Yong Tang. 2023. Efficient multi-view clustering via unified and discrete bipartite graph learning. *IEEE Transactions on Neural Networks and Learning Systems* 35 (2023), 11436–11447.

[9] Hongchang Gao, Feiping Nie, Xuelong Li, and Heng Huang. 2015. Multi-view subspace clustering. In *Proceedings of the IEEE International Conference on Computer Vision*, 4238–4246.

[10] Alexander Gorovits, Ekta Gujral, Evangelos E. Papalexakis, and Petko Bogdanov. 2018. LARC: Learning activity-regularized overlapping communities across time. In *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 1465–1474.

[11] Arthur Gretton, Olivier Bousquet, Alex Smola, and Bernhard Schölkopf. 2005. Measuring statistical dependence with Hilbert-Schmidt norms. In *Proceedings of the International Conference on Algorithmic Learning Theory*. Springer, 63–77.

[12] Wenrui Hu, Dacheng Tao, Wensheng Zhang, Yuan Xie, and Yehui Yang. 2016. The twist tensor nuclear norm for video completion. *IEEE Transactions on Neural Networks and Learning Systems* 28, 12 (2016), 2961–2973.

[13] Yongli Hu, Zuolong Song, Boyue Wang, Junbin Gao, Yanfeng Sun, and Baocai Yin. 2021. AKM3C: Adaptive k-multiple-means for multi-view clustering. *IEEE Transactions on Circuits and Systems for Video Technology* 31, 11 (2021), 4214–4226.

[14] Shudong Huang, Zhao Kang, Ivor W. Tsang, and Zenglin Xu. 2019. Auto-weighted multi-view clustering via kernelized graph learning. *Pattern Recognition* 88 (2019), 174–184.

[15] Vassilis N. Ioannidis, Ahmed S. Zamzam, Georgios B. Giannakis, and Nicholas D. Sidiropoulos. 2019. Coupled graphs and tensor factorization for recommender systems and community detection. *IEEE Transactions on Knowledge and Data Engineering* 33, 3 (2019), 909–920.

[16] Zhao Kang, Zhiping Lin, Xiaofeng Zhu, and Wenbo Xu. 2021. Structured graph learning for scalable subspace clustering: From single view to multiview. *IEEE Transactions on Cybernetics* 52 (2021), 8976–8986.

[17] Zhao Kang, Wangtao Zhou, Zhitong Zhao, Junming Shao, Meng Han, and Zenglin Xu. 2020. Large-scale multi-view subspace clustering in linear time. In *Proceedings of the 34th AAAI Conference on Artificial Intelligence, Vol.* 34, 4412–4419.

[18] Mengcheng Lan, Min Meng, Jun Yu, and Jigang Wu. 2021. Generalized multi-view collaborative subspace clustering. *IEEE Transactions on Circuits and Systems for Video Technology* 32 (2021), 3561–3574.

[19] Ruihuang Li, Changqing Zhang, Huazhu Fu, Xi Peng, Tianyi Zhou, and Qinghua Hu. 2019. Reciprocal multi-layer subspace learning for multi-view clustering. In *Proceedings of the IEEE International Conference on Computer Vision*, 8172–8180.

[20] Xuelong Li, Han Zhang, Rong Wang, and Feiping Nie. 2020. Multiview clustering: A scalable and parameter-free bipartite graph fusion method. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 44, 1 (2020), 330–344.

[21] Yeqing Li, Feiping Nie, Heng Huang, and Junzhou Huang. 2015. Large-scale multi-view spectral clustering via bipartite graph. In *Proceedings of the 29th AAAI Conference on Artificial Intelligence*, 2750–2756.

[22] Zhenglai Li, Chang Tang, Xinwang Liu, Xiao Zheng, Wei Zhang, and En Zhu. 2021. Consensus graph learning for multi-view clustering. *IEEE Transactions on Multimedia* 24 (2021), 2461–2472.

[23] Youwei Liang, Dong Huang, and Chang-Dong Wang. 2019. Consistency meets inconsistency: A unified graph learning framework for multi-view clustering. In *Proceedings of the IEEE International Conference on Data Mining*.
[24] You-Wei Liang, Dong Huang, Chang-Dong Wang, and Philip S. Yu. 2022. Multi-view graph learning by joint modeling of consistency and inconsistency. *IEEE Transactions on Neural Networks and Learning Systems* 35 (2022), 2848–2862.
[25] Suyuan Liu, Xinwang Liu, Siwei Wang, Xin Niu, and En Zhu. 2022. Fast incomplete multi-view clustering with view-independent anchors. *IEEE Transactions on Neural Networks and Learning Systems* 35 (2022), 7740–7751.
[26] Zhen Long, Ce Zhu, Jie Chen, Zihan Li, Yazhou Ren, and Yipeng Liu. 2023. Multi-view MERA subspace clustering. *IEEE Transactions on Multimedia* 5 (2023), 3102–3112.
[27] Shirui Luo, Changqing Zhang, Wei Zhang, and Xiaochun Cao. 2018. Consistent and specific multi-view subspace clustering. In *Proceedings of the 32nd AAAI Conference on Artificial Intelligence*.
[28] Di Ming and Chris Ding. 2019. Robust flexible feature selection via exclusive L21 regularization. In *Proceedings of the 28th International Joint Conference on Artificial Intelligence*, 3158–3164.
[29] Feiping Nie, Guohao Cai, and Xuelong Li. 2017. Multi-view clustering and semi-supervised classification with adaptive neighbours. In *Proceedings of the 31st AAAI Conference on Artificial Intelligence*, 2408–2414.
[30] Feiping Nie, Jing Li, and Xuelong Li. 2017. Self-weighted multiview clustering with multiple graphs. In *Proceedings of the 26th International Joint Conference on Artificial Intelligence*, 2564–2570.
[31] Qilun Luo, Ming Yang, Wen Li, and Mingqing Xiao. 2023. Hyper-Laplacian regularized multi-view clustering with exclusive L21 regularization and tensor log-determinant minimization approach. *ACM Transactions on Intelligent Systems and Technology* 14 (2023), 1–29.
[32] Fatemeh Sheikholeslami, Brian Baingana, Georgios B. Giannakis, and Nikolaos D. Sidiropoulos. 2016. Egonet tensor decomposition for community identification. In *Proceedings of the 2016 IEEE Global Conference on Signal and Information Processing (GlobalSIP)*. IEEE, 341–345.
[33] Fatemeh Sheikholeslami and G. Giannakis. 2018. Robust overlapping community detection via constrained egonet tensor decomposition. In *Proceedings of the 11th ACM International Conference on Web Search and Data Mining*.
[34] Mengjing Sun, Pei Zhang, Siwei Wang, Sihang Zhou, Wenxuan Tu, Xinwang Liu, En Zhu, and Changjian Wang. 2021. Scalable multi-view subspace clustering with unified anchors. In *Proceedings of the 29th ACM International Conference on Multimedia*, 3528–3536.
[35] Hao Wang, Yan Yang, and Bing Liu. 2019. GMC: Graph-based multi-view clustering. *IEEE Transactions on Knowledge and Data Engineering* 32 (2019), 1116–1129.
[36] Qianqian Wang, Zhiqiang Tao, Quanxue Gao, and Licheng Jiao. 2022. Multi-view subspace clustering via structured multi-pathway network. *IEEE Transactions on Neural Networks and Learning Systems* 35, 5 (2022), 7244–7250.
[37] Siwei Wang, Xinwang Liu, Suyuan Liu, Wenxuan Tu, and En Zhu. 2024. Scalable and structural multi-view graph clustering with adaptive anchor fusion. *IEEE Transactions on Image Processing* 33 (2024), 4627–4639.
[38] Siwei Wang, Xinwang Liu, Xinzhong Zhu, Pei Zhang, Yi Zhang, Feng Gao, and En Zhu. 2021. Fast parameter-free multi-view subspace clustering with consensus anchor guidance. *IEEE Transactions on Image Processing* 31 (2021), 556–568.
[39] Xiaobo Wang, Xiaojie Guo, Zhen Lei, Changqing Zhang, and Stan Z. Li. 2017. Exclusivity-consistency regularized multi-view subspace clustering. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 923–931.
[40] Jianlong Wu, Zhouchen Lin, and Hongbin Zha. 2019. Essential tensor learning for multi-view spectral clustering. *IEEE Transactions on Image Processing* 28, 12 (2019), 5910–5922.
[41] Wei Xia, Quanxue Gao, Qianqian Wang, Xinbo Gao, Chris Ding, and Dacheng Tao. 2022. Tensorized bipartite graph learning for multi-view clustering. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 45 (2022), 5187–5202.
[42] Jierui Xie, Stephen Kelley, and Boleslaw K. Szymanski. 2013. Overlapping community detection in networks: The state-of-the-art and comparative study. *ACM Computing Surveys* 45, 4 (2013), 1–35.
[43] Yuan Xie, Dacheng Tao, Wensheng Zhang, Yan Liu, Lei Zhang, and Yanyun Qu. 2018. On unifying multi-view self-representations for clustering by tensor multi-rank minimization. *International Journal of Computer Vision* 126, 11 (2018), 1157–1179.
[44] Yuan Xie, Wensheng Zhang, Yanyun Qu, Longquan Dai, and Dacheng Tao. 2018. Hyper-Laplacian regularized multilinear multiview self-representations for clustering and semisupervised learning. *IEEE Transactions on Cybernetics* 50, 2 (2018), 572–586.
[45] Ben Yang, Xuetao Zhang, Zhiping Lin, Feiping Nie, Badong Chen, and Fei Wang. 2022. Efficient and robust multi-view clustering with anchor graph regularization. *IEEE Transactions on Circuits and Systems for Video Technology* 32 (2022), 6200–6213.
[46] Kun Zhan, Chaoxi Niu, Changlu Chen, Feiping Nie, Changqing Zhang, and Yi Yang. 2018. Graph structure fusion for multiview clustering. *IEEE Transactions on Knowledge and Data Engineering* 31 (2018), 1984–1993.
[47] Changqing Zhang, Huazhu Fu, Qinghua Hu, Xiaochun Cao, Yuan Xie, Dacheng Tao, and Dong Xu. 2018. Generalized latent multi-view subspace clustering. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 42 (2018), 86–99.

[48] Changqing Zhang, Qinghua Hu, Huazhu Fu, Pengfei Zhu, and Xiaochun Cao. 2017. Latent multi-view subspace clustering. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 4279–4287.
[49] Guang-Yu Zhang, Dong Huang, and Chang-Dong Wang. 2022. Facilitated low-rank multi-view subspace clustering. *Knowledge-Based Systems* (2022), 110141.
[50] Guang-Yu Zhang, Chang-Dong Wang, Dong Huang, Wei-Shi Zheng, and Yu-Ren Zhou. 2018. TW-Co-k-means: Two-level weighted collaborative k-means for multi-view clustering. *Knowledge-Based Systems* 150 (2018), 127–138.
[51] Yongshan Zhang, Jia Wu, Chuan Zhou, Zhihua Cai, Jian Yang, and Philip S. Yu. 2019. Multi-view fusion with extreme learning machine for clustering. *ACM Transactions on Intelligent Systems and Technology* 10, 5 (2019), 1–23.
[52] Zheng Zhang, Li Liu, Fumin Shen, Heng Tao Shen, and Ling Shao. 2018. Binary multi-view clustering. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 41, 7 (2018), 1774–1782.

Received 20 March 2023; revised 3 February 2025; accepted 16 April 2025

Dual-level Facilitated Multi-view Contrastive Graph Clustering

Guangyu Zhang
South China Agricultural University
Guangzhou, China
guangyuzhg@foxmail.com

Zi-Ying Li
South China Agricultural University
Guangzhou, China
ziying.li36@foxmail.com

Hai-Yan Wang*
South China Agricultural University
Guangzhou, China
cshywang@scau.edu.cn

Dong Huang
South China Agricultural University
Guangzhou, China
huangdonghere@gmail.com

Changdong Wang
Sun Yat-sen University
Guangzhou, China
changdongwang@hotmail.com

Yang Liu
Sun Yat-sen University
Guangzhou, China
liuy856@mail.sysu.edu.cn

2025 IEEE 31th International Conference on Parallel and Distributed Systems (ICPADS) | 979-8-3315-4901-5/25/$31.00 ©2025 IEEE | DOI: 10.1109/ICPADS67057.2025.11323229

***Abstract*—Multi-view attributed graph clustering (MAGC) has recently experienced impressive attention in the graph exploration literature. Although several excellent achievements have been made, previous MVAGC approaches merely consider the homogeneous information across different views, easily resulting in the compromised results when faced with the heterogeneous graph scenarios. Further, many of them rely on the static neighborhood connection from original attributed graphs, which ignores the dynamical structural relationship for enhancing cross-view contrastive learning. To deal with these drawbacks, this paper derives a Dual-level Facilitated Multi-view Contrastive Graph Clustering (DF-MCGC) approach. Specifically, we design a hybrid graph filter by considering the homogeneity hidden in individual view. Further, the view-consistent topology invariant matrix is derived to exploit the topological relationship among different embedded representations. This design progressively helps to construct the cluster-wise sample pairs for cross-view contrastive learning. Especially, the overall network is incorporated with a dual-level self-supervised paradigm, among which the self-supervised signals can be efficiently facilitated in a mutually enhanced manner. Experiments on heterogeneous benchmarks have confirmed the advantages of our DF-MCGC approach in comparison with the advanced competitors.**

***Index Terms*—Multi-view Attributed Graph clustering, Cross-view Contrastive Learning, Dual-level Self-Supervised Strategy.**

I. Introduction

The technological achievements bring about the substantial emergence of multi-view graph data in various scenarios, ranging from community detection to biological data mining. Specifically, the multi-view graph data are simultaneously obtained from diverse sources, where each source may contain two types of descriptions, namely the attributed features and corresponding affinity graph. As an instance, each ACM paper generally has own specific keywords and two kinds of topological graphs (like co-author and co-term networks) to reflect its academic relationship. Unlike the traditional settings, the multi-view graph data presents a challenging research hotspot yet great scientific opportunity in unsupervised learning scenario. Inspired by this, some data analysts frequently employ the clustering technology to exploit the meaningful dual-level information (i.e., both the attributed features and structural graph) from multi-view graph datasets.

Lately several endeavors have been dedicated to the study of multi-view attributed graph clustering (MAGC). In accordance with the employed strategies, the current mainstream can roughly split into two categories, namely the matrix factorization-based MAGC approaches [1]–[3] and deep learning-based MAGC approaches (also abbreviated as deep MAGC approaches) [4]–[6]. In particular, the matrix factorization-based branch typically focuses on the shadow semantic information from node attribute values and multiview topological graphs. However, such linear factorization framework may lead to the performance degradation when confronted with the complex scenarios (like the heterogeneous graph clustering task). With the prosperous development of neural networks, the deep MAGC literature has drawn increasing attention over the past years. In terms of the heterogeneous perspectives, data researchers proposed a series of deep MAGC representatives, such as MAGCN [5] and VGMGC [7]. By breaking through the conventional practice, these approaches help to explore the underlying semantic patterns hidden in multiple views, which provides a favourable tool for graph clustering analysis from deeper perspective.

Contrastive learning has been a emerging research trends in self-supervised learning. Up to date, merely a few studies have leveraged the useful contrastive learning method for enhancing the multi-view attributed graph clustering. As an early attempt, Pan et al. [1] combined the matrix factorization MAGC approach with subspace learning framework, among which a contrastive graph loss is derived to generate a clustering-friendly representation across heterogeneous views. In additional, Chen

et al. [8] recently presented a contrastive-oriented deep MAGC approach (namely CMAGC) with adaptive encoders. Specifically, this approach attempted to exploit the inherent topological relationships from both inter-view and intra-view viewpoints. Benefiting from the prominent performance of elegant network paradigm, deep MAGC approaches have made remarkable success in many challenging scientific problems (especially for complex graph network field). While verified to be effective, two essential limitations still require to be considered for future explorations. (i) How to go beyond the homophily graph assumption, enabling the existing deep MAGC approaches for heterophilous issue. (ii) How to discover the dynamical neighborhood connections instead of the static topological structure assumption?

To mitigate these limitations, a novel contrastive-oriented deep MAGC approach, termed Dual-level Facilitated Multi-view Contrastive Graph Clustering (DF-MCGC) is developed in this research. By employing a weighting-shared network as foundation, the proposed clustering network consists of three reciprocally enhanced components. In the first component, we utilize a hybrid graph filter to adaptively model the graph heterogeneity within each view, in especially the low-pass and high-pass components are jointly considered for enhancing representation learning. In the second component, a view-consistent topology-invariant matrix is formulated to guide the generation of positive and negative sample set, by dramatically depicting the latent clustering distribution among different feature representations. In the last component, we further incorporate the dual-level self-supervision mechanism into cross-view contrastive learning, which provides an effective guidance for facilitating the overall network robustness. The flowchart of our proposed network as illustrated in Fig. 1, while the corresponding contributions are summarized as follows:

- This paper derives a new deep MAGC approach named DF-MCGC, which inherits the favorable benefits of homophily-driven hybrid graph filter, view-consistent structural preserving and cross-view contrastive learning with dual-level self-supervision mechanism.
- This paper presents a newly designed multi-view graph network characterized by a streamlined architecture and impressive effectiveness.
- Substantial experiments on both homogeneous and heterogeneous benchmarks demonstrate the effectiveness and advantages of of our DF-MCGC approach.

II. Related Works

The brief overview of related research will be introduced in this section, primarily concerning with multi-view attributed graph clustering (MAGC) and contrastive learning techniques.

Attribute graphs extend traditional graph structures by augmenting nodes with feature representations, enabling richer modeling of real-world relational data. Unlike conventional graphs that only encode topological relationships, attribute graphs integrate both structural and semantic information, which enhances their effectiveness for tasks such as node classification, link prediction, and clustering. Early work in attribute graph analysis relied on spectral methods and matrix factorization to jointly optimize structural and feature-based similarities [9]–[11]. Later, graph embedding techniques such as DeepWalk [12] and node2vec [13] introduced unsupervised representation learning by combining random walks with skip-gram models. However, these methods often separate topology and attribute, limiting their ability to capture complex interactions. The advent of Graph Neural Networks (GNNs) revolutionized attribute graph learning by leveraging message-passing mechanisms to aggregate neighborhood information. Frameworks like GCN [14] and GAT [15] achieved state-of-the-art performance by propagating and transforming node features across the graph structure. Subsequent advances in attribute graph clustering further improved partitioning quality through techniques like self-supervised learning. These methods primarily operate under a single-view assumption, neglecting the multi-faceted nature of real-world data gap addressed by multi-view attributed graph clustering.

Many real-world systems (e.g., social networks, biomolecular graphs) exhibit multi-modal data, where nodes are described by multiple feature sets or interaction types. Multi-view attributed graphs (MVAG) [16], [17] formalize this by associating each node with heterogeneous feature spaces, enabling more robust clustering through complementary information. Existing MVAG clustering methods fall into three broad categories: (i) Spectral and Matrix Factorization Approaches: Early techniques extended spectral clustering [18] by fusing multiple affinity matrices, often via co-regularization [19] or consensus learning [20]. While effective for small graphs, these methods struggle with scalability and nonlinear feature interactions. (ii) Deep Embedding and Autoencoder-Based Methods: Models like MGAE [21] and [22] employ deep autoencoders to learn a unified latent representation by reconstructing both graph structure and multi-view features. These approaches improve nonlinear modeling but often lack explicit mechanisms for cross-view consistency. (iii) Graph Neural Network (GNN) Methods Recent GNN-based frameworks, such as MVGRL [23] and DMGCN [24], dynamically aggregate multi-view features through attention or gating mechanisms. For instance, contrastive learning techniques align views by maximizing mutual information, while adaptive weighting strategies mitigate redundancy.

III. The Proposed Approach

The proposed DF-MCGC approach will be described in this section, where three essential components are seamlessly leveraged, namely homophily-driven hybrid graph

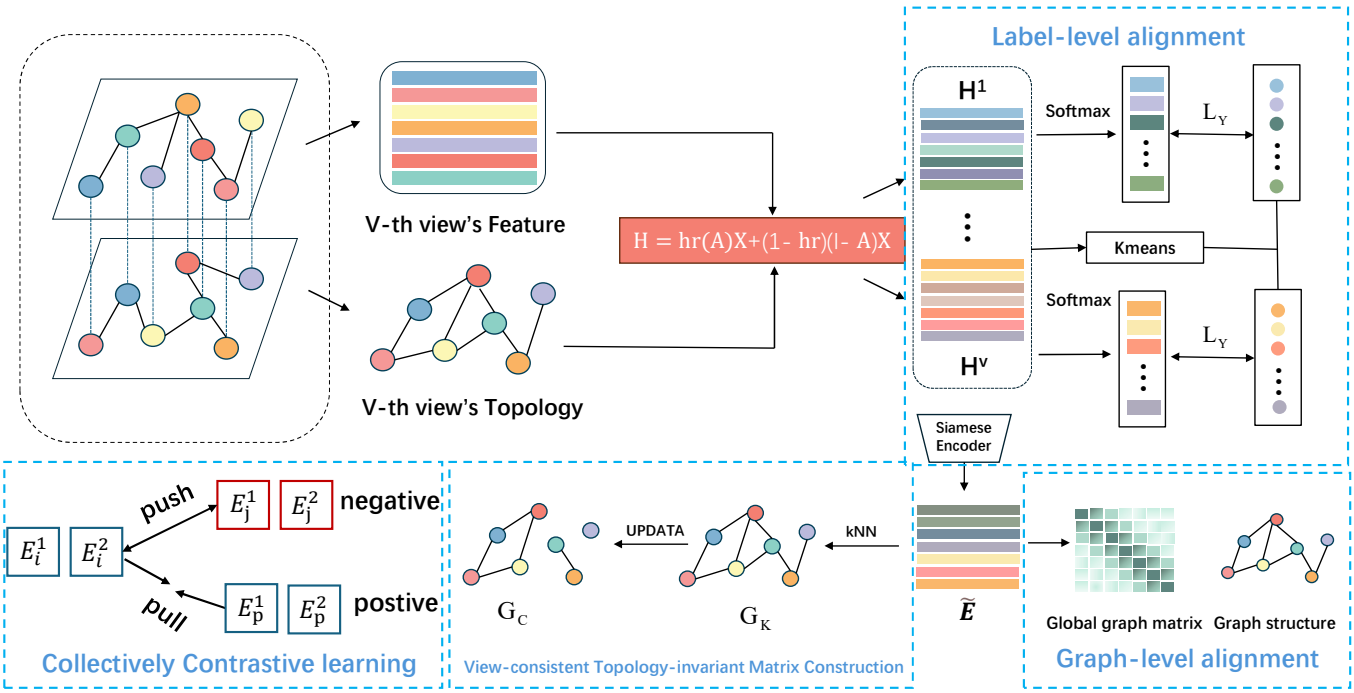


Fig. 1: The flowchart of our DF-MCGC approach.

augmentation, view-consistent topology-invariant matrix construction and collectively contrastive learning with dual-level self-supervised paradigm.

A. Notations and primary objective

Consider a input graph data $\mathcal{G} = (\mathrm{X}, \{\mathrm{A}^{(v)}\}_{v=1}^{V})$ with V views, in which $\mathrm{X} \in \mathbb{R}^{n \times d}$ represents the node features shared across multiple views, while $\mathrm{A}^{(v)} \in \mathbb{R}^{n \times n}$ represents the v-th view adjacent matrix encoding the pairwise relationships between node samples. Especially, the adjacency matrix $\mathrm{A}^{(v)}$ of each view is normalized as $\tilde{\mathrm{A}}^{(v)} = \left(\mathrm{D}^{(v)}\right)^{-1} \mathrm{A}^{(v)}$, where $\mathrm{D}^{(v)}$ denotes the corresponding degree matrix with diagonal element $\mathrm{D}_{ii}^{(v)} = \sum_j a_{ij}^{v}$. As a result, the normalized graph Laplacian matrix could be obtained by $\tilde{\mathrm{L}}^{(v)} = \mathrm{I} - \tilde{\mathrm{A}}^{(v)}$, among which I being the identity matrix. The main objective of multi-view attributed graph clustering relies to partition the node samples into disconnected groups by effectively exploiting both semantic and structural cues from multiple views.

B. Homophily-driven hybrid graph augmentation

Graph augmentation serves as a promising technique in contrastive learning. Inherited from this motivation, our first component aims to generate reliable graph augmentations for subsequent representation learning. Remarkably, some studies have revealed that the smoothed attributed graphs naturally maintain both low-frequency and high-frequency spectral signals. Motivated by this, a hybrid filter could be formulated in what follows.

$$\mathrm{H}^{(v)} = \mu(\tilde{\mathrm{A}}^{(v)})^{r}\mathrm{X} + (1-\mu)(\mathrm{I} - \tilde{\mathrm{A}}^{(v)})^{r}\mathrm{X}, \forall v. \quad (1)$$

Notice that $\tilde{\mathrm{A}}^{(v)}$ and $(\mathrm{I} - \tilde{\mathrm{A}}^{(v)})$ represent the low-pass and high-pass components in hybrid graph filter. Additionally, r is the filtering order, while μ is a balancing parameter that measures the impact of of low-pass (i.e., neighbor connections) and high-pass components (i.e., neighbor disconnections). In multi-view graph clustering scenario, empirical evidences reveal that each view may exhibit the specific graph structural patterns (from homophily to heterogeneity). Thereafter, how to well model the homophily across multiple graphs (i.e., views) still remains an under-explored challenge. In view of this, we

devise a pseudo-label guided homophily metric, which could be defined in what follows.

$$\mathrm{hr}^{(v)} = \frac{\delta(\tilde{\mathrm{A}}^{(v)} \odot \mathrm{YY}^T - \mathrm{I})}{\delta(\tilde{\mathrm{A}}^{(v)} - \mathrm{I})}, \forall v. \tag{2}$$

In particular, metric $\mathrm{hr}^{(v)}$ can help to detect the homophily ratio of view-specific graph $\tilde{\mathrm{A}}^{(v)}$, where δ is a summation operator and $\odot$ denotes the Hadamard product. Besides, $\mathrm{Y} \in \{0,1\}^{n\times c}$ is the the pseudo-label guided matrix of the current iteration, by applying k-means algorithm on the view-consistent embedding $\hat{\mathrm{E}}$. In terms of the homophily ratio perspective, we further propose a homophily-driven hybrid graph filter for each view, which consequently results in the following formulation:

$$\mathrm{H}^{(v)} = \mathrm{hr}^{(v)}(\tilde{\mathrm{A}}^{(v)})^r\mathrm{X} + (1 - \mathrm{hr}^{(v)})(\mathrm{I} - \tilde{\mathrm{A}}^{(v)})^r\mathrm{X}, \forall v. \tag{3}$$

This design adaptively balances the contributions of low-pass and high-pass components based on the homophily ratio of each view-specific graph $\tilde{\mathrm{A}}^{(v)}$. Subsequently, the twin encoders are employed to capture the latent semantics embedded within multi-view graph augmentations. The corresponding formulation is given as below.

$$\mathrm{E}^{(v)} = SiameseEncoder(\mathrm{H}^{(v)}), \forall v, \tag{4}$$

in which $\mathrm{E}^{(v)}$ denotes the node embeddings of the v-th view. In line with this, the semantic information hidden in multiple views can be further considered as follows:

$$\tilde{\mathrm{E}} = \sum_{v=1}^{V} \omega^{(v)}\mathrm{E}^{(v)}, \tag{5}$$

in which

$$\omega^{(v)} = \left(\frac{eva^{(v)}}{\max(eva^{(1)}, eva^{(2)}, \ldots, eva^{(V)})}\right)^{\rho}.$$

It is worthy that, $\tilde{\mathrm{E}}$ stands for the view-consistent embedding and $\omega^{(v)}$ represents the v-th view weight. Besides, ρ is a parameter that controls the distribution of different weights, while $eva^{(v)}$ is the pre-defined function that measures the similarity between $\mathrm{E}^{(v)}$ and $\tilde{\mathrm{E}}$. By this means, a series of reliable node embeddings (including multi-view embeddings and view-consistent embedding) could be obtained for post-processing components.

C. View-consistent Topology-invariant Matrix Construction

Recently several studies encourage to select the potential sample pairs for strengthening cross-view contrastive learning. Following this lead, our second component strives to derive a topology-invariant matrix grounded in Edge Betweenness Centrality (EBC) principle. Once the view-consistent embeddings are obtained, this matrix is designed to capture the intrinsic clustering structure while effectively filtering out redundant edges.

It is well recognized that neighborhood nodes with similar features tend to belong to the same cluster. Following this principle, this work first build a fully-connected similarity graph G, where pairwise similarity between consistent embeddings $\tilde{\mathrm{E}}$ is measured by the Euclidean distance. Subsequently, we select the top-k most comparable pairs associated with per embedding, these pairs are connected to yield a k-nearest neighbor (kNN) graph, denoted as G_k. Nevertheless, certain intra-cluster connections continue to exist which may affect the effectiveness. To remove redundant edges, we apply edge betweenness centrality (EBC) theory to enhance the clustering structure of the k-nearest neighbor (kNN) graph. Next, we present the formal definition of edge betweenness centrality:

Definition 1: Edge Betweenness Centrality (EBC). *To quantify the importance of an edge, the EBC counts the number of shortest paths through that edge in an undirected graph. Specifically, the shortest path refers to the minimum distance between any two nodes. Formally, EBC can be defined as follows:*

$$\mathrm{c_B}(e) = \sum_{g,l\in V} \frac{\sigma(g,l|e)}{\sigma(g,l)}, \tag{6}$$

in which $\sigma(g,l)$ denotes the total count of shortest paths between node g and l, and $\sigma(g,l|e)$ indicates the number of such shortest paths that travel across the edge e.

When it comes to the constructed kNN graph G_k, those edges with high EBC values exhibit statistically tendency to establish connections between disconnected groups. According to the principle described above, the edge betweenness centrality score(EBC) of each edge is determined, and then we can filter out the t edges with the top EBC values for further analysis, in which $t = \sigma \times |\epsilon|$, $|\epsilon|$ represents the number of edges, and σ denotes the probability of edge removal. By means of the removal process, a clustering-guided k-NN graph G_c across multiple views (referred to view-consistent topology-invariant matrix) could be achieved. By eliminating the redundant edges from onstructed kNN graph, the matrix G_c could reveal the clear clustering structure that primarily composed of intra-cluster edges. Consequently, we could employ matrix G_c as a topology-invariant rule to serve the subsequent collectively contrastive learning process. The relevant details will be explained in the following subsection.

D. Collectively Contrastive learning with Dual-level Self-supervised Paradigm

Contrastive learning and self-supervised alignment have emerged as two pivotal techniques in deep clustering research. Diverging from the standard practice, the third component targets at jointly performing cross-view contrastive learning and dual-level self-supervised alignments. With the topology-invariant matrix G_c and node embeddings as input, we firstly derive into the details of cross-view contrastive learning. Especially, by considering the $\mathrm{E}_i^{(v)}$ as an instance, the corresponding positive sample pairs come from two disjoint sources. (i) The v-th view

adjacent neighbors of $\mathrm{E}_i^{(v)}$, e.g., $\{\mathrm{E}_p^{(v)}|\mathrm{E}_p^{(v)} \in \Delta_i\}$. (ii) The adjacent neighbors of $\mathrm{E}_i^{(v)}$ in the view-consistent space, e.g., $\{\tilde{\mathrm{E}}_p|\tilde{\mathrm{E}}_p \in \Delta_i\}$. By this means, the cardinality of positive sample set is $2|\Delta_i|$, where Δ_i means the adjacent neighborhood of the i-th data instance within topology-invariant matrix (i.e., G_c). Consequently, the cross-view contrastive loss of $\mathrm{E}_i^{(v)}$ is formally given by:

$$\ell(\mathrm{E}_i^{(v)}) = \tag{7}$$
$$-\log \frac{\frac{1}{2|\Delta_i|}\left(\sum_{\mathrm{E}_p^{(v)}\in\Delta_i} e^{\theta(\mathrm{E}_i^{(v)},\mathrm{E}_p^{(v)})/\tau} + \sum_{\tilde{\mathrm{E}}_p\in\Delta_i} e^{\theta(\mathrm{E}_i^{(v)},\tilde{\mathrm{E}}_p)/\tau}\right)}{\sum_{j\neq i} e^{\theta(\mathrm{E}_i^{(v)},\mathrm{E}_p^{(v)})/\tau} + \sum_{j\neq i} e^{\theta(\mathrm{E}_i^{(v)},\tilde{\mathrm{E}}_p)/\tau}},$$

in which $\theta(\cdot)$ represents the pairwise similarity that measured by cosine distance, while the parameter τ serves as a temperature scaling parameter. Combining these components, the total loss function for cross-view contrastive learning can be expressed as follows:

$$\mathcal{L}_C = \frac{\sum_{i=1}^{n}\sum_{v=1}^{V}\ell(\mathrm{E}_i^{(v)})}{nV}. \tag{8}$$

Moving forward, we further go into the specifics of dual-level self-supervised paradigm. In the first phase, k-means algorithm is employed to uncover the latent pseudo-label information. By taking the view-specific embedding $\mathrm{E}^{(v)}$ as input, the corresponding expression is given below.

$$\mathrm{Y}^{(v)} = kmeans(\mathrm{E}^{(v)}, c), \forall v, \tag{9}$$

where c represents the clustering number and $\mathrm{Y}^{(v)}$ represents the v-th view pseudo label. At the same time, the softmax operation is utilized to capture the semantic information hidden within each view, which leads to the following formulation.

$$\Pi^{(v)} = Softmax(\mathrm{E}^{(v)}), \forall v \tag{10}$$

in which $\Pi^{(v)}$ denotes the v-th view semantic label. In terms of the label alignment perspective, we can achieve the view-specific self-supervised loss as follows.

$$L_{\mathrm{Y}} = -\frac{1}{nV}\sum_{v=1}^{V}\sum_{i=1}^{n}\Big\{\log(\mathrm{Y}_i{}^{(v)}\Pi_i{}^{(v)}) - (1-\mathrm{Y}_i{}^{(v)})\log(1-\Pi_i{}^{(v)})\Big\}. \tag{11}$$

This formulation could be viewed as a label-level cross-entropy loss, which encourages to promote the alignment between pseudo label and corresponding semantic information. Apart from the label-level viewpoint, we also consider the graph-level alignment for enhancing the network self-training. To be specific, the seismometer scaler is utilized to resize the view-consistent embedding $\tilde{E}$, and then the global graph matrix S could be defined in what follows.

$$\mathrm{S} = \frac{\tilde{\mathrm{E}}\tilde{\mathrm{E}}^T}{\|\tilde{\mathrm{E}}\|_2^2}. \tag{12}$$

TABLE I: The summary of four datasets adopted in our experiments.

Dataset	Type	Samples	Features	Views	Clusters	hr
Texas	Graph	183	1703	2	5	0.09 0.09
Chameleon	Graph	2277	2325	2	5	0.23 0.23
ACM	Graph	3025	1803	2	3	0.0 0.0
Freebase	Graph	3492	3762	3	4	0.99 1 0.99

Similarity, self-supervised loss could be extended to the graph-level alignment, which leads to the following formulation.

$$L_G = -\frac{1}{n^2}\sum_{v=1}^{V}\sum_{i,j=1}^{n}\left(\tilde{\mathrm{A}}_{ij}^{(v)}\log(\mathrm{S}_{ij}) + (1-\tilde{\mathrm{A}}_{ij}^{(v)})\log(1-\mathrm{S}_{ij})\right). \tag{13}$$

The above formulation could be interpreted as a graph-level cross entropy loss, which enables to align the view-specific graph distributions with their global graph structure. Notably, this dual-level design gives a novel insight into the fields of multi-view self-supervised learning.

More concretely, our proposed DF-MCGC approach optimizes three objectives into one unified paradigm. Accordingly, the general model is composed of the dual-level alignment objective $\mathcal{L}_D = \mathcal{L}_Y + \mathcal{L}_G$ and the contrastive loss $\mathcal{L}_C$:

$$\mathcal{L}_S = \mathcal{L}_D + \lambda\mathcal{L}_C, \tag{14}$$

Here the hyper-parameter λ is introduced to control the importance of two losses. Specifically, the model parameters are optimized by using the Adam algorithm during the training iterations.

E. Experimental Setup

The comprehensive experiments are executed on four widely-used standard datasets with different homophily degree. These datasets are collected from diverse practical scenarios (from homogeneous graphs to heterogeneous graphs), including **Texas**, **Chameleon**, **Synthetic ACM** and **Freebase**. To be specific, more details of these datasets could be found in Table I.

In order to verify the superiority and effectiveness of proposed DF-MCGC approach, several popular baselines are adopted for empirical comparison. (i) **k-means** is the best well-known clustering algorithm during the last decades. (ii)O^2**MAC** [4] is a basic model which aims to extract the shared representation from the most informative graph view, among which the shared representation is used to reconstruct the informative representation of all views. (iii) **MAGC** [25] and **MCGC** [1] represent two state-of-art graph approaches that utilize low-pass filter to achieve the consensus graphs. (iv) **DLGR** [26] leverages dual label guidance to refine the graph structure, and meanwhile improving the clustering robustness. (v)

TABLE II: The aggregated clustering metrics on four benchmark datasets.

Dataset	Metric	k-means	O^2MAC	MAGC	MCGC	DLGR	DCRN	AHGFC	Ours
Texas	ACC	38.8	46.7	54.3	51.9	54.3	52.8	[70.3]	**73.2**
	NMI	23.7	8.7	5.4	12.7	32.6	6.9	[39.3]	**43.5**
	ARI	14.2	14.6	1.1	12.9	[26.0]	1.7	45.3	**49.0**
	F1	29.1	29.1	19.8	32.5	[46.8]	22.9	45.7	**49.9**
Chameleon	ACC	39.8	33.5	29.2	30.0	41.1	24.1	[43.2]	**44.6**
	NMI	18.0	12.3	10.8	9.5	19.5	0.0	**21.5**	[20.3]
	ARI	15.3	8.9	3.3	5.9	16.0	0.0	[16.2]	**16.6**
	F1	34.3	28.6	24.3	19.1	37.7	17.3	[41.6]	**42.9**
ACM	ACC	37.4	55.0	37.1	63.0	[84.8]	84.0	83.8	**85.2**
	NMI	15.2	25.0	4.2	49.8	[55.1]	45.7	54.3	**56.4**
	ARI	10.3	24.7	9.2	42.9	**60.7**	57.3	59.3	[60.5]
	F1	37.1	54.6	35.5	53.5	82.5	83.3	[84.7]	**85.8**
Freebase	ACC	44.4	15.3	53.6	51.2	[55.1]	54.3	44.3	**62.5**
	NMI	0.0	14.1	10.9	13.9	[17.6]	12.7	0.0	**18.7**
	ARI	0.0	[13.4]	9.6	-0.7	10.0	11.2	0.0	**17.8**
	F1	20.6	17.5	[40.2]	36.5	39.6	37.2	20.6	**45.6**

DCRN [27] is a graph clustering approach by means of reducing information correlation. (vi) **AHGFC** [28] makes the attempt to alternately facilitates clustering and data interpolation in a unified framework.

Following the previous works [29]–[32], four commonly used metrics, i.e., normalized mutual information (NMI) [33], adjusted rand index (ARI) [34], accuracy (ACC) [35], and F1-score (F1) [36], are utilized to assess the clustering results.

F. Comparison experiments

The quantitative results of our DF-MCGC approach against seven baselines are presented in Table 2. Notably, the highest scores are highlighted in bold font to emphasize the best results, while the second-best scores are indicated within square brackets for clarity. According to these experiential results, we can acquire following two observations. With regard to the homogeneous graph cases (i.e., Freebase dataset), it is obvious that DF-MCGC approach outperforms other deep graph-based baselines in most of the comparisons. Especially, by means of four metrics (i.e., ACC, NMI, ARI and F1-score), our proposed approach suppresses the best competitor by 7.4%, 1.1%, 4.4% and 5.4%, respectively. With regard to the heterogeneous graph cases, our proposed approach has achieved stable yet promising performance, particularly for Texas and Chameleon dataset. In general, these analysis results highlight the efficacy of the proposed approach on both homogeneous and heterogeneous graph scenarios.

Apart from the quantitative perspective, we take the Chameleon dataset as an example, further providing the visualization evaluation by t-SNE algorithm. As shown in Fig. 2, our DF-MCGC method visually demonstrates distinct cluster separation and enhanced intra-cluster compactness, underscoring its superior advantage in capturing the underlying semantic structures within multi-view graph data.

G. Sensitivity Analysis

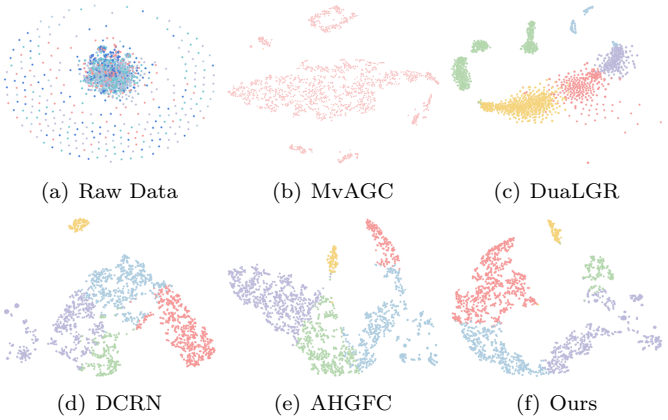
The sensitivity experiments are conducted to investigate the effect of hyper-parameters r and λ. Note that whenever a hyper-parameter is analyzed, the other hyper-parameters are set to static values. As can be seen from the Fig. 3, clustering performance exhibits relatively stable across different r values, while being more sensitive to the choice of λ, especially on the Texas dataset. As for Chameleon dataset, our proposed approach achieves best performance when $r = 10$ and $\lambda = 0.01$. This suggests that the moderate settings of hyper-parameters r and λ always lead to the promising clustering results.

H. Ablation Analysis

The efficacy of three essential components are assessed in this subsection. Particularly, ablation experiments conducted on three benchmarks are reported in Table 3, where "(w/o) HHF", "(w/o) TMC" and "(w/o) DSP" denote Homophily-oriented Hybrid Filter, Topology-invariant Matrix Construction and Dual-level Self-supervised Paradigm are removed, respectively. Especially, the absence of any individual component invariably may easily result in the statistical degradation. This has proved the necessity of these three components for promoting clustering effectiveness and scalability.

IV. Conclusion

This paper proposes a novel approach for multi-view contrastive graph clustering, which consists of three advantageous components for promising performance and effectiveness. The first component employs a hybrid graph filter to generate a collection of homophily-oriented graph augmentations. The second component constructs a view-consistent topology-invariant matrix based on edge betweenness centrality (EBC) theory, aiming to guide the formation of positive sample pairs. The third component introduces a collectively contrastive loss with dual-level self-supervised alignment, in particular the label-level and graph-level cross-entropy losses are incorporated. These



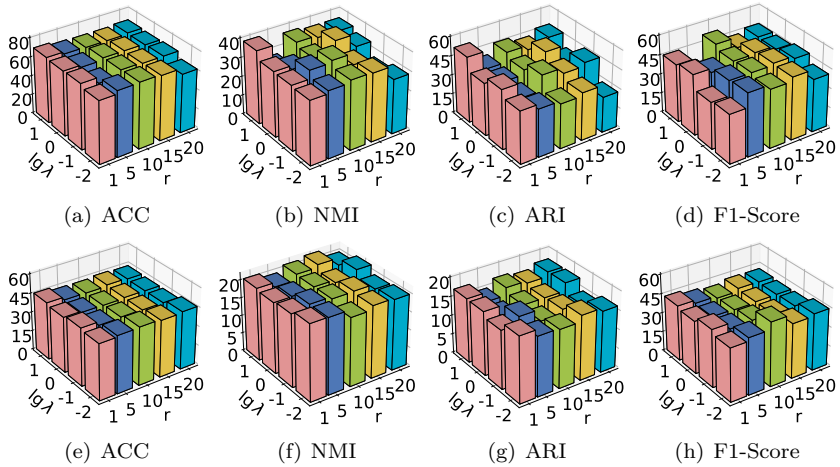
(a) Raw Data (b) MvAGC (c) DuaLGR

(d) DCRN (e) AHGFC (f) Ours

Fig. 2: The t-SNE plots comparing five methods and the raw data on the Chameleon dataset.

TABLE III: Evaluation of the individual contributions of our three components.

Dataset	Model	ACC(%)	NMI(%)	ARI(%)	F1(%)
Texas	(w/o) HHF	65.5	31.9	33.8	48.9
	(w/o) TMC	65.6	28.6	34.8	44.5
	(w/o) DSP	72.3	41.4	45.1	48.0
	Ours	**73.2**	**43.5**	**49.0**	**49.9**
Chameleon	(w/o) HHF	**45.5**	19.6	16.1	44.4
	(w/o) TMC	45.4	19.3	15.2	**45.0**
	(w/o) DSP	43.8	19.5	15.3	41.7
	Ours	44.6	**20.3**	**16.6**	42.9
Freebase	(w/o) HHF	60.1	18.0	16.8	44.3
	(w/o) TMC	62.0	18.0	16.6	43.0
	(w/o) DSP	61.8	18.3	15.3	41.7
	Ours	**62.5**	**18.7**	**17.8**	**45.6**



(a) ACC (b) NMI (c) ARI (d) F1-Score

(e) ACC (f) NMI (g) ARI (h) F1-Score

Fig. 3: The sensitivity analysis of clustering results of DF-MCGC w.r.t. λ and r on (a-d) Texas and (e-f) Chameleon datasets.

mutually beneficial components ensure the efficacy as well as generalizability of our clustering network, which have been validated on both homogeneous and heterogeneous secnarios.

Acknowledgment

This work was partially supported by the National Natural Science Foundation of China (NSFC) under Grants 62206099 and 62276277, the Guangdong Basic and Applied Basic Research Foundation under Grant

2022B1515120059, the Science and Technology Program of Guangzhou, China under Grant 2024A04J4451, as well as the Guangdong Provincial Key Laboratory of Intellectual Property and Big Data under Grant 2018B030322016.

References

[1] E. Pan and Z. Kang, "Multi-view contrastive graph clustering," *Advances in Neural Information Processing Systems*, vol. 34, pp. 2148–2159, 2021.

[2] Z. Lin, Z. Kang, L. Zhang, and L. Tian, "Multi-view attributed graph clustering," *IEEE Transactions on Knowledge and Data Engineering*, vol. 35, no. 2, pp. 1872–1880, 2021.

[3] C. Fettal, L. Labiod, and M. Nadif, "Simultaneous linear multi-view attributed graph representation learning and clustering," in *Proceedings of the Sixteenth ACM International Conference on Web Search and Data Mining*, 2023, pp. 303–311.

[4] S. Fan, X. Wang, C. Shi, E. Lu, K. Lin, and B. Wang, "One2multi graph autoencoder for multi-view graph clustering," in *Proceedings of the Web Conference 2020*, 2020, pp. 3070–3076.

[5] J. Cheng, Q. Wang, Z. Tao, D. Xie, and Q. Gao, "Multi-view attribute graph convolution networks for clustering," in *Proceedings of the Twenty-ninth International Conference on International Joint Conferences on Artificial Intelligence*, 2021, pp. 2973–2979.

[6] J.-Q. Lin, M.-S. Chen, X.-R. Zhu, C.-D. Wang, and H. Zhang, "Dual information enhanced multiview attributed graph clustering," *IEEE Transactions on Neural Networks and Learning Systems*, 2024.

[7] J. Chen, Y. Ling, J. Xu, Y. Ren, S. Huang, X. Pu, Z. Hao, S. Y. Philip, and L. He, "Variational graph generator for multiview graph clustering," *IEEE Transactions on Neural Networks and Learning Systems*, 2025.

[8] M.-S. Chen, X.-R. Zhu, J.-Q. Lin, and C.-D. Wang, "Contrastive multiview attribute graph clustering with adaptive encoders," *IEEE Transactions on Neural Networks and Learning Systems*, 2024.

[9] Y. Wang, L. Wu, X. Lin, and J. Gao, "Multiview spectral clustering via structured low-rank matrix factorization," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 29, no. 10, pp. 4833–4843, 2018.

[10] A. Kumar, P. Rai, and H. Daume, "Co-regularized multi-view spectral clustering," *Advances in Neural Information Processing Systems*, vol. 24, 2011.

[11] R. Xia, Y. Pan, L. Du, and J. Yin, "Robust multi-view spectral clustering via low-rank and sparse decomposition," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 28, no. 1, 2014.

[12] B. Perozzi, R. Al-Rfou, and S. Skiena, "Deepwalk: Online learning of social representations," in *Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2014, pp. 701–710.

[13] A. Grover and J. Leskovec, "node2vec: Scalable feature learning for networks," in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2016, pp. 855–864.

[14] T. N. Kipf and M. Welling, "Semi-supervised classification with graph convolutional networks," *arXiv preprint arXiv:1609.02907*, 2016.

[15] P. Veličković, G. Cucurull, A. Casanova, A. Romero, P. Lio, and Y. Bengio, "Graph attention networks," *arXiv preprint arXiv:1710.10903*, 2017.

[16] C. Xu, D. Tao, and C. Xu, "A survey on multi-view learning," *arXiv preprint arXiv:1304.5634*, 2013.

[17] Y. Yang and H. Wang, "Multi-view clustering: A survey," *Big data mining and analytics*, vol. 1, no. 2, pp. 83–107, 2018.

[18] S. Luo, C. Zhang, W. Zhang, and X. Cao, "Consistent and specific multi-view subspace clustering," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 32, no. 1, 2018.

[19] V. Sindhwani, P. Niyogi, and M. Belkin, "A co-regularization approach to semi-supervised learning with multiple views," in *Proceedings of ICML Workshop on Learning with Multiple Views*, vol. 2005. Citeseer, 2005, pp. 74–79.

[20] Z. Li, C. Tang, X. Liu, X. Zheng, W. Zhang, and E. Zhu, "Consensus graph learning for multi-view clustering," *IEEE Transactions on Multimedia*, vol. 24, pp. 2461–2472, 2021.

[21] C. Wang, S. Pan, G. Long, X. Zhu, and J. Jiang, "Mgae: Marginalized graph autoencoder for graph clustering," in *Proceedings of the 2017 ACM on Conference on Information and Knowledge Management*, 2017, pp. 889–898.

[22] W. Wang, R. Arora, K. Livescu, and J. Bilmes, "On deep multi-view representation learning," in *International Conference on Machine Learning*. PMLR, 2015, pp. 1083–1092.

[23] K. Hassani and A. H. Khasahmadi, "Contrastive multi-view representation learning on graphs," in *International Conference on Machine Learning*. PMLR, 2020, pp. 4116–4126.

[24] S. Pang, Y. Xue, Z. Yan, W. Huang, and J. Feng, "Dynamic and multi-channel graph convolutional networks for aspect-based sentiment analysis," in *Findings of the Association for Computational Linguistics: ACL-IJCNLP 2021*, 2021, pp. 2627–2636.

[25] Z. Lin and Z. Kang, "Graph filter-based multi-view attributed graph clustering." in *Proceedings of the International Joint Conference on Artificial Intelligence*, 2021, pp. 2723–2729.

[26] Y. Ling, J. Chen, Y. Ren, X. Pu, J. Xu, X. Zhu, and L. He, "Dual label-guided graph refinement for multi-view graph clustering," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 37, no. 7, 2023, pp. 8791–8798.

[27] Y. Liu, W. Tu, S. Zhou, X. Liu, L. Song, X. Yang, and E. Zhu, "Deep graph clustering via dual correlation reduction," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 36, no. 7, 2022, pp. 7603–7611.

[28] Z. Wen, Y. Ling, Y. Ren, T. Wu, J. Chen, X. Pu, Z. Hao, and L. He, "Homophily-related: Adaptive hybrid graph filter for multi-view graph clustering," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 38, no. 14, 2024, pp. 15 841–15 849.

[29] X. Yang, Y. Liu, S. Zhou, S. Wang, W. Tu, Q. Zheng, X. Liu, L. Fang, and E. Zhu, "Cluster-guided contrastive graph clustering network," in *Proceedings of the AAAI conference on artificial intelligence*, vol. 37, no. 9, 2023, pp. 10 834–10 842.

[30] Y. Liu, X. Yang, S. Zhou, X. Liu, Z. Wang, K. Liang, W. Tu, L. Li, J. Duan, and C. Chen, "Hard sample aware network for contrastive deep graph clustering," in *Proceedings of the AAAI conference on artificial intelligence*, vol. 37, no. 7, 2023, pp. 8914–8922.

[31] G.-Y. Zhang, D. Huang, and C.-D. Wang, "Tensorized incomplete multi-view kernel subspace clustering," *Neural Networks*, vol. 179, p. 106529, 2024.

[32] ——, "Unified and tensorized incomplete multi-view kernel subspace clustering," *IEEE Transactions on Emerging Topics in Computational Intelligence*, vol. 8, no. 2, pp. 1550–1566, 2024.

[33] ——, "Large-scale tensorized multi-view kernel subspace clustering," *ACM Transactions on Intelligent Systems and Technology*, vol. 16, pp. 1–21, 2025.

[34] G.-Y. Zhang, Y.-D. Huang, D. Huang, C.-D. Wang, Y. Liu, and E. Huang, "Structure-preserving contrastive graph clustering with dual-channel label alignment," *Neural Networks*, p. 107954, 2025.

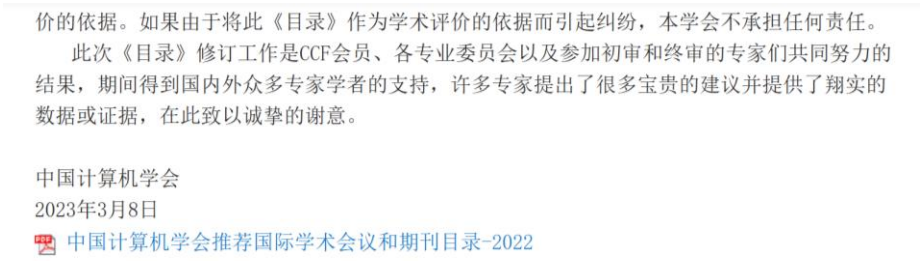
[35] G.-Y. Zhang, C.-B. Guan, D. Huang, Z. Wen, C.-D. Wang, and L. Xiao, "Scalable tri-factorization guided multi-view subspace clustering," *Knowledge-Based Systems*, vol. 312, p. 113119, 2025.

[36] Q. Zhang, G. Zhang, D. Huang, C. Wang, and H. Wang, "Multi-scale multi-order attributed graph clustering," in *International Conference on Intelligent Computing*, 2025, pp. 482–494.

附件：ICPADS 为中国计算机学会 CCF 推荐的 C 类会议论文的佐证

1.中国计算机学会推荐国际学术会议和期刊目录正式发布

https://www.ccf.org.cn/Academic_Evaluation/By_category/2023-03-08/787209.shtml



价的依据。如果由于将此《目录》作为学术评价的依据而引起纠纷，本学会不承担任何责任。

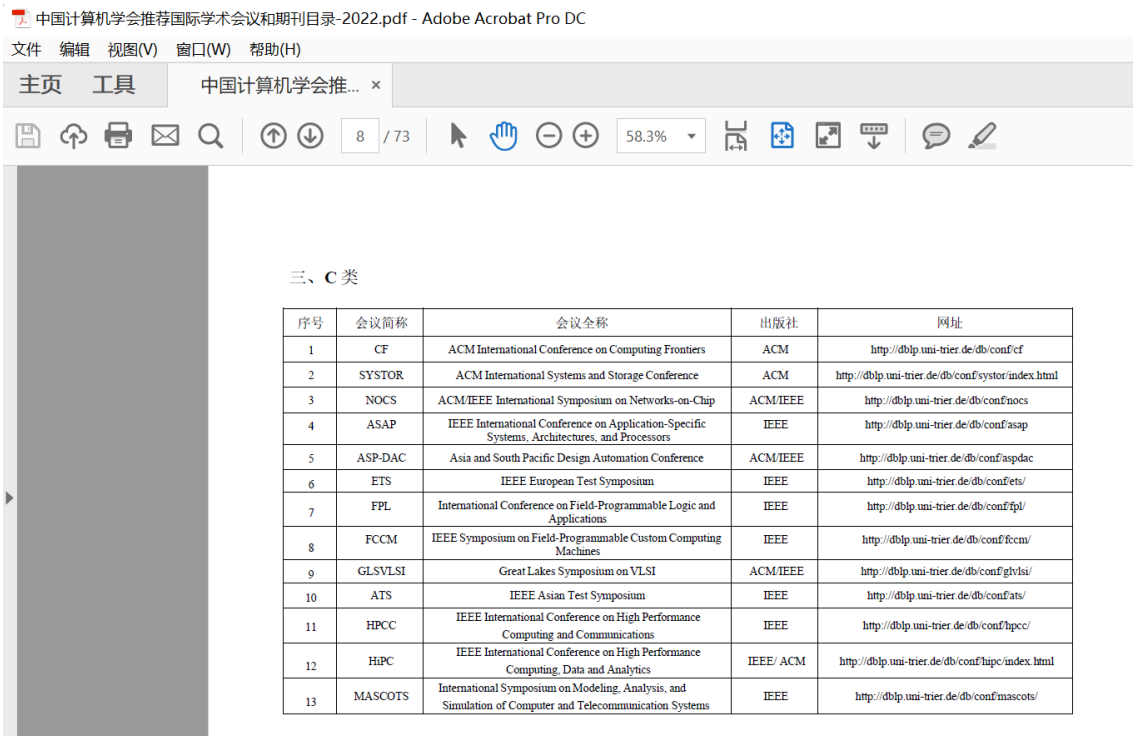
此次《目录》修订工作是CCF会员、各专业委员会以及参加初审和终审的专家们共同努力的结果，期间得到国内外众多专家学者的支持，许多专家提出了很多宝贵的建议并提供了翔实的数据或证据，在此致以诚挚的谢意。

中国计算机学会

2023年3月8日

中国计算机学会推荐国际学术会议和期刊目录-2022

2. 在《中国计算机学会推荐国际学术会议和期刊目录》（CCF 推荐目录）2022 年版本中，第 9 页将 ICPADS 列为 C 类会议。



中国计算机学会推荐国际学术会议和期刊目录-2022.pdf - Adobe Acrobat Pro DC

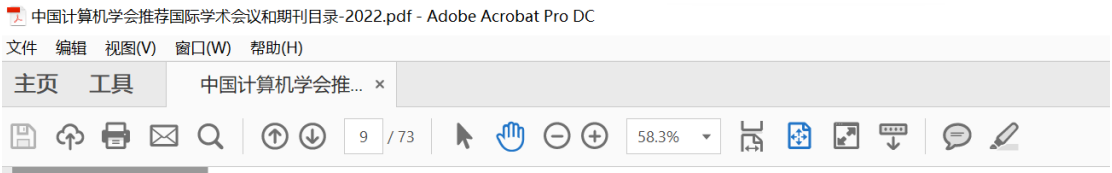
文件 编辑 视图(V) 窗口(W) 帮助(H)

主页 工具 中国计算机学会推... ×

8 / 73 58.3%

三、C 类

序号	会议简称	会议全称	出版社	网址
1	CF	ACM International Conference on Computing Frontiers	ACM	http://dblp.uni-trier.de/db/conf/cf
2	SYSTOR	ACM International Systems and Storage Conference	ACM	http://dblp.uni-trier.de/db/conf/systor/index.html
3	NOCS	ACM/IEEE International Symposium on Networks-on-Chip	ACM/IEEE	http://dblp.uni-trier.de/db/conf/nocs
4	ASAP	IEEE International Conference on Application-Specific Systems, Architectures, and Processors	IEEE	http://dblp.uni-trier.de/db/conf/asap
5	ASP-DAC	Asia and South Pacific Design Automation Conference	ACM/IEEE	http://dblp.uni-trier.de/db/conf/aspdac
6	ETS	IEEE European Test Symposium	IEEE	http://dblp.uni-trier.de/db/conf/ets/
7	FPL	International Conference on Field-Programmable Logic and Applications	IEEE	http://dblp.uni-trier.de/db/conf/fpl/
8	FCCM	IEEE Symposium on Field-Programmable Custom Computing Machines	IEEE	http://dblp.uni-trier.de/db/conf/fccm/
9	GLSVLSI	Great Lakes Symposium on VLSI	ACM/IEEE	http://dblp.uni-trier.de/db/conf/glvlsi/
10	ATS	IEEE Asian Test Symposium	IEEE	http://dblp.uni-trier.de/db/conf/ats/
11	HPCC	IEEE International Conference on High Performance Computing and Communications	IEEE	http://dblp.uni-trier.de/db/conf/hpcc/
12	HiPC	IEEE International Conference on High Performance Computing, Data and Analytics	IEEE/ ACM	http://dblp.uni-trier.de/db/conf/hipc/index.html
13	MASCOTS	International Symposium on Modeling, Analysis, and Simulation of Computer and Telecommunication Systems	IEEE	http://dblp.uni-trier.de/db/conf/mascots/



14	ISPA	IEEE International Symposium on Parallel and Distributed Processing with Applications	IEEE	http://dblp.uni-trier.de/db/conf/ispa/
15	CCGRID	IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing	ACM/IEEE	http://dblp.uni-trier.de/db/conf/ccgrid/
16	NPC	IFIP International Conference on Network and Parallel Computing	Springer	http://dblp.uni-trier.de/db/conf/npc/
17	ICA3PP	International Conference on Algorithms and Architectures for Parallel Processing	IEEE	http://dblp.uni-trier.de/db/conf/ica3pp/
18	CASES	International Conference on Compilers, Architectures, and Synthesis for Embedded Systems	ACM	http://dblp.uni-trier.de/db/conf/cases/index.html
19	FPT	International Conference on Field-Programmable Technology	IEEE	http://dblp.uni-trier.de/db/conf/fpt/
20	ICPADS	International Conference on Parallel and Distributed Systems	IEEE	http://dblp.uni-trier.de/db/conf/icpads/
21	ISCAS	IEEE International Symposium on Circuits and Systems	IEEE	http://dblp.uni-trier.de/db/conf/iscas/
22	ISLPED	International Symposium on Low Power Electronics and Design	ACM/IEEE	http://dblp.uni-trier.de/db/conf/islped/
23	ISPD	International Symposium on Physical Design	ACM	http://dblp.uni-trier.de/db/conf/ispd/
24	HOTI	IEEE Symposium on High-Performance Interconnects	IEEE	http://dblp.uni-trier.de/db/conf/hoti/
25	VTS	IEEE VLSI Test Symposium	IEEE	http://dblp.uni-trier.de/db/conf/vts/
26	ITC-Asia	International Test Conference in Asia	IEEE	https://dblp.org/db/conf/itc-asia/index.html

Multi-view Clustering via Flexible Dual-level Fusion

Guangyu Zhang
South China Agricultural University
Guangzhou, China
guangyuzhg@foxmail.com

Jiaming Deng
South China Agricultural University
Guangzhou, China
jiaming9413@163.com

Zihao Wen*
South China Agricultural University
Guangzhou, China
zihao.wen@hotmail.com

Dong Huang
South China Agricultural University
Guangzhou, China
huangdonghere@gmail.com

Changdong Wang
Sun Yat-sen University
Guangzhou, China
changdongwang@hotmail.com

2025 IEEE 31th International Conference on Parallel and Distributed Systems (ICPADS) | 979-8-3315-4901-5/25/$31.00 ©2025 IEEE | DOI: 10.1109/ICPADS67057.2025.11322983

***Abstract*—The anchor-based multi-view clustering is a prominent focus in massive data analysis. Although it has witnessed considerable progress, two critical limitations can still be observed in recent research. First, most existing clustering frameworks construct one anchor graph independently for each view, which may which may overlook the versatile diversity across multi-view latent distributions. In additional, they typically suffer from the two-stage issue, further restricting their flexibility in complex practical scenarios. By considering of this, we derive a new Multi-view Clustering approach via Flexible Dual-level Fusion (MC-FDF). Unlike the traditional one-anchor-graph-per-view practice, the proposed approach first produces a series of diversified anchor graphs from multiple views. Subsequently, these diversified anchor graphs are further fused from a scale-aware perspective, followed by the rotation of scale alignments to capture the consistent clustering structure across different views. With multiple views extended to the multi-scale multi-view paradigm, two levels of granularity (i.e. diversified anchor graph fusion and multi-scale late fusion) are seamlessly formulated into a mutually beneficial framework, in especially a facilitated optimization algorithm with linear time complexity is also provided. Comprehensive evaluations on heterogeneous benchmark datasets confirm that the proposed approach achieves superior performance and scalability over the state-of-the-art baselines.**

***Index Terms*—Multi-view Clustering, flexible dual-level fusion, diversified anchor graph fusion, Multi-scale late fusion.**

I. Introduction

The last decade has witnessed an exponential emergence of multi-view data has been witnessed in the past decade across different disciplines, ranging from data analysis to medical image processing [1]. Notably, multi-view observations are inherently originated from diverse modalities, each of which tends to capture distinct view-specific semantic characteristics. For instance, each medical image could be naturally represented by CT scans (providing anatomical structure) and PET scans (showing metabolic activity). As a result, how to jointly capture the cross-view complementarity and within-view specificity has posed a critical challenge for multi-view data analysis [2]. Building upon this, many data analysts frequently employ the clustering technology to harness the latent multi-view information for more comprehensive solutions.

The present multi-view clustering literature primarily falls into four paradigms, namely the kernel-driven approaches, the matrix factorization approaches [3], [4], the subspace learning [5], [6] and graph-oriented appraochess [7], [8]. Among the diverse branches of multi-view clustering, graph-oriented approaches (also known as spectral-based approaches) have garnered substantial interest owing to their provable guarantee and scalable architecture. In terms of distinct perspectives, an extensive range of graph-oriented approaches have been well investigated, mainly including MLAN [9] and GMC [10]. These graph-based studies have capacity in discovering the nonlinear topology from multiple views. Nevertheless, most of them still encounter computational bottlenecks (such as building affinity graphs and other matrix operations), seriously hindering their practical implementation in handing complex large-scale problems.

Aiming to alleviate the computational burden, recently widespread attention has been directed toward the research of scalable multi-view graph clustering. Rather than constructing a general graph (of size $n \times n$) for an individual view, the scalable multi-view graph clustering approaches (also abbreviated as the anchor graph-oriented approaches) endeavor to effectively capture the latent bipartite distribution by a limited set of anchor representations (of size $n \times p$), where p indicates the number of anchors and n represents the sample size. By deriving the inspiration from anchor graph formulation, a large amount of anchor graph-oriented approaches have been devised in the past years, primarily including LMVSC [11] and SMVSC [12]. Despite the advantageous improvement of efficiency made by the previous anchor graph-oriented approaches, in especially the computational complexity (of traditional practice) has been improved from $\mathrm{O}(n^3)$ to $\mathrm{O}(n)$, there still remain two essential issues that needs to be addressed. (i) How to go beyond the traditional practice, so as to better reflect the underlying relationships between anchors and samples in multi-scale scenarios? (ii) How to bridge the gaps between two mutually beneficial granularities for more promising clustering performance?

In order to deal with these issues, a new multi-view clustering approach with a flexible dual-level fusion framework is proposed. Specifically, our proposed approach first generates a collection of diversified anchor representations for each individual view, which helps to remedy the potential instability of traditional one-anchor-graph-per-view practice. Then, these diversified anchor representations are further fused from a scale-aware perspective, followed by the multi-scale late fusion to capture the consistent clustering partition across multiple views. Inherited from this, two levels of granularity are flexibly formed to exploit the hidden information underlying the multi-view data distributions. This design provides a flexible dual-level tool such that two granularities of multi-scale fusion tasks could mutually promote each other. Besides, a convergent four-stage alternating optimization procedure is designed to implement the proposed approach with linear time complexity. The main contributions of this paper is outlined as follows:

- With multiple views extended to multi-scale multiple views, this paper derives a flexible dual-level fusion framework for scalable yet effective multi-view clustering.
- The proposed dual-level fusion framework flexibly uncovers the inherent connection between diversified anchor graph fusion and multi-scale latent fusion, which seeks to directly obtain the balanced clustering partition from a hierarchical perspective.
- Experiments on nine real-world datasets, spanning from hundreds to over one hundred thousand instances, have demonstrated the superior clustering performance (in efficiency and scalability) achieved by our method relative to leading techniques.

II. Related work

Multi-view clustering (MVC) is an important methodology that leverages multiple data representations or perspectives to improve clustering results. Especially, MVC has been widely applied in domains including social network analysis, bioinformatics, and image processing. Compared to the single-view practice, MVC integrates information from heterogeneous sources or modalities, enhancing clustering accuracy and robustness. As different views often contain complementary information, the effective integration of these views remains a significant challenge. Recent advancements have introduced various methods to enhance MVC performance. For example, Cheng et al. [13] incorporated auxiliary knowledge to effectively determine the importance of each view. Nonetheless, this approach requires predefined knowledge, thus limiting its flexibility and stability. Moreover, Kumar et al. [14] proposed a co-regularized spectral clustering framework which ensures consistency among multiple views, leading to more coherent results. In addition, Wang et al. [15] further designed a co-training-guided spectral clustering approach that iteratively refines clustering results by enabling information exchange among views, improving the stability of clustering. Besides, Liu et al. [2] introduced low-rank constraints into subspace learning to preserve global structure while ensuring consistency across views, reducing noise and redundancy, and enhancing discriminative representations. Moving forward, Xu et al. [16] employed tensor decomposition to construct a shared subspace representation, improving the compatibility across views and facilitating the extraction of higher-order correlations, thus enhancing clustering performance. To address the high computational cost of extensive datasets, Qiang et al. [17] proposed an anchor graph-oriented clustering approach that directly computes discrete cluster assignments with linear time complexity. Additionally, de Kergorlay & Higham [18] provided theoretical consistency guarantees for anchor graph-oriented clustering in asymptotic regimes.

In the federated multi-view clustering domain, Chen et al. [19] derived a framework to control heterogeneous hybrid-view clients by combining local synergistic contrastive learning with global specific weighting aggregation for robust cluster learning. Meanwhile, Feng et al. [20] designed a tensorized approach that leverages tensor factorization and federated optimization to enable efficient and privacy-aware clustering. Furthermore, recent advancements have seen the emergence of several bipartite graph multi-view clustering approaches. Li et al. [21] introduced a scalable multi-view clustering framework without parameter tuning through self-adaptive graph integration, highlighting the potential of bipartite graphs for multi-view clustering.

III. Methodology

This section describes the details of our proposed approach, termed multi-view clustering via flexible dual-level fusion.

A. Notations and Main Objective

Formally, let us consider a multi-view dataset consisting of V sources as well as n instances, represented as $\mathcal{X} = [\mathbf{X}^{(1)}; \ldots; \mathbf{X}^{(V)}]$, in which $\mathbf{X}^{(v)} \in \mathbb{R}^{d_v \times n}$ indicates the data sub-matrix of the v-th view. For convenience of the later descriptions, the symbols adopted are summarized in Table I.

The main aim of traditional multi-view clustering is to merge the informational complementarity among diverse sources. In large-scale clustering, when the sample size n grows substantially, it presents a key challenge to efficiently comine heterogeneous representations while preserving promising scalability, which serves as the main focus of the subsequent subsections.

B. Constructing Diversified Anchor Graph Representations from Various Sources

The conventional anchor multi-view clustering tends to pursue a individual anchor representation per view, which always assumes that different views should share the collective anchor representations with fixed dimension. However, this assumption may present three types of difficulties when facing complex scenarios. In the first place, because of the distribution diversity among various sources, it is difficult to build a set of common anchor graph representations for all

TABLE I: Symbols Adopted Throughout the Paper

Notation	Meaning
n	The number of data samples
V	The number of sources
k	The number of clusters
M	The number of scales
P_m	The anchor size of the m-th scale
d_v	The dimension of the v-th view
$\mathbf{X}^{(v)} \in \mathbb{R}^{n\times d_v}$	The data matrix of the v-th view
$\mathbf{B}_m^{(v)} \in \mathbb{R}^{n\times P_m}$	The anchor graph matrix of the v-th view at the m-scale
$\mathbf{R}_m^{(v)} \in \mathbb{R}^{P_m\times P_m}$	The anchor rotation matrix of the v-th view at the m-scale
$\mathbf{H}_m \in \mathbb{R}^{n\times P_m}$	The representative anchor graph matrix at the m-th scale
$\mathbf{C}_m \in \mathbb{R}^{P_m\times k}$	The scale rotation matrix at the m-th scale
$\mathbf{Y} \in \{0,1\}^{n\times k}$	The consensus cluster indicator matrix

views. More importantly, given a set of common anchor graph representations, it is difficult to characterize the hierarchical structure hidden in multi-view data. Furthermore, many of them construct the common anchor graphs by k-means algorithm or random selection, and this could easily result in the inherently unstable performance.

To alleviate the above three difficulties of singe-anchor-per-view practice, we aim to produce many diversified anchor graphs from multiple views, which makes the proposed flexible dual-level fusion possible. Suppose $\mathcal{X} = [\mathbf{X}^{(1)}; \ldots; \mathbf{X}^{(V)}]$ is a given multi-view dataset, the v-th view data samples could be characterized by a set of selected anchor representatives, which motivates the following formulation.

$$\mathbf{X}^{(v)} \approx \mathbf{G}^{(v)}\mathbf{B}^{(v)}, \forall v, \tag{1}$$

where $\mathbf{G}^{(v)}$ represents the anchor set in the v-th view and $\mathbf{B}^{(v)}$ indicates the corresponding anchor graph representation. Different from the present studies, our approach employs the variance decorrelation anchor selection (VDA) strategy to select representative anchors (i.e., $\{\mathbf{G}^{(v)}\}_{v=1}^{V}$) from multiple views. On this basis, we further construct the robust anchor graph representations (i.e., $\{\mathbf{B}^{(v)}\}_{v=1}^{V}$) by using Gaussian kernels, which effectively encode the compact similarity among data samples and anchors within each view.

Following this line, we aim to jointly acquire many diversified anchor graphs with various magnitudes. Concretely, by repeating the VDA-based anchor generation process with diverse anchor sizes, a group of M anchor sets could be obtained on each view. The resultant anchor graph sets for the v-th view are formulated as $\{\mathbf{G}_1^{(v)}, \cdots, \mathbf{G}_M^{(v)}\}$, where $\mathbf{G}_m^{(v)}$ represents the m-th base anchor set in the v-th view. Meanwhile, the corresponding anchor representation is expressed as $\{\mathbf{B}_1^{(v)}, \cdots, \mathbf{B}_M^{(v)}\}$, where $\mathbf{B}_m^{(v)}$ represents the m-th base anchor embedding of the v-th view. In contrast to the previous studies, this design produces a series of $M * V$ diversified anchor representations for all views, which can effectively model the structural diversity by transitioning from multi-view representations to a hierarchical multi-scale clustering perspective.

C. Flexible Dual-level Fusion for Multi-view Clustering

With the many diversified anchor graph representations generated, our goal is to incorporate two levels of multi-scale fusion into a one-stage framework. The detailed descriptions of each component are introduced in what follows.

1) Diversified Anchor Graph Fusion: Suppose $\{\mathcal{B}^{(1)}, \cdots, \mathcal{B}^{(V)}\}$ are the diversified anchor graph representations that are generated from multiple views, where $\mathcal{B}^{(v)} = \{\mathbf{B}_1^{(v)}, \cdots, \mathbf{B}_M^{(v)}\}$ represents the v-th view anchor graph at M different scales. In our flexible dual-level fusion paradigm, the early stage seeks to integrate the complementary yet rich information hidden in heterogeneous views, yielding a set of multi-scale representatives that conveys comprehensive knowledge. Notably, it is reasonable to assume that the diversified anchor graph representations (within the same scale) should capture the shared scale-wise representative. Inspired by the recent advancements in multiple graph fusion technology, we can obtain the following formulation:

$$\min_{\mathbf{Y},\mathbf{C}_m^{(v)},\mathbf{H}_m,\mathbf{R}_m^{(v)}} \sum_{m=1}^{M}\sum_{v=1}^{V} \|\mathbf{H}_m - \mathbf{B}_m^{(v)}\mathbf{R}_m^{(v)}\|_F^2 \tag{2}$$

$$\text{s.t.} \quad \mathbf{H}_m \geq 0, \mathbf{H}_m\mathbf{1} = \mathbf{1}, (\mathbf{R}_m^{(v)})^\top\mathbf{R}_m^{(v)} = \mathbf{I}$$

where $\mathbf{B}_m^{(v)}$ indicates the v-th view anchor graph representation at the m-th scale, and $\mathbf{R}_m^{(v)}$ represents the corresponding rotation matrix. Besides, $\mathbf{H}_m$ indicates the fused anchor graph representation at the m-th scale (i.e., the representative at the m-th scale), where the simplex constraint $\mathbf{H}_m \geq 0, \mathbf{H}_m\mathbf{1} = \mathbf{1}$ restricts each entry to convey probability property. By means of the above formulation, representative $\mathbf{H}_m$ could be efficiently approximated by the combinations of $\mathbf{B}_m^{(v)}\mathbf{R}_m^{(v)}$ (which can also be interpreted as the anchor rotations of $\{\mathbf{B}_m^{(v)}\}_{v=1}^{V}$). In the following, this formulation will serve as the early stage (i.e., view fusion awareness perspective) of our flexible dual-level fusion framework.

2) Multi-scale Late Fusion: Moving forward, we formulate the details of multi-scale late fusion. Besides the diversified anchor graph fusion, the second stage excepts to generate the optimal clustering partition at multi-scale fusion level. Previous studies have indicated that the late fusion strategy can efficiently reduce the computational complexity while enhancing clustering effectiveness. Following this principle, the consensus partition matrix should be well aligned with the multi-scale representatives $\{\mathbf{H}_m\}_{m=1}^{M}$. Concretely, the corresponding formulation can be expressed as what follows.

$$\begin{aligned} &\max_{\mathbf{C}_m,\mathbf{Y}} \sum_{m=1}^{M} \text{Tr}\left((\mathbf{Y}^\top\mathbf{Y})^{-\frac{1}{2}}\mathbf{Y}^\top\mathbf{H}_m\mathbf{C}_m\right) \\ \Leftrightarrow &\min_{\mathbf{C}_m,\mathbf{Y}} -\sum_{m=1}^{M} \text{Tr}\left((\mathbf{Y}^\top\mathbf{Y})^{-\frac{1}{2}}\mathbf{Y}^\top\mathbf{H}_m\mathbf{C}_m\right) \end{aligned} \tag{3}$$

$$\text{s.t.} \quad \mathbf{Y} \in \{0,1\}^{n\times k}, \mathbf{Y}\mathbf{1}_k = \mathbf{1}_n, \mathbf{C}_m^\top\mathbf{C}_m = \mathbf{I}_k$$

where $\mathbf{Y}$ represents the one-hot clustering indicator matrix, and $\mathbf{Y}_{il}$ is 1 if the i-th sample is assigned to the l-th cluster. Notice that, the matrix $(\mathbf{Y}^\top\mathbf{Y})^{-\frac{1}{2}}$ here is used to balance

the distribution of different clusters. Moreover, $\mathbf{C}_m$ represents the v-th orthogonal transformation matrix that enables $\mathbf{H}_m$ better fit the clustering indicator matrix $\mathbf{Y}$. Especially, this formulation helps to integrate the hierarchical information in a late fusion manner, such that the multi-scale representatives can well serve for the generation of robust clustering labels. As a result, the above formulation favors the late stage (i.e., multi-scale fusion perspective) of our flexible dual-level fusion framework.

D. The Proposed Model

With the combination of two-stage fusion into a unified framework, the complete objective function can be expressed as follows:

$$\begin{aligned}\min_{\mathbf{Y},\mathbf{C}_m,\mathbf{H}_m,\mathbf{R}_m^{(v)}} &\sum_{m=1}^{M}\sum_{v=1}^{V}\|\mathbf{H}_m-\mathbf{B}_m^{(v)}\mathbf{R}_m^{(v)}\|_F^2 \\ &-\alpha\sum_{m=1}^{M}\mathrm{Tr}\Big((\mathbf{Y}^\top\mathbf{Y})^{-\frac{1}{2}}\mathbf{Y}^\top\mathbf{H}_m\mathbf{C}_m\Big)\end{aligned} \tag{4}$$

$$\begin{aligned}\text{s.t.}\quad &\mathbf{Y}\in\{0,1\}^{n\times k},\mathbf{Y}\mathbf{1}_k=\mathbf{1}_n,\mathbf{C}_m^\top\mathbf{C}_m=\mathbf{I}_k,\\ &\mathbf{H}_m\geq 0,\mathbf{H}_m\mathbf{1}_{P_m}=\mathbf{1}_n,(\mathbf{R}_m^{(v)})^\top\mathbf{R}_m^{(v)}=\mathbf{I}_{P_m}\end{aligned}$$

where α is a parameter that balances the weight of diversified anchor graph fusion and multi-scale late fusion. By going beyond the traditional practice, the proposed objective function provides a flexible solution to simultaneously optimize the consensus clustering partition, the multi-scale representative anchor graph representations, and the rotation matrices into a joint framework. In this way, the two fusion procedures can interact to reach an optimal solution, resulting in enhanced clustering performance.

IV. Optimization and Theoretical Analysis

A. Optimization of Problem

In this subsection, an alternating optimization scheme is adopted to iteratively update $\mathbf{Y}$, $\{\mathbf{C}_m\}_{m=1}^M$, $\{\mathbf{H}_m\}_{m=1}^M$, and $\{\mathbf{R}_m^{(v)}\}_{m=1,\,v=1}^{M,\,V}$ respectively, while keeping the others fixed. The update rules for each variable are summarized as follows.

1) Update $\mathbf{Y}$: With other variables fixed, the subproblem depends solely on $\mathbf{Y}$, and can be formulated as

$$\max_{\mathbf{Y}}\ \mathrm{Tr}\Big((\mathbf{Y}^\top\mathbf{Y})^{-\frac{1}{2}}\mathbf{Y}^\top\mathbf{Q}\Big), \tag{5}$$

subject to $\mathbf{Y}\in\{0,1\}^{n\times k}$ and $\mathbf{Y}\mathbf{1}_k=\mathbf{1}_n$, where $\mathbf{Q}=\sum_{m=1}^M\mathbf{H}_m\mathbf{C}_m$.

Following [22], $\mathbf{Y}$ is updated by iteratively maximizing the gain Δ_{ij}:

$$\begin{aligned}\Delta_{ij}=&\frac{\sum_{s=1}^n y_{sj}q_{sj}+q_{ij}(1-y_{ij})}{\sqrt{y_j^\top y_j+(1-y_{ij})}}\\ &-\frac{\sum_{s=1}^n y_{sj}q_{sj}+q_{ij}y_{ij}}{\sqrt{y_j^\top y_j-y_{ij}}},\end{aligned} \tag{6}$$

where y_j means the j-th column of $\mathbf{Y}$. The update rule is given by

$$y_{ij}=\left\langle j=\arg\max_{l\in\{1,\ldots,k\}}\Delta_{il}\right\rangle, \tag{7}$$

where $\langle\cdot\rangle$ equals 1 if the condition is satisfied and 0 otherwise.

2) Update $\{\mathbf{C}_m\}_{m=1}^M$: With other variables fixed, the subproblem depends solely on $\mathbf{C}_m$, and can be expressed as:

$$\max_{\mathbf{C}_m^\top\mathbf{C}_m=\mathbf{I}_k}\sum_{m=1}^{M}\mathrm{Tr}\left(\mathbf{C}_m^\top\mathbf{H}_m\mathbf{Y}(\mathbf{Y}^\top\mathbf{Y})^{-\frac{1}{2}}\right). \tag{8}$$

The solution to problem 8 is derived via a compact singular value decomposition (SVD) of the matrix $\mathbf{H}_m\mathbf{Y}(\mathbf{Y}^\top\mathbf{Y})^{-\frac{1}{2}}$. Specifically, let

$$\mathbf{H}_m\mathbf{Y}(\mathbf{Y}^\top\mathbf{Y})^{-\frac{1}{2}}=\mathbf{U}\mathbf{\Sigma}\mathbf{V}^\top, \tag{9}$$

where $\mathbf{U}\in\mathbb{R}^{P_m\times K}$ and $\mathbf{V}\in\mathbb{R}^{K\times K}$ are orthogonal matrices. The optimal $\mathbf{C}_m$ is then given by

$$\mathbf{C}_m=\mathbf{U}\mathbf{V}^\top. \tag{10}$$

3) Update $\mathbf{H}_m$: With other variables fixed, the subproblem depends solely on $\mathbf{H}_m$, and can be expressed as:

$$\begin{aligned}\min_{\mathbf{H}_m}\ &\sum_{v=1}^{V}\left\|\mathbf{H}_m-\mathbf{B}_m^{(v)}\mathbf{R}_m^{(v)}\right\|_F^2-\alpha\,\mathrm{Tr}\Big((\mathbf{Y}^\top\mathbf{Y})^{-\frac{1}{2}}\mathbf{Y}^\top\mathbf{H}_m\mathbf{C}_m\Big)\\ \text{s.t. }&\mathbf{H}_m\geq 0,\quad \mathbf{H}_m\mathbf{1}_{P_m}=\mathbf{1}_n.\end{aligned} \tag{11}$$

By setting the derivative of the objective with respect to $\mathbf{H}_m$ to zero and relaxing the constraints, the unconstrained solution is computed as:

$$\widetilde{\mathbf{H}}_m=\frac{1}{V}\sum_{v=1}^{V}\mathbf{B}_m^{(v)}\mathbf{R}_m^{(v)}+\frac{\alpha}{2V}\,\mathbf{Y}(\mathbf{Y}^\top\mathbf{Y})^{-1/2}\mathbf{C}_m^\top. \tag{12}$$

To enforce the non-negativity and row-stochastic constraints, each row of $\widetilde{\mathbf{H}}_m$ is projected onto the probability simplex

$$\Delta_{P_m}=\{\mathbf{x}\in\mathbb{R}^{P_m}\mid\mathbf{x}\geq 0,\ \mathbf{1}_{P_m}^\top\mathbf{x}=1\}$$

The projection is performed using the efficient algorithm described in [23], which ensures that the updated $\mathbf{H}_m$ remains a valid non-negative and normalized representation.

4) Update $\left\{\mathbf{R}_m^{(v)}\right\}_{m=1,\,v=1}^{M,\,V}$: With other variables fixed, the subproblem depends solely on $\mathbf{R}_m^{(v)}$, and can be formulated as:

$$\max_{\mathbf{R}_m^{(v)T}\mathbf{R}_m^{(v)}=\mathbf{I}_{P_m}}\sum_{v=1}^{V}\mathrm{Tr}\left(\mathbf{R}_m^{(v)T}\mathbf{B}_m^{(v)T}\mathbf{H}_m\right). \tag{13}$$

The solution to problem 13 is derived via a compact SVD of the matrix $\mathbf{B}_m^{(v)T}\mathbf{H}_m$. Specifically, let

$$\mathbf{B}_m^{(v)T}\mathbf{H}_m=\mathbf{U}\mathbf{\Sigma}\mathbf{V}^T, \tag{14}$$

in which $\mathbf{U}\in\mathbb{R}^{P_m\times K}$ and $\mathbf{V}\in\mathbb{R}^{K\times K}$ are orthogonal matrices. Then, the optimal $\mathbf{R}_m^{(v)}$ is given by

$$\mathbf{R}_m^{(v)}=\mathbf{U}\mathbf{V}^T. \tag{15}$$

To enhance comprehensibility, Algorithm 1 summarizes the full procedure of our MC-FDF approach.

Algorithm 1 MC-FDF

Input: Dataset $\mathbf{X} = \{\mathbf{X}^{(1)}, \cdots, \mathbf{X}^{(V)}\}$ generated from V sources; maximum number of iterations $t_{\max}$ (set to 30); number of multi-scale anchors M; and parameter α.
Parameter Setup: Set $\alpha = 10^{-5}$ and $M = 4$.

Initialization: Set $t = 1$. Initialize $\mathbf{R}_{\mathbf{m}}^{(\mathbf{v})} = \mathbf{I}_{P_m}$; $\mathbf{H}_{\mathbf{m}} = \frac{1}{V}\sum_{v=1}^{V} \mathbf{B}_{\mathbf{m}}^{(\mathbf{v})}$; use K-means to initialize $\mathbf{Y}$; initialize $\mathbf{C}_{\mathbf{m}}$ using Eq. (10).
repeat
  Acquire $\mathbf{Y}$ by utilizing Eq. (7).
  Acquire $\mathbf{C}_{\mathbf{m}}$ by utilizing Eq. (10).
  Acquire $\mathbf{H}_m$ by utilizing Eq. (12).
  Acquire $\mathbf{R}_{\mathbf{m}}^{(\mathbf{v})}$ by utilizing Eq. (15).
  $t = t + 1$.
until The convergence condition is met or $t > t_{\max}$.
Output: The final clustering label $\mathbf{Y}$.

B. Computational Complexity Analysis

The computational cost of MC-FDF mainly arises from constructing the anchor graph representations and calculating the optimization variables. Constructing the anchor graph representations $\{\mathbf{B}_m{}^{(v)}\}_{v=1}^{V}$ from the data matrix takes $O(MnpD)$, where M is the number of scales, n is the number of samples, p is the number of anchors, and D is the feature dimensionality. Notably, each iteration involves four part. First, updating $\mathbf{Y}$ needs $O((MnpK + \text{IterY} \cdot nK))$. Second, updating $\{\mathbf{C}_\mathbf{m}\}_{m=1}^{M}$ needs $O(M(npK + pK^2 + K^3))$. Third, updating $\{\mathbf{H}_m\}_{m=1}^{M}$ needs $O(MVnp^2 + Mnp\log p)$. Finally, updating $\{\mathbf{R}_\mathbf{m}^{(\mathbf{v})}\}_{m,v=1}^{M,V}$ needs $O(MV(p^2n + p^3))$, where V is the number of views, K is the number of clusters, and IterY is the number of inner iterations for Y. The overall bottleneck complexity is $O(TMVnp^2 + MnpD)$, where T is the number of outer iterations. Because T, V, p, D are typically small and $M \ll n$, the approach scales linearly with the number of samples n.

V. Experiments

A. Benchmarks and Metrics

In the experimental section, we evaluate the proposed approach along with several baselines on nine real-world benchmarks, involve MSRC-v1, Movies, Outdoor Scene, AwA, and VGG2, with sample sizes ranging from 210 to 124,880. The details of these benchmarks are summarized in Table II. For performance evaluation, three frequently adopted clustering metrics are employed, namely accuracy (ACC) [24], purity (PUR) [25], and normalized mutual information (NMI) [26].

B. Experimental Settings

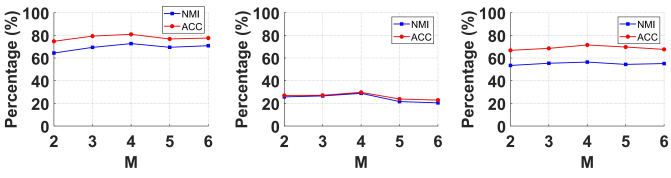
Moving forward, eight recently proposed baselines are selected to ensure the complementarity and fairness of the practical comparison. Additionally, the baselines can be further categorized into the graph-oriented approaches (i.e., FDAGF [27], UOMVSC [28] and SFMC [21]) and subspace learning-based approaches (i.e., AWMVC [29], OMSC [30], SMVSC [11], FPMVS [31] and CFMC [32]). The details of these state-of-arts could be found in the original papers.

TABLE II: **Dataset Clustering Information**

No.	Dataset	Sample	View	Cluster
1	MSRC-v1	210	4	7
2	Movies	617	2	17
3	Outdoor Scene	2688	4	8
4	AwA	30076	6	50
5	VGG2_cropped3_up_50	34027	4	50
6	cifer10_mv_up_big_up_50	60000	4	10
7	cifer100_mv_up_big_up_50	60000	4	100
8	VGG2_cropped3_up_100	65324	4	100
9	VGG2_cropped3_up_200	124880	4	200

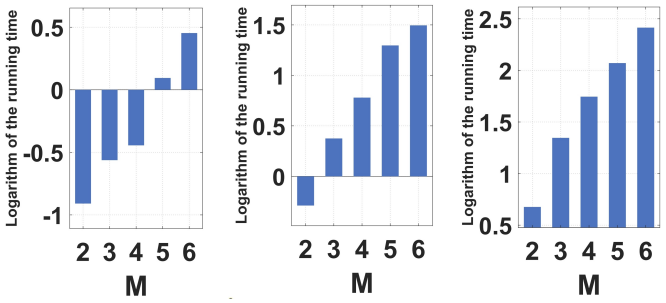
For each benchmark, every comparison approach is executed twenty times, while the mean performance is reported in terms of NMI, ACC, and PUR. For our MC-FDF approach, the trade-off parameter is searched over the range of $[10^{-5}, 10^{-4}, \ldots, 10^{5}]$. Similarly, the hyperparameters of the baselines are also searched over the same range of $[10^{-5}, 10^{-4}, \ldots, 10^{5}]$, unless a specific tuning range is provided in their respective papers. In our experiments, the parameter M is set to 4, as this value provides an optimal trade-off between accuracy and computational cost.

As shown in Fig. 1, the clustering performance (measured by NMI and ACC) across different datasets peaks at M=4, while smaller values degrade representational capability and thus reduce accuracy. Conversely, larger values of M produce only marginal accuracy gains while significantly increasing computational costs. Fig. 2 demonstrates this trend by showing running time grows substantially with M. The value M=4 achieves the best balance between effectiveness and efficiency.



(a) MSRC-v1 (b) Movies (c) Outdoor Scene

Fig. 1: **Clustering performance under varying parameter M measured by NMI and ACC.**



(a) MSRC-v1 (b) Movies (c) Outdoor Scene

Fig. 2: **Running time of the algorithm for different values of M.**

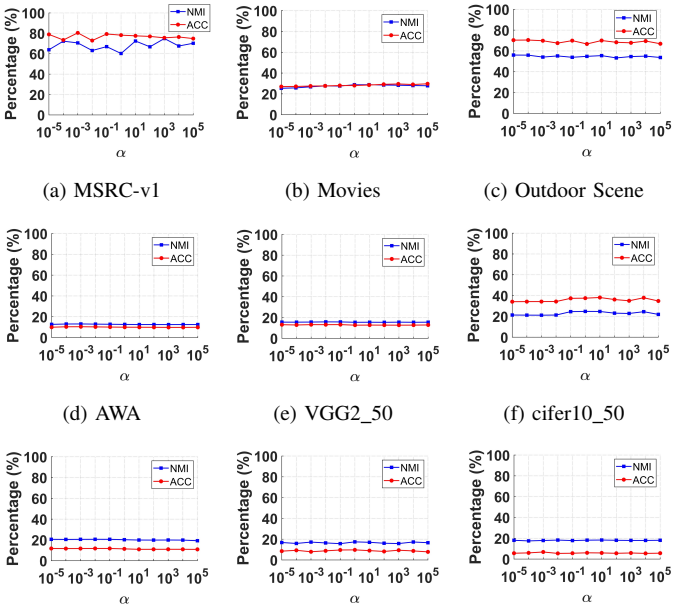
C. Comparison Results and Analysis

In this subsection, we present and examine the performance outcomes of MC-FDF relative to eight baselines over nine tested benchmarks. The clustering outcomes, measured by ACC, NMI and PUR, are reported in Table III. Notably, the top-performing results are shown in bold, whereas the second-highest scores are indicated in square brackets. It should be noted that a dash - in the table denotes that the corresponding approach failed to complete due to an out-of-memory (OOM) error. As shown in Table III, MC-FDF achieves the highest ACC on most of the benchmarks. Unlike conventional graph-oriented approaches that keep anchor points fixed after initialization, our approach adaptively capture a group of anchor representations throughout the optimization process, resulting in consistently higher NMI scores across all datasets. With respect to ACC and PUR, MC-FDF achieves either the highest or highly competitive scores on most benchmarks. Overall, the results in Table III clearly validate the superiority of MC-FDF compared to existing promising subspace- and spectral-based clustering baselines.

D. Parameter Sensitivity and Convergence Analysis

Next, we conduct experiments to examine the influence of the trade-off parameter α and the convergence behavior of MC-FDF across three representative datasets. The sensitivity analysis of α is presented in Fig. 3. Moderate or small values of α generally yield superior clustering performance (measured by NMI and ACC), indicating a balanced trade-off between bipartite graph construction and partitioning. The ablation study further demonstrates that the clustering results vary smoothly across a wide range of α values, without abrupt fluctuations. This indicates that MC-FDF is relatively insensitive to α, maintaining robust performance. Such stability arises from the multi-scale anchor graph fusion and alignment mechanism, which effectively balances the influence of graph construction and partitioning.

Fig. 4 depicts the convergence behavior of the objective function on three benchmarks. The value of the objective steadily declines and reaches convergence after only a few iterations, indicating a consistent convergence pattern under diverse data distributions. These findings demonstrate that the optimization algorithm achieves both superior efficiency and robust stability, while the multi-scale anchor graph fusion and alignment framework reliably yields high-quality clustering results. This pronounced convergence property ensures that MC-FDF sustains stable performance on large-scale datasets, effectively preventing unstable or divergent optimization.

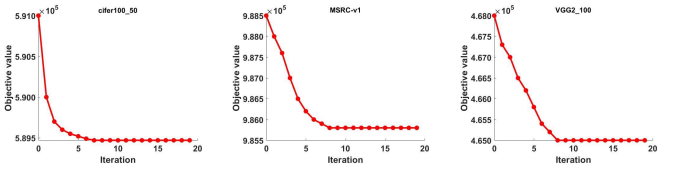


(a) MSRC-v1 (b) Movies (c) Outdoor Scene

(d) AWA (e) VGG2_50 (f) cifer10_50

(g) cifer100_50 (h) VGG2_100 (i) VGG2_200

Fig. 3: **Parameter analysis on α.**



(a) cifer100_50 (b) MSRC-v1 (c) VGG2_100

Fig. 4: **Convergence curves on three datasets.(a) cifer100_50; (b) MSRC-v1; (c) VGG2_100.**

E. Execution Time

As illustrated in Fig. 5, MC-FDF exhibits efficient and stable computational performance across benchmarks of different scales, achieving faster execution on small datasets and consistent efficiency on large ones without memory failures.

VI. Conclusion

In this paper, we introduce a flexible dual-level integration framework for scalable multi-view clustering (MC-FDF). Unlike the conventional practice of constructing the single anchor graph for each view, MC-FDF endeavors to learn a set of diversified anchor graph representations with multiple granularities, which mitigates the potential instability associated with the single-anchor graph learning and substantially preserves highly-competitive efficiency. Specifically, the proposed dual-level fusion framework enhances the mutual negotiation

TABLE III: The clustering performance comparison on nine benchmarks.

Dataset	Metric	FDAGF	AWMVC	OSMC	UOMvSC	SMVSC	SFMC	FPMVS	CFMC	Ours
MSRC-v1	ACC	76.2	[78.1]	60.9	77.6	67.1	56.7	60.0	66.7	**80.5**
	NMI	68.2	[71.6]	56.1	56.5	61.1	52.3	56.6	61.8	**71.6**
	PUR	76.2	78.1	62.4	[81.4]	69.0	59.5	62.8	71.9	**81.7**
Movies	ACC	23.4	**33.9**	20.2	19.1	21.4	8.9	17.0	13.0	[29.7]
	NMI	26.3	**33.4**	15.6	17.5	16.8	22.9	14.5	11.1	[28.8]
	PUR	**57.4**	[37.0]	21.2	20.1	23.5	29.0	18.6	14.9	33.1
Outdoor Scene	ACC	57.7	67.4	65.6	[71.1]	63.3	47.8	63.6	55.9	**71.5**
	NMI	46.0	**67.4**	52.9	60.6	51.7	39.7	53.0	47.9	[56.0]
	PUR	62.2	62.1	67.2	[71.0]	66.0	48.2	67.1	58.4	**71.8**
AwA	ACC	9.0	-	[9.2]	-	9.1	6.6	9.1	8.6	**10.4**
	NMI	[11.0]	-	10.2	-	9.6	5.4	9.7	11.4	**13.1**
	PUR	[11.1]	-	9.8	-	9.6	7.2	9.6	10.6	**12.3**
VGG2_50	ACC	12.5	[13.7]	9.6	-	9.8	8.1	10.7	11.6	**14.2**
	NMI	15.3	**17.4**	11.8	-	11.1	7.8	12.3	14.2	[15.8]
	PUR	**17.1**	[15.7]	10.1	-	6.7	8.5	11.1	12.7	14.2
cifer10_50	ACC	26.9	[30.1]	31.5	-	29.2	11.5	29.6	29.8	**38.1**
	NMI	14.0	16.0	[18.4]	-	18.4	1.8	18.4	17.0	**24.5**
	PUR	29.5	30.5	[33.3]	-	32.4	11.9	32.8	30.2	**39.0**
cifer100_50	ACC	11.0	[11.3]	8.7	-	8.1	6.9	9.0	10.4	**11.8**
	NMI	18.2	18.3	14.1	-	13.9	12.2	13.8	[19.3]	**20.6**
	PUR	[13.6]	13.0	9.1	-	8.6	7.6	9.4	12.1	**13.8**
VGG2_100	ACC	10.0	**11.8**	7.3	-	6.6	6.1	7.2	9.0	[10.1]
	NMI	[17.6]	**19.3**	12.2	-	11.9	10.2	11.8	16.8	17.4
	PUR	**13.3**	[12.5]	7.6	-	6.7	6.6	7.3	9.9	10.6
VGG2_200	ACC	-	-	-	-	4.1	4.2	4.2	[5.9]	**6.8**
	NMI	-	-	-	-	11.1	11.4	9.6	[17.2]	**17.9**
	PUR	-	-	-	-	4.4	4.5	4.3	[6.8]	**7.5**

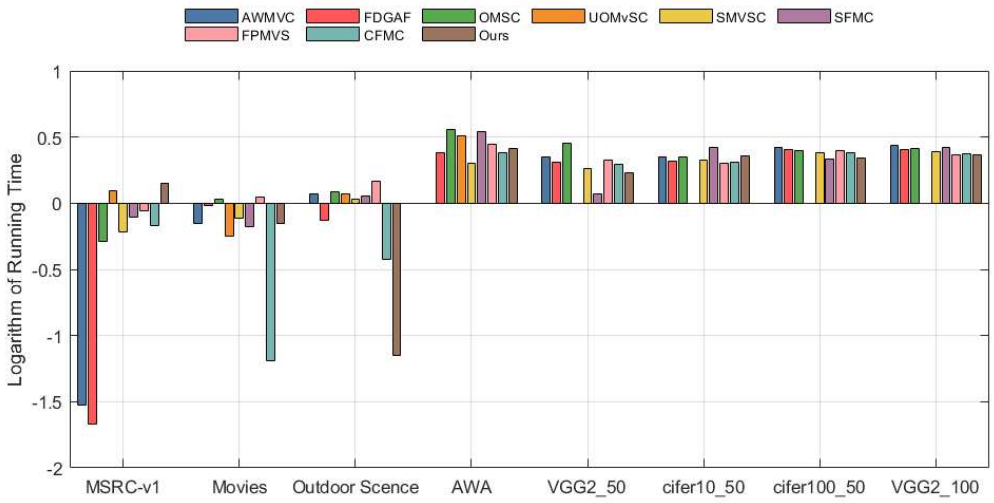


Fig. 5: **Logarithm of the running time on eight datasets (in Sec.).**

between multi-scale representative anchor graph and consensus clustering partition. The results from experiments on real-world datasets validate the effectiveness and computational efficiency of the MC-FDF method.

REFERENCES

[1] Xia, R., Pan, Y., Du, L., Yin, J.: Robust multi-view spectral clustering via low-rank and sparse decomposition. In: Proceedings of the AAAI conference on artificial intelligence. vol. 28 (2014)

[2] Liu, J., Wang, C., Gao, J., Han, J.: Multi-view clustering via joint nonnegative matrix factorization. In: Proceedings of the 2013 SIAM international conference on data mining. pp. 252–260. SIAM (2013)

[3] Zhang, G.Y., Guan, C.B., Huang, D., Wen, Z., Wang, C.D., Xiao, L.: Scalable tri-factorization guided multi-view subspace clustering. Knowledge-Based Systems **312**, 113119 (2025)

[4] Zhang, G.Y., Huang, D., Wang, C.D.: Large-scale tensorized multi-view kernel subspace clustering. ACM Transactions on Intelligent Systems and Technology **16**, 1–21 (2025)

[5] Zhang, G.Y., Huang, D., Wang, C.D.: Facilitated low-rank multi-view subspace clustering. Knowledge-Based Systems **260**, 110141 (2023)

[6] Zhang, G.Y., Huang, D., Wang, C.D.: Unified and tensorized incomplete multi-view kernel subspace clustering. IEEE Transactions on Emerging Topics in Computational Intelligence **8**(2), 1550–1566 (2024)

[7] Liang, Y., Huang, D., Wang, C.D., Yu, P.S.: Multi-view graph learning by joint modeling of consistency and inconsistency. IEEE transactions on neural networks and learning systems **35**(2), 2848–2862 (2022)

[8] Wang, R., Wang, P., Wu, D., Sun, Z., Nie, F., Li, X.: Multi-view

and multi-order structured graph learning. IEEE Transactions on Neural Networks and Learning Systems **35**(10), 14437–14448 (2023)
[9] Nie, F., Cai, G., Li, X.: Multi-view clustering and semi-supervised classification with adaptive neighbours. In: Proceedings of the AAAI conference on artificial intelligence. vol. 31 (2017)
[10] Wang, H., Yang, Y., Liu, B.: Gmc: Graph-based multi-view clustering. IEEE Transactions on Knowledge and Data Engineering **32**(6), 1116–1129 (2019)
[11] Sun, M., Zhang, P., Wang, S., Zhou, S., Tu, W., Liu, X., Zhu, E., Wang, C.: Scalable multi-view subspace clustering with unified anchors. In: Proceedings of the 29th ACM international conference on multimedia. pp. 3528–3536 (2021)
[12] Chen, X., Cai, D.: Large scale spectral clustering with landmark-based representation. In: Proceedings of the AAAI Conference on Artificial Intelligence. vol. 25, pp. 313–318 (2011)
[13] Cheng, Y., Zhao, R.: Multiview spectral clustering via ensemble. In: 2009 IEEE international conference on granular computing. pp. 101–106. IEEE (2009)
[14] Kumar, A., Rai, P., Daume, H.: Co-regularized multi-view spectral clustering. Advances in neural information processing systems **24** (2011)
[15] Wang, H., Nie, F., Huang, H.: Multi-view clustering and feature learning via structured sparsity. In: International conference on machine learning. pp. 352–360. PMLR (2013)
[16] Xu, C., Tao, D., Xu, C.: Multi-view learning with incomplete views. IEEE Transactions on Image Processing **24**(12), 5812–5825 (2015)
[17] Qiang, Q., Zhang, B., Wang, F., Nie, F.: Fast multi-view discrete clustering with anchor graphs. In: Proceedings of the AAAI Conference on Artificial Intelligence. vol. 35, pp. 9360–9367 (2021)
[18] de Kergorlay, H.L., Higham, D.J.: Consistency of anchor-based spectral clustering. arXiv preprint arXiv:2006.13984 (2020)
[19] Chen, X., Ren, Y., Xu, J., Lin, F., Pu, X., Yang, Y.: Bridging gaps: Federated multi-view clustering in heterogeneous hybrid views. Advances in Neural Information Processing Systems **37**, 37020–37049 (2024)
[20] Feng, W., Wu, Z., Wang, Q., Dong, B., Tao, Z., Gao, Q.: Federated multi-view clustering via tensor factorization. In: Proceedings of the Thirty-Third International Joint Conference on Artificial Intelligence, IJCAI-24. pp. 3962–3970 (2024)
[21] Li, X., Zhang, H., Wang, R., Nie, F.: Multiview clustering: A scalable and parameter-free bipartite graph fusion method. IEEE transactions on pattern analysis and machine intelligence **44**(1), 330–344 (2020)
[22] Huang, J., Nie, F., Huang, H.: Spectral rotation versus k-means in spectral clustering. In: Proceedings of the AAAI conference on artificial intelligence. vol. 27, pp. 431–437 (2013)
[23] Duchi, J., Shalev-Shwartz, S., Singer, Y., Chandra, T.: Efficient projections onto the l 1-ball for learning in high dimensions. In: Proceedings of the 25th international conference on Machine learning. pp. 272–279 (2008)
[24] Zhang, G.Y., Huang, D., Wang, C.D.: Tensorized incomplete multi-view kernel subspace clustering. Neural Networks **179**, 106529 (2024)
[25] Liang, Y., Huang, D., Wang, C.D.: Consistency meets inconsistency: A unified graph learning framework for multi-view clustering. In: 2019 IEEE international conference on data mining (ICDM). pp. 1204–1209 (2019)
[26] Zhang, G.Y., Huang, Y.D., Huang, D., Wang, C.D., Liu, Y., Huang, E.: Structure-preserving contrastive graph clustering with dual-channel label alignment. Neural Networks p. 107954 (2025)
[27] Zhang, P., Wang, S., Li, L., Zhang, C., Liu, X., Zhu, E., Liu, Z., Zhou, L., Luo, L.: Let the data choose: Flexible and diverse anchor graph fusion for scalable multi-view clustering. In: Proceedings of the AAAI Conference on Artificial Intelligence. vol. 37, pp. 11262–11269 (2023)
[28] Tang, C., Li, Z., Wang, J., Liu, X., Zhang, W., Zhu, E.: Unified one-step multi-view spectral clustering. IEEE Transactions on Knowledge and Data Engineering **35**(6), 6449–6460 (2022)
[29] Wan, X., Liu, X., Liu, J., Wang, S., Wen, Y., Liang, W., Zhu, E., Liu, Z., Zhou, L.: Auto-weighted multi-view clustering for large-scale data. In: Proceedings of the AAAI Conference on Artificial Intelligence. vol. 37, pp. 10078–10086 (2023)
[30] Chen, M.S., Wang, C.D., Huang, D., Lai, J.H., Yu, P.S.: Efficient orthogonal multi-view subspace clustering. In: Proceedings of the 28th ACM SIGKDD conference on knowledge discovery and data mining. pp. 127–135 (2022)
[31] Wang, S., Liu, X., Zhu, X., Zhang, P., Zhang, Y., Gao, F., Zhu, E.: Fast parameter-free multi-view subspace clustering with consensus anchor guidance. IEEE Transactions on Image Processing **31**, 556–568 (2021)
[32] Chen, M.S., Wang, C.D., Huang, D., Lai, J.H., Yu, P.S.: Concept factorization based multiview clustering for large-scale data. IEEE Transactions on Knowledge and Data Engineering **36**(11), 5784–5796 (2024)

附件：ICPADS 为中国计算机学会 CCF 推荐的 C 类会议论文的佐证

1.中国计算机学会推荐国际学术会议和期刊目录正式发布

https://www.ccf.org.cn/Academic_Evaluation/By_category/2023-03-08/787209.shtml

价的依据。如果由于将此《目录》作为学术评价的依据而引起纠纷，本学会不承担任何责任。

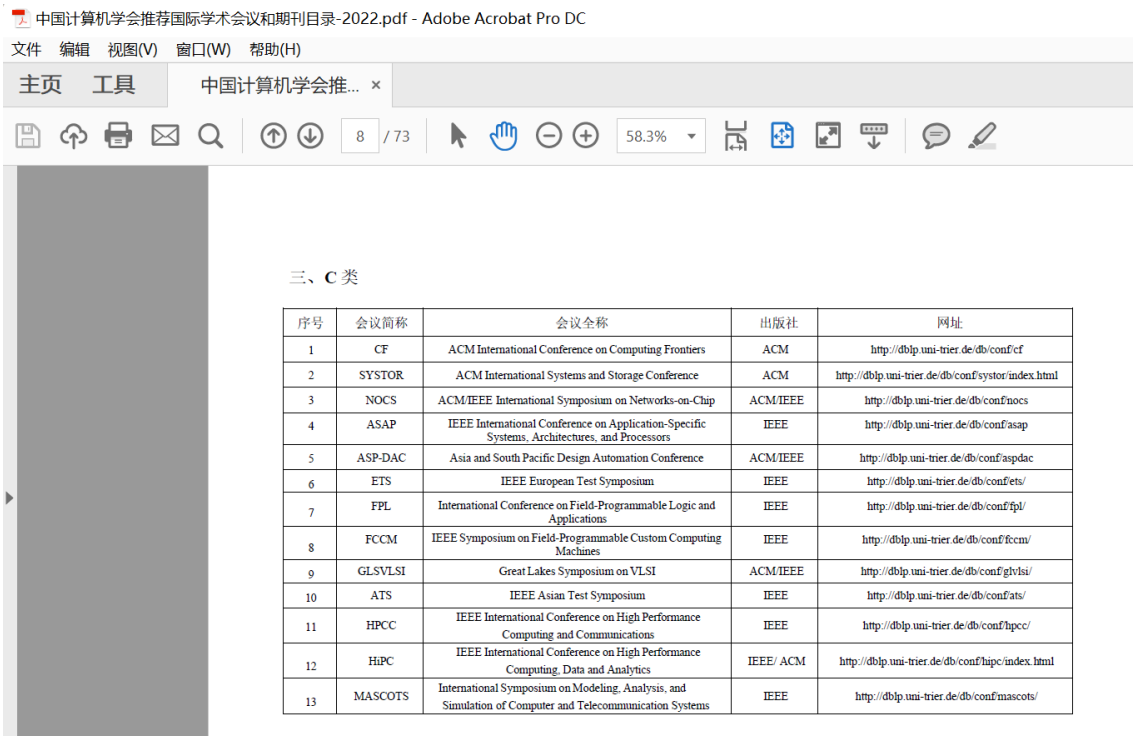
此次《目录》修订工作是CCF会员、各专业委员会以及参加初审和终审的专家们共同努力的结果，期间得到国内外众多专家学者的支持，许多专家提出了很多宝贵的建议并提供了翔实的数据或证据，在此致以诚挚的谢意。

中国计算机学会

2023年3月8日

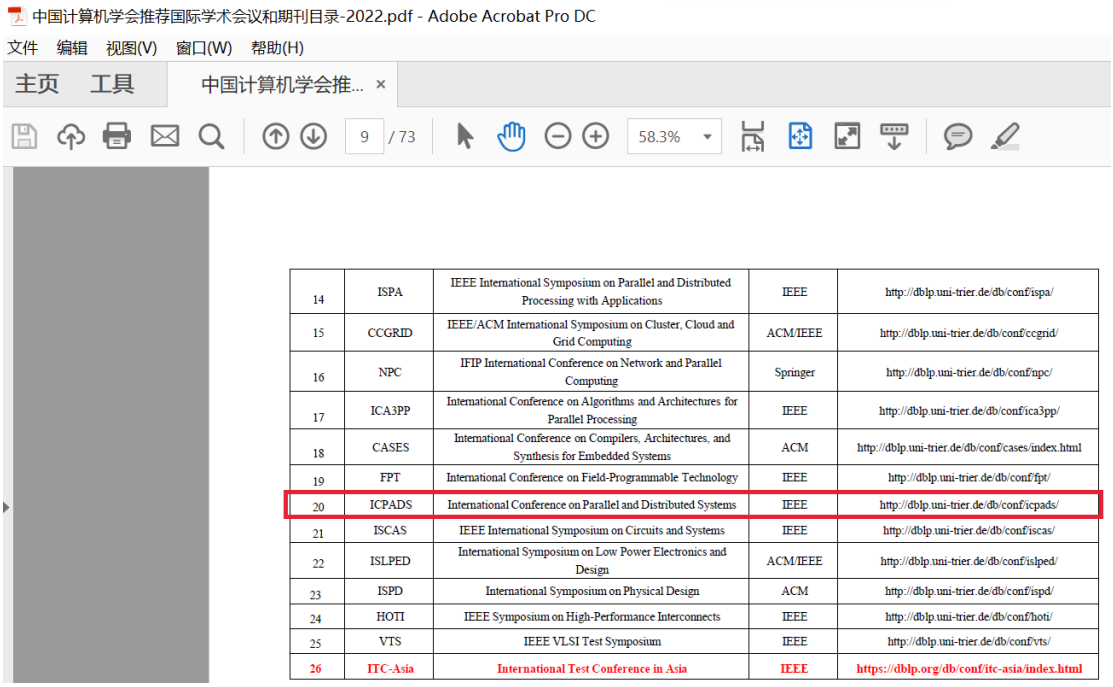
中国计算机学会推荐国际学术会议和期刊目录-2022

2. 在《中国计算机学会推荐国际学术会议和期刊目录》（CCF 推荐目录）2022 年版本中，第 9 页将 ICPADS 列为 C 类会议。



三、C 类

序号	会议简称	会议全称	出版社	网址
1	CF	ACM International Conference on Computing Frontiers	ACM	http://dblp.uni-trier.de/db/conf/cf
2	SYSTOR	ACM International Systems and Storage Conference	ACM	http://dblp.uni-trier.de/db/conf/systor/index.html
3	NOCS	ACM/IEEE International Symposium on Networks-on-Chip	ACM/IEEE	http://dblp.uni-trier.de/db/conf/nocs
4	ASAP	IEEE International Conference on Application-Specific Systems, Architectures, and Processors	IEEE	http://dblp.uni-trier.de/db/conf/asap
5	ASP-DAC	Asia and South Pacific Design Automation Conference	ACM/IEEE	http://dblp.uni-trier.de/db/conf/aspdac
6	ETS	IEEE European Test Symposium	IEEE	http://dblp.uni-trier.de/db/conf/ets/
7	FPL	International Conference on Field-Programmable Logic and Applications	IEEE	http://dblp.uni-trier.de/db/conf/fpl/
8	FCCM	IEEE Symposium on Field-Programmable Custom Computing Machines	IEEE	http://dblp.uni-trier.de/db/conf/fccm/
9	GLSVLSI	Great Lakes Symposium on VLSI	ACM/IEEE	http://dblp.uni-trier.de/db/conf/glvlsi/
10	ATS	IEEE Asian Test Symposium	IEEE	http://dblp.uni-trier.de/db/conf/ats/
11	HPCC	IEEE International Conference on High Performance Computing and Communications	IEEE	http://dblp.uni-trier.de/db/conf/hpcc/
12	HiPC	IEEE International Conference on High Performance Computing, Data and Analytics	IEEE/ ACM	http://dblp.uni-trier.de/db/conf/hipc/index.html
13	MASCOTS	International Symposium on Modeling, Analysis, and Simulation of Computer and Telecommunication Systems	IEEE	http://dblp.uni-trier.de/db/conf/mascots/



14	ISPA	IEEE International Symposium on Parallel and Distributed Processing with Applications	IEEE	http://dblp.uni-trier.de/db/conf/ispa/
15	CCGRID	IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing	ACM/IEEE	http://dblp.uni-trier.de/db/conf/ccgrid/
16	NPC	IFIP International Conference on Network and Parallel Computing	Springer	http://dblp.uni-trier.de/db/conf/npc/
17	ICA3PP	International Conference on Algorithms and Architectures for Parallel Processing	IEEE	http://dblp.uni-trier.de/db/conf/ica3pp/
18	CASES	International Conference on Compilers, Architectures, and Synthesis for Embedded Systems	ACM	http://dblp.uni-trier.de/db/conf/cases/index.html
19	FPT	International Conference on Field-Programmable Technology	IEEE	http://dblp.uni-trier.de/db/conf/fpt/
20	ICPADS	International Conference on Parallel and Distributed Systems	IEEE	http://dblp.uni-trier.de/db/conf/icpads/
21	ISCAS	IEEE International Symposium on Circuits and Systems	IEEE	http://dblp.uni-trier.de/db/conf/iscas/
22	ISLPED	International Symposium on Low Power Electronics and Design	ACM/IEEE	http://dblp.uni-trier.de/db/conf/islped/
23	ISPD	International Symposium on Physical Design	ACM	http://dblp.uni-trier.de/db/conf/ispd/
24	HOTI	IEEE Symposium on High-Performance Interconnects	IEEE	http://dblp.uni-trier.de/db/conf/hoti/
25	VTS	IEEE VLSI Test Symposium	IEEE	http://dblp.uni-trier.de/db/conf/vts/
26	ITC-Asia	International Test Conference in Asia	IEEE	https://dblp.org/db/conf/itc-asia/index.html

Confidence-oriented Contrastive Graph Clustering

Yan-Di Huang
College of Mathematics and Informatics
South China Agricultural University
Guangzhou, China
1084536492@qq.com

Guang-Yu Zhang✉,
College of Mathematics and Informatics
South China Agricultural University
Guangzhou, China
guangyuzhg@foxmail.com

Dong Huang,
College of Mathematics and Informatics
South China Agricultural University
Guangzhou, China
huangdonghere@gmail.com

Chang-Dong Wang,
School of Computer Science and Engineering
Sun Yat-sen University
Guangdong Provincial Key Laboratory of Intellectual Property and Big Data.
Guangzhou, China
changdongwang@hotmail.com

Yang Liu,
School of Computer Science and Engineering
Sun Yat-sen University
Guangzhou, China
liuy856@mail.sysu.edu.cn

Enbo Huang,
Guangxi Key Lab of Human-machine interaction and intelligent Decision
Nanning Normal University
Nanning, China
lovemusicge@163.com

2024 International Joint Conference on Neural Networks (IJCNN) | 979-8-3503-5931-2/24/$31.00 ©2024 IEEE | DOI: 10.1109/IJCNN60899.2024.10650993

***Abstract*—Contrastive clustering has recently been an emerging topic in deep unsupervised learning. Nevertheless, the previous works mostly adopt the stochastic data augmentations, which easily leads to the semantic drift problem by limited transformations. Moreover, these approaches ignore the data distribution information when generating the positive and negative pair-wise samples. In light of this, this paper proposes a simple yet effective unsupervised clustering network termed Confidence-oriented Contrastive Graph Clustering (CoCGC). Particularly, we design an end-to-end network paradigm with un-shared weights, among which a hybrid graph filter is utilized to generate two views of reliable augmentations. Guided by the non-dominated sorting theory, we further construct a confidence-oriented sample set from the latent data distribution perspective. By considering the local density and cluster distribution of the embedding representations, the discriminative sample pairs can be derived from the confidence-oriented sets in a two-view contrastive manner. Finally, a cross-view neighbor contrastive loss is devised for better exploiting the self-supervised network signals. Extensive experimental results on five benchmark datasets demonstrate the effectiveness of our method against the existing state-of-the-art deep graph clustering methods.**

***Index Terms*—Deep graph clustering, Graph filtering, Contrastive learning, High-confidence samples.**

I. Introduction

Clustering analysis plays a vital role in unsupervised learning scenario, which seeks to optimally partition the data samples into a certain number of disjoint groups. According to the prior knowledge in some specific domains, the traditional clustering approaches, like k-means-based and hierarchical-based clustering approaches typically lean on the hand-crafted features. However, this may easily lead to the sub-optimal issues when facing with the high-dimensional complex data (such as medical images or genome sequence), due to the lack of representation flexibility for extracting the latent semantic information.

To alleviate this vital issue, many efforts have been devoted towards the clustering approaches based on deep neural networks, referred to as deep clustering (DC). In terms of different perspectives, the existing deep clustering approaches can be roughly categorized into three classes, namely AutoEncoder-based approaches (AE-DC), Generative Adversarial Networks-based approaches (GAN-DC) and Graph Neural Network-based approaches (GNN-DC). Among these various approaches, the GNN-DC approaches has become the mainstream paradigm, which has proven to be a promising technique for unsupervised clustering of the graph data. Representative GNN-DC approach mainly include Deep Attentional Embedding Graph Clustering (DAEGC) [1], Structural Deep Clustering Network (SDCN) [2], Deep Graph Clustering via Proximity Generative Adversarial Network (DGC-ProGAN) [3] and Attributed graph clustering with Dual Redundancy Reduction (AGC-DRR) [4]. Although remarkable progress has been achieved in the past years, these GNN-DC approaches still face with one common drawback. That is, many of them tend to guide the unsupervised network training by some clustering losses, neglecting the sample-wise (or augmentation-wise) relationships for enhancing better clustering performance.

Inspired by this, some researchers incorporate the contrastive learning technique into GNN-DC literature, which comes to the Graph Neural Network-based Contrastive Clustering (GNN-CC). The past two years have witness the rapid

This work was supported by the NSFC (62206099, 62276277), and in part by the Science and Technology Program of Guangzhou, China (2024A04J4451). Correspond author: Guang-Yu Zhang.

progress of GNN-CC research. For example, Zhong et al. [5] presented a novel deep graph contrastive clustering approach that consists of two graph contrastive modules. Notably, this approach utilizes the Laplacian loss to capture the clustering-friendly features while inventing a graph contrastive learning strategy to explore the compact clustering assignments. Following this line, Liu et al. [6] proposed a hard sample aware network for deep graph contrastive clustering, among which a dynamic sample weighing strategy is utilized to weight both hard positive and negative sample pairs. Very recently, Sun et al. [7] presented a deep graph clustering approach that generates the deep representations via the product of the Riemannian graph convolutional nets. Within different self-supervised contrastive framework, both the positive and negative sample pairs are constructed for characterize the sample-wise relationships effectively. However, many of them use the classic graph augmentation strategies (i.e., graph diffusion, random edge shedding or attribute masking) to generate the multiple views, which may destroy the graph topology structure and easily lead to the semantic drift. Moreover, these GNN-CC approaches mostly neglect the latent distribution information (i.e., the correlation between the node representations) when generating the positive and negative pair-wise samples.

In light of this, this paper presents a new contrastive deep graph clustering approach termed Confidence-oriented Contrastive Graph Clustering (CoCGC). Under the un-shared weighted encoder paradigm, our proposed network mainly consists of three facilitated modules. In the first module, we design a hybrid graph filter to generate two views of reliable augmentations. This ensures that the hybrid graph signals (i.e., the low-and-high frequency information) could be retained from the high-order and low-order perspectives. In the second module, we further construct a confidence-oriented sample set that well adapts to the subsequent contrastive learning. Based on this, a cross-view neighbor contrastive loss is designed in the final module, so as to enhance the clustering discriminative of the network topology. The main contributions of this paper are as follows:

- This paper develops a novel contrastive deep graph clustering approach in terms of the reliable graph augmentation, the confidence-oriented set construction and the cross-view neighbor contrastive loss.
- Within the simple yet effective network structure, our proposed approach enjoys high efficiency and competitive performances.
- Extensive experiments on five benchmark datasets demonstrate the efficiency and superiority of our CoCGC approach.

II. Related Work

Deep learning has provided an advantageous tool for clustering analysis of very complex data. A variety of deep clustering approaches have been proposed in terms of different network losses, such as the losses of autoencoder network (AE-based approaches), the losses of generative adversarial network (GAN-based approaches) and the losses of Graph Neural Network (GNN-based approaches).

The pioneering works in deep clustering scenario mainly focus on the AE-based approaches, which generally update the network structure by both the reconstruction loss and some clustering loss. As an early attempt, Yang et al. [8] presented a deep clustering network that jointly models the reconstruction loss and k-means term into a unified framework. Furthermore, Guo et al. designed an improved embedding clustering network with the local structure preserved. In [9], Xie et al. proposed to map the learned features in the embedding low-dimensional space, where the KL-divergence loss is utilized to capture the distribution cluster structure. By introducing the self-expressive layer, Ji et al. [10] developed an AE-based approach for deep subspace clustering. In addition, Dizaji et al. [11] proposed the deep regularized clustering via AE embedding and relative entropy minimization. In [12], Jiang et al. made an attempt to utilize the VAE to regularize the network training. In order to follow some predefined distributions, Dilokthanakul et al. [13] proposed a deep clustering approach that based on the Gaussian mixture variational autoencoder and variational Bayes.

The GAN-based approaches seek to boost the network generalization performance by adversarial game theory. For instance, Springenberg et al. [14] proposed a categorical generative adversarial clustering network, among which the GAN and regularized information maximization (RIM) are jointly guide the network training. Following this approach, Chen et al. [15] utilized the information maximizing generative adversarial network (InfoGAN) to extract the discriminatively embedded features for deep clustering. Moreover, Yu et al. [16] performed deep clustering by adopting two dual GANs to model the real data and corresponding representations. In [17], Mukherjee et al. designed a clusterGAN network that considers the discrete-continuous mixtures for sampling noise variables.

The GNN-based approaches have become the mainstream in the present literature, whose target is to discover the graph underlying semantic information and divide the graph nodes into several clusters. In [18], Bianchi et al. proposed a deep GNN-based spectral clustering, where the message-passing mechanism is followed by the MinCutPool layer. Moving forward, Qiu et al. [19] designed a scale GNNN-based deep clustering approach, which combines the local clustering and graph neural network with theoretical guarantees. Furthermore, Ciortan et al. [20] designed a GNN-based embedding approach for scRNA-seq data clustering, where the ZINB loss, a weighted soft K-means loss and a KL (Kullback-Leibler) divergence is integrated into the combined loss. Very recently, Tsitsulin et al. [21] further devised a deep modularity networks (DMoN) to recover high quality clusters. The significant success of contrastive learning in the field of unsupervised clustering has inspired some researchers to focus on the Graph Neural Network-based Contrastive Clustering. For instance, Zhao et al. [22] presented a graph debiased contrastive learning method to reduce false negative samples. Additionally, Lee et al.

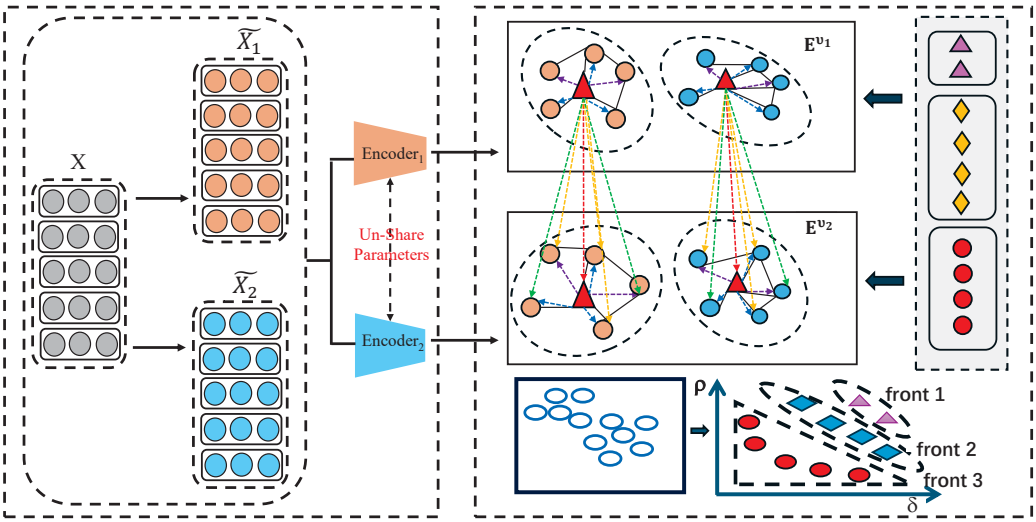


Fig. 1. The flowchart of the proposed approach.

[23] attempted to use the K-nearest-neighbor (k-NN) graph to discover the positive samples of each node representation within two-views. Besides, Liu et al. [24] introduced a dual information correlation reduction mechanism to enhance the discriminative ability of latent features, effectively alleviating the problem of representation collapse. Recently, Yang et al. [25] aimed to improve the quality of positive and negative samples by mining the high confidence clustering information.

III. Methodology

In the section, we describe the proposed CoCGC approach in detail, which enjoys the high effectiveness as well as low efficiency. Guided by the un-shared weighted paradigm, our network mainly composes of three key modules, namely hybird graph augmentation, confidence-oriented set construction and cross-view neighbor contrastive learning. In specific, the overall network structure is illustrated in Figure 1.

A. Notations

Formally, let $\mathrm{G} = (\mathrm{X}, \mathrm{A})$ be an input attributed graph data with k classes, where $\mathrm{X} \in \mathbb{R}^{n \times d}$ and $\mathrm{A} \in \{0,1\}^{n \times n}$ denote the attribute matrix and the original adjacency matrix, respectively. Accordingly, we denote the degree matrix of adjacency matrix A as $\mathrm{D} = \mathrm{diag}(d_1, d_2, \cdots, d_n) \in \mathbb{R}^{n \times n}$, among which the graph Laplacian matrix is defined as $\mathrm{L} = \mathrm{D} - \mathrm{A}$. With the renormalization trick, i.e., $\hat{\mathrm{A}} = \mathrm{A} + \mathrm{I}$ is given, the symmetric normalized graph Laplacian matrix can be denoted as $\tilde{\mathrm{L}} = \mathrm{I} - \hat{\mathrm{D}}^{-\frac{1}{2}} \hat{\mathrm{A}} \hat{\mathrm{D}}^{-\frac{1}{2}}$.

B. Hybrid gaph augmentation

In the first module, our target relies on generating two views of complementary data augmentations for strengthen deep contrastive clustering. Before getting into the network backbone, many GDNC works [6], [25] tend to utilize the low-pass filter for smoothing the attribute graph data. The corresponding formulation is given as follows:

$$\bar{\mathrm{X}} = (\mathrm{I} - \frac{\tilde{\mathrm{L}}}{2})^r \mathrm{X}, \tag{1}$$

where $(\mathrm{I} - \frac{\tilde{\mathrm{L}}}{2})$ is a low-pass graph filter and r is the filtering order. According to the signal processing theory, the low-pass frequency components have the disadvantage in aggregating the neighborhood node information. However, some existing studies indicate that the high-frequency components also play an essential role in some certain graph-specific scenarios, which can effectively detect the change abruptly and characterize discontinuities. Thereafter, we design a hybrid graph filter as follows:

$$\bar{\mathrm{X}} = [(\mathrm{I} - \frac{\tilde{\mathrm{L}}}{2})^r + \lambda(\frac{\tilde{\mathrm{L}}}{2})^r]\mathrm{X}, \tag{2}$$

where $\frac{\tilde{\mathrm{L}}}{2}$ is a high-pass filter, and parameter λ is used to balance the importance of the high- and low-pass frequency components. As can be seen, the proposed graph filter is not only able to aggregate the neighborhood node attributes, but also has the ability to retain the uniqueness of each node representation when performing the neighborhood propagation procedure.

Moving forward, two types of data augmentations can be achieved through the following formulation:

$$\begin{aligned} \bar{\mathrm{X}}_1 &= [(\mathrm{I} - \frac{\tilde{\mathrm{L}}}{2})^{r_1} + \lambda(\frac{\tilde{\mathrm{L}}}{2})^{r_1}]\mathrm{X}; \\ \bar{\mathrm{X}}_2 &= [(\mathrm{I} - \frac{\tilde{\mathrm{L}}}{2})^{r_2} + \lambda(\frac{\tilde{\mathrm{L}}}{2})^{r_2}]\mathrm{X}, \end{aligned} \tag{3}$$

where r_1 and r_2 represent the balancing parameters that corresponds to two types of data augmentations. Particularly, we set the parameter r_1 as larger integer (i.e., [7,9]) while fixing the parameter r_2 as smaller integer ([1,3]). By this means, two views of augmentations (i.e., $\bar{\mathrm{X}}_1$ and $\bar{\mathrm{X}}_2$) are

respectively generated from the high-and low-order neighborhood propagation viewpoints. Furthermore, we adopt the Siamese encoders to discover the deep semantics hidden in $\bar{\mathrm{X}}_1$ and $\bar{\mathrm{X}}_2$, which can be formulated as follows:

$$\mathrm{E}^{\mathrm{v}_1} = Encoder(\bar{\mathrm{X}}_1), \mathrm{E}^{\mathrm{v}_2} = Encoder(\bar{\mathrm{X}}_2), \tag{4}$$

where $\mathrm{E}^{\mathrm{v}_1}$ and $\mathrm{E}^{\mathrm{v}_2}$ denote the two views of node embeddings. Notice that the Siamese encoders share the same architecture but with independent parameters. After that, two views of node embeddings, i.e., $\mathrm{E}^{\mathrm{v}_1}$ and $\mathrm{E}^{\mathrm{v}_2}$ can be normalized into the following ℓ^2-norm :

$$\mathrm{E}^{\mathrm{v}_1} = \frac{\mathrm{E}^{\mathrm{v}_1}}{\|\mathrm{E}^{\mathrm{v}_1}\|_2}, \mathrm{E}^{\mathrm{v}_2} = \frac{\mathrm{E}^{\mathrm{v}_2}}{\|\mathrm{E}^{\mathrm{v}_2}\|_2}. \tag{5}$$

Finally, by incorporating the semantics of these two-view node embeddings, we have the following formulation:

$$\widetilde{\mathrm{E}} = \frac{(\mathrm{E}^{\mathrm{v}_1} + \mathrm{E}^{\mathrm{v}_2})}{2}. \tag{6}$$

Here $\widetilde{\mathrm{E}}$ represents the consistent embedding of two views.

C. Confidence-oriented set construction

In the second module, our aim is to construct the confidence-oriented set for facilitating the consequent contrastive learning. Borrowing the idea from density clustering [26] and its variants, we evaluate the confidence of each consistent embedding by two distribution metrics, namely the local density and the relatively minimum distance. To be specific, the local density ρ_i of the i-th consistent embedding $\widetilde{\mathrm{E}}_i$ is as follows:

$$\rho_i = \sum_j \chi(d_{ij} - d_c), \tag{7}$$

where d_{ij} represents the Euclidean distance between $\widetilde{\mathrm{E}}_i$ and $\widetilde{\mathrm{E}}_j$, and d_c is a pre-defined radius parameter. Note that $\chi(\cdot)$ is a cutoff function, i.e., $\chi(\mathrm{x})$ = 1 if $\mathrm{x} < 0$ and $\chi(\mathrm{x})$ = 0 otherwise. As a matter of fact, larger ρ_i means that the i-th consistent embedding $\widetilde{\mathrm{E}}_i$ is surrounded by neighbors with lower density.

Besides, the relatively minimum distance δ_{i} of the i-th consistent embedding $\widetilde{\mathrm{E}}_i$ is as follows:

$$\delta_{\mathrm{i}} = \min_{j:\rho_j > \rho_i} (d_{ij}). \tag{8}$$

Here the larger δ_i indicates that the $\widetilde{\mathrm{E}}_i$ lies at a larger distance from other consistent embedding with higher density.

When evaluating the confidence of consistent node embeddings, we should consider the local density ρ and the relatively minimum distance δ simultaneously. Thereby, these distribution metrics can be treated as two compatible objectives, which naturally leads to a non-dominated sorting problem. Particularly, we first give the definitions of Pareto improvement as follows.

Definition 1: (Pareto Improvement). *Given two different consistent node embeddings $\widetilde{\mathrm{E}}_i$ and $\widetilde{\mathrm{E}}_j$, we say $\widetilde{\mathrm{E}}_i$ is a Pareto improvement over $\widetilde{\mathrm{E}}_j$ if and only if the following condition holds,*

$$(\rho_i \geq \rho_j) \wedge (\delta_i > \delta_j) \ or \ (\rho_i > \rho_j) \wedge (\delta_i \geq \delta_j). \tag{9}$$

When $\widetilde{\mathrm{E}}_i$ is a Pareto improvement over $\widetilde{\mathrm{E}}_j$, we can also say that $\widetilde{\mathrm{E}}_i$ dominates $\widetilde{\mathrm{E}}_j$, or $\widetilde{\mathrm{E}}_j$ is dominated by $\widetilde{\mathrm{E}}_i$. The corresponding notation is as follows:

$$\widetilde{\mathrm{E}}_i \succ_{(\rho,\delta)} \widetilde{\mathrm{E}}_j \tag{10}$$

Besides, we have the following definition of Pareto set.

Definition 2: (Pareto Set). *Given a set $\Omega = \{\widetilde{\mathrm{E}}_1, \cdots, \widetilde{\mathrm{E}}_p\}$, we say Ω is a Pareto set if and only if the following condition holds,*

$$\forall \widetilde{\mathrm{E}}_i \in \Omega, \neg \exists \widetilde{\mathrm{E}}_j \in \Omega, \text{ such that } \widetilde{\mathrm{E}}_j \succ_{(\rho,\delta)} \widetilde{\mathrm{E}}_i. \tag{11}$$

In other words, we can also say that Ω is a non-dominated set. Subsequently, the overall procedure of fast non-dominated sorting algorithm is listed in Algorithm 1.

Algorithm 1 Fast non-dominated sorting

Input: The input set P that contains the consistent node embeddings $\mathrm{P} = \{\widetilde{\mathrm{E}}_1, \cdots, \widetilde{\mathrm{E}}_n\}$.
Initialization: Initialize the set $\mathrm{S}_i = \{\emptyset\}, i = 1, \cdots, n$, Initialize the set $\mathrm{F}_1 = \{\emptyset\}$. Set the counter $m_i = 0, i = 1, \cdots, n$

```
for each Ẽ_p ∈ P do
    Caculate the set S_p that contains the consistent node embeddding dominated by Ẽ_p.
    Cacluate the number of consistent node embeddding m_p that dominates Ẽ_p.
    if m_p = 0 then
        r_p = 1
        F_1 = F_1 ∪ {Ẽ_p}
    end if
end for
i =1
while F_i ≠ {∅} do
    Q = {∅}
    for each Ẽ_p ∈ F_i do
        for each Ẽ_q ∈ S_p do
            m_q = m_q − 1
            if m_q = 0 then
                r_q = i+1
                Q = Q ∪ { Ẽ_q}
            end if
        end for
        i = i+1
        F_i = Q
    end for
end while
```

Output: A series of Pareto set $\{\mathrm{F}_i\}_{i=1}$ and the front rank r_p for each consistent node embedding.

After applying the Algorithm 1 (i.e., fast non-dominated sorting), we can obtain a series of Pareto sets (also known as Pareto fronts) and the correspond Pareto rank for each consistent node embedding, as illustrated in Figure **??**. Therefore, we can choose the top τ-Pareto fronts to construct the high-confidence set, which is well-adapted to the subsequent contrastive learning.

D. Cross-view neighbor contrastive learning

In the third module, our propose lies to conduct the neighbor-guide contrastive learning by a mutually beneficial manner. Most of the current GDC approaches consider the cross-view node representation as the positive sample while regarding the other node representations as the negative samples. Despite of this, these GDC approaches fail to explore the neighborhood structure hidden in multiple views, which may destroy the homophily assumption of graph neural network. Therefore, we devise a neighbor contrastive loss that constructs the positive and negative samples by detecting the network topology as the supervised signals.

Suppose $\widetilde{\mathrm{E}}_h$ be a consistent node embedding that belongs to the co-confidence set, through which we can denote $\mathrm{E}_h^{v_1}$ and $\mathrm{E}_h^{v_2}$ as the two-view normalized embeddings of $\widetilde{\mathrm{E}}_h$, respectively. Especially, take the $\mathrm{E}_h^{v_1}$ as an example, the positive samples come from three disjoint sources: (1) The inter-view node embedding $\mathrm{E}_h^{v_2}$ in the $v2$-th view. (2) The intra-view neighbors of $\mathrm{E}_h^{v_1}$ in the $v1$-th view, i.e., $\{\mathrm{E}_k^{v_1}|\mathrm{x}_k \in \Delta_k\}$. (3) The inter-view neighbors of $\mathrm{E}_h^{v_2}$ in the v_2-th view, i.e., $\{\mathrm{E}_k^{v_2}|\mathrm{x}_k \in \Delta_k\}$. In this case, the size of positive samples is $2|\Delta_k|$ + 1, where $|\Delta_{\mathrm{k}}|$ represents the neighbor size of the h-th data sample of the affinity matrix A. Hence, the cross-view neighbor contrastive loss of the h-th $\mathrm{E}_h^{v_1}$ is formulated as follows:

$$\ell(\mathrm{E}_{\mathrm{h}}^{\mathrm{v}_1}) = -\log \frac{\frac{1}{2|\Delta_h|+1}(e^{\theta(\mathrm{E}_h^{v_1},\mathrm{E}_h^{v_2})/\tau} + \Psi_h^{v_1})}{e^{\theta(\mathrm{E}_h^{v_1},\mathrm{E}_h^{v_2})/\tau} + \sum_{g\neq h}(e^{\theta(\mathrm{E}_h^{v_1},\mathrm{E}_g^{v_1})/\tau} + e^{\theta(\mathrm{E}_h^{v_1},\mathrm{E}_g^{v_2})/\tau})}. \tag{12}$$

and

$$\Psi_h^{v_1} = \sum_{x_h\in\Delta_h}(e^{\theta(\mathrm{E}_h^{v_1},\mathrm{E}_k^{v_1})/\tau} + e^{\theta(\mathrm{E}_h^{v_1},\mathrm{E}_k^{v_2})/\tau} \tag{13}$$

where $\theta(\cdot)$ denotes the pair-wise metric of the cosine similarity, and the parameter τ is a temperature parameter.

In summary, the overall neighbor contrastive loss is formulated as follow:

$$\mathrm{L} = \frac{\sum_{i=1}^{\eta}[\ell(\mathrm{E}_h^{v_1}) + \ell(\mathrm{E}_h^{v_2})]}{2\eta}. \tag{14}$$

where η represents the size of co-confidence set. In particular, the detailed learning process of CoCGC approach is shown in Algorithm 2.

Algorithm 2 CoCGC

Input: The input feature matrix X and adjacency graph A. The iteration number T. The hyper-parameters r_1, r_2, d_c, τ, η.
Output: The clustering result R.

```
Obtain two different smoothed attribute matrix X̄1, X̄2 with Eq. (3).
for t = 1 to T do
    Encode X̄1, X̄2 into two views with parameter shared Siamese encoders via Eq. (4).
    Normalize the embeddings E^v1, E^v2 via Eq. (5).
    Fuse E^v1 and E^v2 to obtain Ẽ via Eq.(6).
    Calculate ρ and δ for each consistent node Embedding via Eq.(7) and Eq. (8).
    Perform the Algorithm 1 to select the top τ samples as the high-confidence samples.
    Compute neighbor contrastive loss L via Eq.(14).
    Update the whole network parameters via Adam optimizer .
end for
Perform K-means on Ẽ to obtain the final clustering result R.
return R.
```

IV. Experiments

In this section, we extensively evaluate the performance of our method from three aspects, i.e., comparison experiments, parameter analysis, and convergence analysis. All the experiments are implemented with an Intel 3.4-GHz CPU and 64-GB RAM.

TABLE I
THE STATISTICS OF FIVE DATASETS USED IN THE EXPERIMENTS.

Dataset	Type	Samples	Features	Edges	Classes
DBLP	Graph	4057	334	3528	4
CORA	Graph	2708	1433	5429	7
CITESEER	Graph	3327	3703	4732	6
BAT	Graph	131	81	1038	4
UAT	Graph	1190	239	13599	4

A. Benchmark Datasets

In the experimental part, extensive experiments are carried out on five widely-used datasets, including DBLP, CORA, CITESEER, BAT and UAT. To be specific, the summarized information of these datasets is listed in Table I.

B. Experiment Setup

The experiments are implemented on the desktop computer with the Intel(R) Core(TM) i9-12900K CPU, one NVIDIA GeForce RTX 3090 GPU, 128GB RAM, and the PyTorch deep learning platform. In our experiments, the number of training epochs is set as 400 while each comparison experiment is executed 10 times for fairness and stableness. The training process of our proposed network consists of two stages. In the first stage, the aim is to generate two discriminative views during the 50 epochs. In the second stage, the goal relies in constructing the co-confidence set that fits to the subsequent contrastive learning.

The learning rate is fixed as 10^{-2} for BAT dataset, while is set as 10^{-4} for DBLP,CORA,CITESEER and UAT datasets. In the first module, the designed hybrid high-low order filter has three parameters. Especially, the hyper-parameter r_1 is tuned within the range of [1, 2, 3], the hyper-parameter r_2 is tuned within the range of [5, 6, 7, 8], while the trade-off parameter λ is tuned in the range of [0.01,0.1]. Besides, we adopt four public metrics to evaluate the clustering performance of all compared approaches, involving Clustering Accuracy (ACC),

TABLE II
THE AVERAGE CLUSTERING PERFORMANCE OF TEN RUNS ON FIVE BENCHMARK DATASETS. ON EACH DATASET, THE BEST SCORE IS HIGHLIGHTED IN **BOLD**, WHILE THE SECOND BEST ONE IN [BRACKETS].

Dataset	Metric	DAEGC	ARGA	SDCN	DFCN	AGE	MVGRL	AutoSSL	AGC-DRR	DCRN	AFGRL	GDCL	ProGCL	CCGC	Ours
DBLP	ACC	62.05	64.83	68.05	76.00	75.52	42.73	0.0	[80.41]	79.66	0.0	0.0	0.0	65.06	**80.48**
	NMI	32.49	29.42	39.50	43.70	45.33	15.41	0.0	[49.77]	48.95	0.0	0.0	0.0	34.94	**50.64**
	ARI	21.03	27.99	39.15	47.00	48.08	08.22	0.0	[55.39]	53.60	0.0	0.0	0.0	35.51	**55.62**
	F1	61.75	64.97	67.71	75.70	0.0	40.52	0.0	[79.90]	79.28	0.0	0.0	0.0	62.45	**79.92**
CORA	ACC	70.43	71.04	35.60	36.33	[73.50]	70.47	63.81	40.62	61.93	26.25	70.83	57.13	**73.88**	72.51
	NMI	52.89	51.06	14.28	19.36	57.58	55.57	47.62	18.74	45.13	12.36	56.30	41.02	**56.45**	[56.15]
	ARI	49.63	47.71	07.78	04.67	[50.10]	48.70	38.92	14.80	33.15	14.32	48.05	30.71	**52.51**	49.49
	F1	68.27	69.27	24.37	26.16	[69.28]	67.15	56.42	31.23	49.50	30.20	52.88	45.68	**70.98**	68.71
CITESEER	ACC	64.54	61.07	65.96	69.50	69.73	62.83	66.76	68.32	**70.86**	31.45	66.39	65.92	69.84	[70.15]
	NMI	36.41	34.40	38.71	43.90	[44.93]	40.69	40.67	43.28	**45.86**	15.17	39.52	39.59	44.33	44.58
	ARI	37.78	34.32	40.17	[45.50]	45.31	34.18	38.73	45.34	**47.64**	14.32	41.07	36.36	45.68	45.29
	F1	62.20	58.23	63.62	64.30	64.45	59.54	58.22	[64.82]	**65.83**	30.20	61.12	57.89	62.71	62.88
BAT	ACC	52.67	67.86	53.05	55.73	56.68	37.56	42.43	47.79	67.94	50.92	45.42	55.73	[75.04]	**76.41**
	NMI	21.43	49.09	25.74	48.77	36.04	29.33	17.84	19.91	47.23	27.55	31.70	28.69	[50.23]	**53.43**
	ARI	18.18	42.02	21.04	37.76	26.59	13.45	13.11	14.59	39.76	21.89	19.33	21.84	[46.95]	**51.19**
	F1	52.23	67.02	46.45	50.90	55.07	29.64	34.84	42.33	67.40	46.53	39.94	56.08	[74.90]	**75.70**
UAT	ACC	52.29	49.31	52.25	33.61	52.37	44.16	42.52	42.64	49.92	41.50	48.70	45.38	[56.34]	**60.12**
	NMI	21.33	25.44	21.61	26.49	23.64	21.53	17.86	11.15	24.09	17.33	25.10	22.04	[28.15]	**30.03**
	ARI	20.50	16.57	21.63	11.87	20.39	17.12	13.13	09.50	17.17	13.62	21.76	14.74	[25.52]	**30.03**
	F1	50.33	50.26	45.59	25.79	50.15	39.44	34.94	35.18	44.81	36.52	45.69	39.30	[55.24]	**58.38**

Average Rand Index (ARI), Normalized Mutual Information (NMI), and macro F1-score (F1) [27]–[29].

C. Comparison Performance

In order to demonstrate the superiority of the proposed approach, we comprehensively compare our CoCGC approach against 14 baselines. On the one hand, four non-contrastive deep graph clustering approaches, i.e., DAEGC [1], ARGA [30], SDCN [2] and DFCN [31] are chosen in our experiments. On the other hand, nine advanced contrastive deep graph clustering approaches, i.e., AGE [32], MVGRL [33], AutoSSL [34], AGC-DRR [4], DCRN [24], AFGRL [23], GDCL [22], ProGCL [35] and CCGC [25] are selected to validate their effectiveness. The source code of these baselines is directly downloaded in the open access, where the corresponding hyper-parameters are tuned according to the original paper.

The comparison results over four metrics are reported in Table II. By analyzing these clustering results, we can draw the following three conclusions. First, most non-contrastive approaches have achieved comparable or even better performance over the contrastive approaches, especially on the CORA and UAT datasets. Second, the proposed approach consistently outperforms the four non-contrastive approaches by significant margins. Take the DBLP dataset as an example, it has surpassed the best competitor (i.e., DFCN) over 6.94% NMI and 8.62% ACC, respectively. Third, our proposed approach achieves the best performance except the CORA dataset. For instance, it has obtained 3.78%, 4.51%, and 3.14% improvements over ACC, ARI and F1, respectively. Overall, our CoCGC approach is always ranked at the first potion on all the benchmark datasets, which suggests the superiority of the proposed clustering network with three simple yet necessity modules.

D. Ablation Study

In this section, we verify the effect of two key modules in the proposed approach, i.e., Hybrid Filter (HF) and Co-confidence set construction (CoSC). The ablation results are recorded in Table III, where "(w/o)HF" and "(w/o)CoSC" here denote the modules of Hybrid Filter and co-confidence construction are removed, respectively. It can be seen that, our proposed approach has obtained the best clustering performance by considering both two modules. Conversely, when we merely adopt one module like "(w/o)HF" or "(w/o)CoSC", the clustering performance becomes deteriorated on all the datasets. Subsequently, the above analysis has validated that these both modules (HF and CoSC) play a crucial role in enhancing the deep clustering performance.

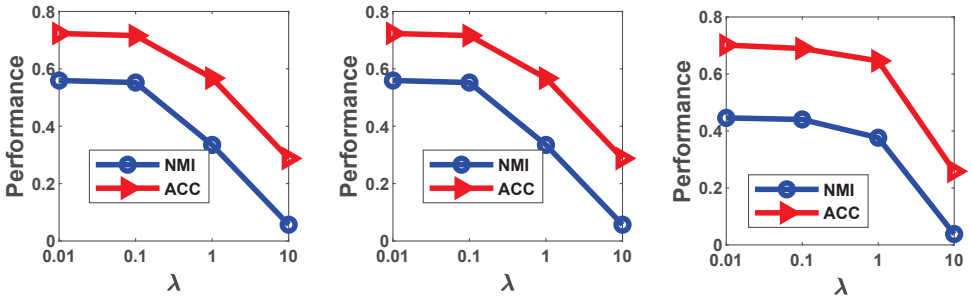
E. Sensitivity Analysis

Moving forward, we further conduct the experiments of the proposed network with different hyper-parameters setting. Notice that there are three hyper-parameters λ, r_1 and r_2 in our network. First of all, we investigate the impact of hyper-parameter λ by setting the other two hyper-parameters as static values. The clustering results with different λ (within the range of [0.01, 0.1, 1, 10]) are illustrated in Figure 2. From this figure, we can see that the network model becomes gradually decreasing when the hyper-parameter λ goes from 0.01 to 10. Especially, when we set the λ as small value (i.e., 0.01 or 0.1), our proposed approach can achieve the relatively promising performance on DBLP, CORA and CITESEER datasets.

Next, we study the clustering performance with various combinations of hyper-parameters r_1 and r_2 on BAT dataset. As can be observed in Figure 3, our approach shows relatively stale performance with different combinations of r_1 and r_2. This indicates that the proposed CoCGC approach is not very sensitive to the hyper-parameters r_1 and r_2 in various application scenarios.

TABLE III
ABLATION EXPERIMENTS.

Dataset	Model	ACC(%)	NMI(%)	ARI(%)	F1(%)
DBLP	(w/o) HF	76.96$_{\pm 0.50}$	44.49$_{\pm 0.42}$	49.53$_{\pm 0.76}$	76.27$_{\pm 0.48}$
	(w/o) CoSC	79.61$_{\pm 0.30}$	49.29$_{\pm 0.33}$	54.40$_{\pm 0.54}$	79.10$_{\pm 0.30}$
	Ours	**80.48**$_{\pm 0.41}$	**50.64**$_{\pm 0.53}$	**55.62**$_{\pm 0.76}$	**79.92**$_{\pm 0.40}$
CORA	(w/o) HF	71.23$_{\pm 0.29}$	54.64$_{\pm 0.45}$	47.95$_{\pm 0.70}$	67.66$_{\pm 0.99}$
	(w/o) CoSC	71.55$_{\pm 0.72}$	54.73$_{\pm 0.94}$	48.18$_{\pm 1.08}$	68.14$_{\pm 0.69}$
	Ours	**72.51**$_{\pm 0.40}$	**56.15**$_{\pm 0.38}$	**49.49**$_{\pm 0.91}$	**68.71**$_{\pm 0.21}$
CITESEER	(w/o) HF	68.87$_{\pm 0.48}$	43.93$_{\pm 0.62}$	44.92$_{\pm 0.73}$	61.23$_{\pm 1.01}$
	(w/o) CoSC	70.14$_{\pm 0.66}$	44.50$_{\pm 0.79}$	44.97$_{\pm 0.75}$	62.23$_{\pm 1.39}$
	Ours	**70.15**$_{\pm 0.69}$	**44.58**$_{\pm 0.57}$	**45.29**$_{\pm 0.51}$	**62.88**$_{\pm 1.54}$
BAT	(w/o) HF	74.65$_{\pm 1.22}$	51.98$_{\pm 1.12}$	48.55$_{\pm 1.08}$	74.06$_{\pm 1.30}$
	(w/o) CoSC	69.61$_{\pm 3.01}$	47.28$_{\pm 2.66}$	42.25$_{\pm 3.96}$	68.88$_{\pm 3.30}$
	Ours	**76.41**$_{\pm 0.72}$	**53.43**$_{\pm 0.91}$	**51.19**$_{\pm 1.11}$	**75.70**$_{\pm 0.85}$
UAT	(w/o) HF	57.07$_{\pm 1.54}$	29.83$_{\pm 1.66}$	29.12$_{\pm 2.65}$	55.62$_{\pm 1.74}$
	(w/o) CoSC	58.10$_{\pm 1.42}$	27.55$_{\pm 0.62}$	27.51$_{\pm 1.25}$	56.43$_{\pm 1.05}$
	Ours	**60.12**$_{\pm 0.55}$	**30.03**$_{\pm 0.65}$	**30.03**$_{\pm 0.69}$	**58.38**$_{\pm 0.93}$



(a) *DBLP* (b) *CORA* (c) *CITESEER*

Fig. 2. Sensitivity Analysis of hyper-parameter λ.

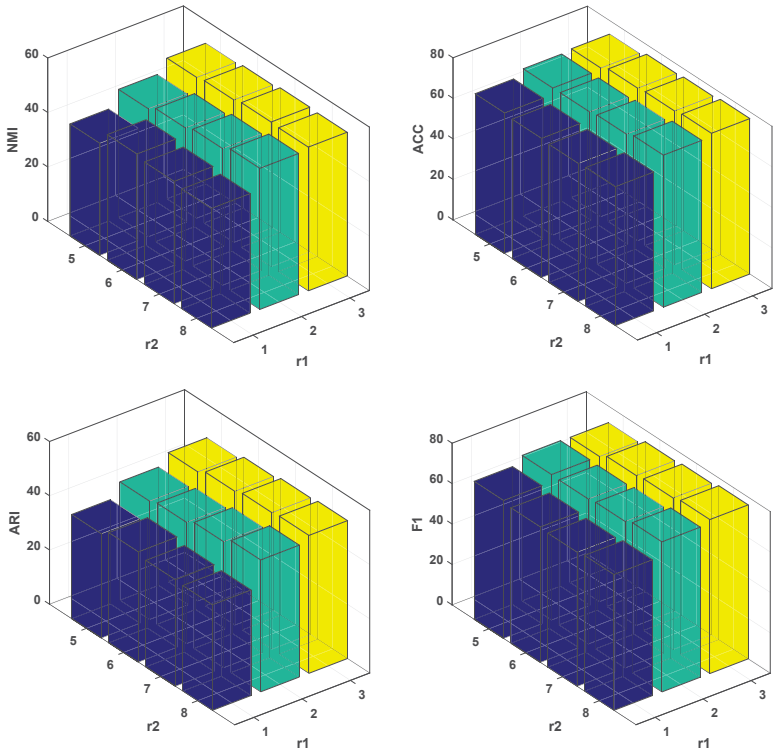


Fig. 3. Sensitivity Analysis of hyper-parameter r_1 and r_2.

V. Conclusion

This paper develops a new contrastive deep graph clustering approach termed CoCGC, which incorporates three key modules into a simple yet effective network framework. In the first module, we devise a hybrid graph filter to generate two complementary augmentations from high-and-low order viewpoints, through which two reliable views can be further achieved by an un-shared weighted Siamese encoder. In the second module, we construct a co-confident sample set that is benefited to the following contrastive learning. In the third module, we present a neighbor contrastive loss in a two-view manner, by using the network topology as the supervised signals. Extensive experiments on five datasets demonstrate the effectiveness of our proposed approach.

References

[1] C. Wang, S. Pan, R. Hu, G. Long, J. Jiang, and C. Zhang, "Attributed graph clustering: A deep attentional embedding approach," *arXiv preprint arXiv:1906.06532*, 2019.

[2] D. Bo, X. Wang, C. Shi, M. Zhu, E. Lu, and P. Cui, "Structural deep clustering network," in *Proceedings of the web conference 2020*, 2020, pp. 1400–1410.

[3] H. Gao, J. Pei, and H. Huang, "Progan: Network embedding via proximity generative adversarial network," in *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, 2019, pp. 1308–1316.

[4] L. Gong, S. Zhou, W. Tu, and X. Liu, "Attributed graph clustering with dual redundancy reduction," in *Proc. of IJCAI*, 2022.

[5] H. Zhong, J. Wu, C. Chen, J. Huang, M. Deng, L. Nie, Z. Lin, and X.-S. Hua, "Graph contrastive clustering," in *Proceedings of the IEEE/CVF international conference on computer vision*, 2021, pp. 9224–9233.

[6] Y. Liu, X. Yang, S. Zhou, X. Liu, Z. Wang, K. Liang, W. Tu, L. Li, J. Duan, and C. Chen, "Hard sample aware network for contrastive deep graph clustering," in *Proceedings of the AAAI conference on artificial intelligence*, vol. 37, no. 7, 2023, pp. 8914–8922.

[7] L. Sun, F. Wang, J. Ye, H. Peng, and P. S. Yu, "Congregate: contrastive graph clustering in curvature spaces," *Proceedings of the 32nd IJCAI*, pp. 2296–2305, 2023.

[8] B. Yang, X. Fu, N. D. Sidiropoulos, and M. Hong, "Towards k-means-friendly spaces: Simultaneous deep learning and clustering," in *international conference on machine learning*. PMLR, 2017, pp. 3861–3870.

[9] J. Xie, R. Girshick, and A. Farhadi, "Unsupervised deep embedding for clustering analysis," in *International conference on machine learning*. PMLR, 2016, pp. 478–487.

[10] P. Ji, T. Zhang, H. Li, M. Salzmann, and I. Reid, "Deep subspace clustering networks," *Advances in neural information processing systems*, vol. 30, 2017.

[11] K. Ghasedi Dizaji, A. Herandi, C. Deng, W. Cai, and H. Huang, "Deep clustering via joint convolutional autoencoder embedding and relative entropy minimization," in *Proceedings of the IEEE international conference on computer vision*, 2017, pp. 5736–5745.

[12] Z. Jiang, Y. Zheng, H. Tan, B. Tang, and H. Zhou, "Variational deep embedding: An unsupervised and generative approach to clustering," *arXiv preprint arXiv:1611.05148*, 2016.

[13] N. Dilokthanakul, P. A. Mediano, M. Garnelo, M. C. Lee, H. Salimbeni, K. Arulkumaran, and M. Shanahan, "Deep unsupervised clustering with gaussian mixture variational autoencoders," *arXiv preprint arXiv:1611.02648*, 2016.

[14] J. T. Springenberg, "Unsupervised and semi-supervised learning with categorical generative adversarial networks," *arXiv preprint arXiv:1511.06390*, 2015.

[15] X. Chen, Y. Duan, R. Houthooft, J. Schulman, I. Sutskever, and P. Abbeel, "Infogan: Interpretable representation learning by information maximizing generative adversarial nets," *Advances in neural information processing systems*, vol. 29, 2016.

[16] Z. Yu, Z. Zhang, W. Cao, C. Liu, C. P. Chen, and H.-S. Wong, "Gan-based enhanced deep subspace clustering networks," *IEEE Transactions on Knowledge and Data Engineering*, vol. 34, no. 7, pp. 3267–3281, 2020.

[17] S. Mukherjee, H. Asnani, E. Lin, and S. Kannan, "Clustergan: Latent space clustering in generative adversarial networks," in *Proceedings of the AAAI conference on artificial intelligence*, vol. 33, no. 01, 2019, pp. 4610–4617.

[18] F. M. Bianchi, D. Grattarola, and C. Alippi, "Spectral clustering with graph neural networks for graph pooling," in *International conference on machine learning*. PMLR, 2020, pp. 874–883.

[19] J. Qiu, Y. Cen, Q. Chen, C. Zhou, J. Zhou, H. Yang, and J. Tang, "Local clustering graph neural networks," 2020.

[20] M. Ciortan and M. Defrance, "Gnn-based embedding for clustering scrna-seq data," *Bioinformatics*, vol. 38, no. 4, pp. 1037–1044, 2022.

[21] A. Tsitsulin, J. Palowitch, B. Perozzi, and E. Müller, "Graph clustering with graph neural networks," *Journal of Machine Learning Research*, vol. 24, no. 127, pp. 1–21, 2023.

[22] H. Zhao, X. Yang, Z. Wang, E. Yang, and C. Deng, "Graph debiased contrastive learning with joint representation clustering." in *IJCAI*, 2021, pp. 3434–3440.

[23] N. Lee, J. Lee, and C. Park, "Augmentation-free self-supervised learning on graphs," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 36, no. 7, 2022, pp. 7372–7380.

[24] Y. Liu, W. Tu, S. Zhou, X. Liu, L. Song, X. Yang, and E. Zhu, "Deep graph clustering via dual correlation reduction," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 36, no. 7, 2022, pp. 7603–7611.

[25] X. Yang, Y. Liu, S. Zhou, S. Wang, W. Tu, Q. Zheng, X. Liu, L. Fang, and E. Zhu, "Cluster-guided contrastive graph clustering network," *arXiv preprint arXiv:2301.01098*, 2023.

[26] A. Rodriguez and A. Laio, "Clustering by fast search and find of density peaks," *science*, vol. 344, no. 6191, pp. 1492–1496, 2014.

[27] S. Zhou, X. Liu, M. Li, E. Zhu, L. Liu, C. Zhang, and J. Yin, "Multiple kernel clustering with neighbor-kernel subspace segmentation," *IEEE transactions on neural networks and learning systems*, vol. 31, no. 4, pp. 1351–1362, 2019.

[28] S. Wang, X. Liu, X. Zhu, P. Zhang, Y. Zhang, F. Gao, and E. Zhu, "Fast parameter-free multi-view subspace clustering with consensus anchor guidance," *IEEE Transactions on Image Processing*, vol. 31, pp. 556–568, 2021.

[29] L. Li, S. Wang, X. Liu, E. Zhu, L. Shen, K. Li, and K. Li, "Local sample-weighted multiple kernel clustering with consensus discriminative graph," *IEEE Transactions on Neural Networks and Learning Systems*, 2022.

[30] S. Pan, R. Hu, S.-f. Fung, G. Long, J. Jiang, and C. Zhang, "Learning graph embedding with adversarial training methods," *IEEE transactions on cybernetics*, vol. 50, no. 6, pp. 2475–2487, 2019.

[31] W. Tu, S. Zhou, X. Liu, X. Guo, Z. Cai, E. Zhu, and J. Cheng, "Deep fusion clustering network," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 35, no. 11, 2021, pp. 9978–9987.

[32] G. Cui, J. Zhou, C. Yang, and Z. Liu, "Adaptive graph encoder for attributed graph embedding," in *Proceedings of the 26th ACM SIGKDD international conference on knowledge discovery & data mining*, 2020, pp. 976–985.

[33] K. Hassani and A. H. Khasahmadi, "Contrastive multi-view representation learning on graphs," in *International conference on machine learning*. PMLR, 2020, pp. 4116–4126.

[34] W. Jin, X. Liu, X. Zhao, Y. Ma, N. Shah, and J. Tang, "Automated self-supervised learning for graphs," *arXiv preprint arXiv:2106.05470*, 2021.

[35] J. Xia, L. Wu, G. Wang, J. Chen, and S. Z. Li, "Progcl: Rethinking hard negative mining in graph contrastive learning," *arXiv preprint arXiv:2110.02027*, 2021.

附件：IJCNN 为中国计算机学会 CCF 推荐的 C 类会议论文的佐证

1.中国计算机学会推荐国际学术会议和期刊目录正式发布

https://www.ccf.org.cn/Academic_Evaluation/By_category/2023-03-08/787209.shtml

价的依据。如果由于将此《目录》作为学术评价的依据而引起纠纷，本学会不承担任何责任。

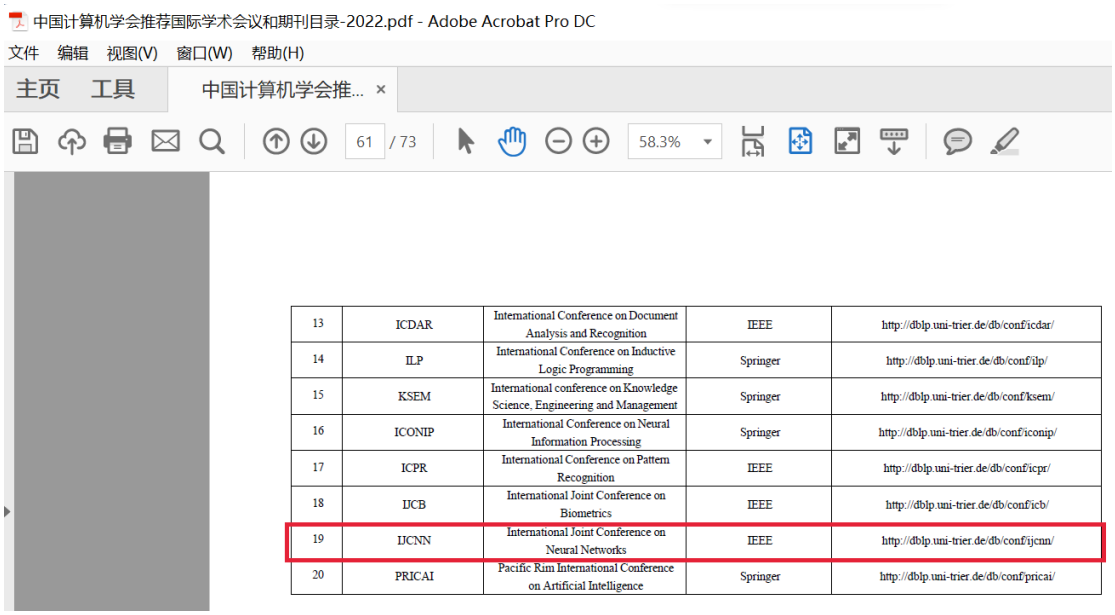
此次《目录》修订工作是CCF会员、各专业委员会以及参加初审和终审的专家们共同努力的结果，期间得到国内外众多专家学者的支持，许多专家提出了很多宝贵的建议并提供了翔实的数据或证据，在此致以诚挚的谢意。

中国计算机学会
2023年3月8日
中国计算机学会推荐国际学术会议和期刊目录-2022

2. 在《中国计算机学会推荐国际学术会议和期刊目录》（CCF 推荐目录）2022 年版本中，第 61 页将 IJCNN 列为 C 类会议。

三、C 类

序号	会议简称	会议全称	出版社	网址
1	AISTATS	International Conference on Artificial Intelligence and Statistics	JMLR	http://dblp.uni-trier.de/db/conf/aistats/
2	ACCV	Asian Conference on Computer Vision	Springer	http://dblp.uni-trier.de/db/conf/accv/
3	ACML	Asian Conference on Machine Learning	JMLR	http://dblp.uni-trier.de/db/conf/acml/
4	BMVC	British Machine Vision Conference	British Machine Vision Association	http://dblp.uni-trier.de/db/conf/bmvc/
5	NLPCC	CCF International Conference on Natural Language Processing and Chinese Computing	Springer	https://dblp.uni-trier.de/db/conf/nlpcc/
6	CoNLL	Conference on Computational Natural Language Learning	Association for Computational Linguistics	http://dblp.uni-trier.de/db/conf/conll
7	GECCO	Genetic and Evolutionary Computation Conference	ACM	http://dblp.uni-trier.de/db/conf/gecco/
8	ICTAI	IEEE International Conference on Tools with Artificial Intelligence	IEEE	http://dblp.uni-trier.de/db/conf/ictai/
9	IROS	IEEE/RSJ International Conference on Intelligent Robots and Systems	IEEE	http://dblp.uni-trier.de/db/conf/iros/
10	ALT	International Conference on Algorithmic Learning Theory	Springer	http://dblp.uni-trier.de/db/conf/alt/
11	ICANN	International Conference on Artificial Neural Networks	Springer	http://dblp.uni-trier.de/db/conf/icann/
12	FG	IEEE International Conference on Automatic Face and Gesture Recognition	IEEE	http://dblp.uni-trier.de/db/conf/fgr/



13	ICDAR	International Conference on Document Analysis and Recognition	IEEE	http://dblp.uni-trier.de/db/conf/icdar/
14	ILP	International Conference on Inductive Logic Programming	Springer	http://dblp.uni-trier.de/db/conf/ilp/
15	KSEM	International conference on Knowledge Science, Engineering and Management	Springer	http://dblp.uni-trier.de/db/conf/ksem/
16	ICONIP	International Conference on Neural Information Processing	Springer	http://dblp.uni-trier.de/db/conf/iconip/
17	ICPR	International Conference on Pattern Recognition	IEEE	http://dblp.uni-trier.de/db/conf/icpr/
18	IJCB	International Joint Conference on Biometrics	IEEE	http://dblp.uni-trier.de/db/conf/icb/
19	IJCNN	International Joint Conference on Neural Networks	IEEE	http://dblp.uni-trier.de/db/conf/ijcnn/
20	PRICAI	Pacific Rim International Conference on Artificial Intelligence	Springer	http://dblp.uni-trier.de/db/conf/pricai/

Multi-scale Multi-order Attributed Graph Clustering

Qi Zhang[1], Guangyu Zhang[1,3(✉)], Dong Huang[1], Changdong Wang[2], and Haiyan Wang[1]

[1] College of Mathematics and Informatics, South China Agricultural University, Guangzhou 510642, China
guangyuzhg@foxmail.com

[2] School of Computer Science and Engineering, Sun Yat-Sen University, Guangzhou 510006, China

[3] Ministry of Agriculture and Rural Affairs, Key Laboratory of Smart Agricultural Technology in Tropical South China, Guangzhou 510640, China

Abstract. Attributed graph clustering (AGC) is a prominent research focus in graph data exploration. Although considerable efforts have been made, there still remains two essential issues in recent studies. First, the mainstream AGC approaches mostly derive inspiration from the deep learning models, which may inherit the expensive computation burdens and unnecessary complex structures. Second, they often neglect the potential high-order relationship hidden in graph structural data. In view of this, we develops a lightweight approach termed Multi-scale Multi-order Attributed Graph Clustering (MM-AGC) with nearly linear complexity. Specifically, our approach generates a set of multi-step propagate matrices as multiple views, thereby simultaneously capturing the first-order and high-order topological relationships for subsequent clustering module. Inherited from this foundation, a linear graph filter-based model is purposed with tri-factorization guidance. To be specific, this design seamlessly leverages the linear graph autoencoder and scalable bipartite graph learning into a unified framework. By going beyond the conventional constraints, the multi-view bipartite representations are further extended with multiple diversified anchor sets, which helps to flexibly explore the hierarchical information from multi-scale perspective. Extensive experiments have been conducted to validate the effectiveness and efficiency of our MM-AGC approach against several state-of-the-art competitors.

Keywords: Linear Graph Convolution Networks · Tri-factorization · Bipartite graph learning · Multi-order Clustering

1 Introduction

Attributed graph data are broadly prevalent in a family of real-world scenarios, such as social network discovery [1], medical data diagnosis, and protein-protein interaction analysis tasks [2]. Take the neural network as an example, each Facebook user can be described by different attributed features (i.e., age, interests, location) and corresponding

D.-S. Huang et al. (Eds.): ICIC 2025, LNCS 15850, pp. 482–494, 2025.
https://doi.org/10.1007/978-981-96-9884-4_40

co-friendship graph network. Specifically, how to jointly harness the graph structure as well as node attributes from attributed graph data, has become a critical challenge for attributed graph analysis.

In recent years, attribute graph clustering (AGC) has experienced rapid development. According to the target models, attributed graph clustering methods can be broadly classified into deep learning-based and matrix factorization-based approaches. With the guidance of graph conventional networks, many deep learning-based approaches [3, 4] achieved the significant success in various fields. For instance, DAEGC [5] stands out as a classical deep attributed graph clustering approach. This approach adopts a unified self-supervised framework to performs node embedding and graph clustering in an iterative manner. SDCN [6] employs a deep network with feature reconstruction loss to obtain a cluster assignment distribution, which is subsequently used to guide the clustering process conducted by a two-layer graph convolutional network (GCN). These deep-learning based approaches have provided a prominent tool for clustering analysis of attributed graph data. However, the training procedure of deep network model may suffer from a severe computational burden.

In order to break through the computational burden, some researchers adopt the traditional matrix analysis technique to model the neighborhood graph relationships, which inspires the research of matrix-factorization-based approaches [7–9]. GCC [10] employs a lightweight GCN to generate convolved node embeddings, and performs node clustering by minimizing the discrepancy between these embeddings and their corresponding reconstructed cluster centroids. Instead of using the deep learning networks, the existing matrix factorization-based AGC approaches have efficiently boosted the computational efficiency. However, there still remain two challenging questions according to our observations. (i) How to fully exploit the high-order neighbourhood relationships from attributed graph data? (ii) How to break through the traditional practice (e.g., only focus on the specific anchor space) for better clustering performance?

In light of this, a Multi-scale Multi-order Attributed Graph Clustering (MM-AGC) approach is developed in this paper. Particularly, we first produce a series of high-order propagation matrices as multiple views, such that these views could convey rich yet complementary information to each other. With the multiple constructed views, a tri-factorization guided model is formulated with theoretical guarantee. In specific, the proposed tri-factorization framework can serve as a bridge to breakthrough the gap between two mutually beneficial modules, namely linear graph feature reduction as well as scalable bipartite graph clustering. Instead of utilizing the fixed anchor size for each view, the multiple views are further fed to the multi-scale bipartite graph paradigm, upon which the versatile yet hierarchical information can be flexibly explored. Extensive experiments have been conducted on several practical benchmarks, which have verified the robustness and efficiency of our MM-AGC approach when compared to the other state-of-the-arts. The main contributions are as follows.

This paper proposes an AGC approach from multi-scale multi-order viewpoint, which provides a brilliant insight into the existing literature.

Theoretical analysis reveals that the proposed tri-factorization paradigm has the inherent correlations with two mutually enhanced modules.

An efficient optimization algorithm is developed to solve our designed model. Experiments on five real-world benchmark datasets clarify the efficiency as well as effectiveness of the proposed MM-AGC approach.

2 The Proposed Method

2.1 Notations and Problem Definition

Formally, let $\mathcal{G} = (\mathcal{V}, \mathbf{X}, \mathbf{A})$ be an undirected graph data with k classes, where $\mathbf{X} \in \mathbb{R}^{d\times n}$ is a node features matrix and $\mathbf{A} \in \{0,1\}^{n\times n}$ is an original adjacency matrix. Accordingly, we denote the degree matrix of adjacency matrix $\mathbf{A}$ as $\mathbf{D} = \mathrm{diag}(d_1, d_2, \cdots, d_n) \in \mathbb{R}^{n\times n}$. Given the modified affinity matrix $\widetilde{\mathbf{A}} = \mathbf{A} + \beta\mathbf{I}$, the normalized graph Laplacian is defined as $\widetilde{\mathbf{L}} = \mathbf{I} - \widetilde{\mathbf{D}}^{-1}\widetilde{\mathbf{A}}$, where $\mathbf{I}$ is the identity matrix.

2.2 Multi-order Convolutional Module

GCN is essential to solving the AGC problem. Especially, the mainstream AGC approaches tend to conduct the k-order graph convolutional operation via a low-pass filter, among which the high-frequency information could be removed to generate a smooth representation. This k-order graph convolution can efficiently discover the specific neighborhood information from the attributed graph data, which, however, still fails to capture the comprehensive information across multi-order feature propagations.

Thereafter, we construct the distinct high-order propagation matrices as multiple views, so as to simultaneously capture the different multi-order neighborhood information in depth. By considering the node feature matrix $\mathbf{X}$ and the corresponding modified affinity matrix $\widetilde{\mathbf{A}}$, we can produce a set of smooth representations in what follows:

$$\begin{aligned} &\mathbf{H}_1 = \mathbf{X}(\mathbf{W}), \mathbf{H}_2 = \mathbf{X}(\mathbf{WW}), \cdots, \\ &\mathbf{H}_v = \mathbf{X}\left(\underbrace{\mathbf{W}\cdots\mathbf{W}}_{v\text{times}}\right), v = 1, 2, \cdots, V. \end{aligned} \tag{1}$$

Here $\mathbf{W} = \widetilde{\mathbf{D}}^{-1}\widetilde{\mathbf{A}}$ denotes the normalized adjacency matrix. In terms of the message propagation viewpoint, the multi-scale graph context representations $[\mathbf{H}_{(1)}; \mathbf{H}_{(2)}; \cdots; \mathbf{H}_{(v)}]$ not only characterize the local neighborhood information but also incorporate the distant high-order node features for enhancing the downstream graph clustering.

2.3 Multi-scale Multi-order Attributed Graph Clustering

Starting from the matrix factorization perspective, the data feature matrix can be factorized into an anchor graph matrix and the corresponding bipartite graph representation. Thereafter, by leveraging the scalable subspace clustering with multi-order smooth representations, we can conclude the following equations:

$$\min_{Z,\,\mathbf{F}_v,\,\mathbf{P}_v} \quad \sum\nolimits_{v=1}^{v} \left(\|\mathbf{H}_v - \mathbf{A}_v\mathbf{Z}\|_F^2 \right), \tag{2}$$

where $\mathbf{A}_v$ denotes the v-th view anchor set, and $\mathbf{Z}$ represents the common anchor representation across various views. In each view, this design help capture local relationships between smooth representations and selected anchors. We design a tri-factorization scheme that decomposes the anchor set matrix into two parts, formulated as:

$$\mathbf{A}_v \approx \mathbf{P}_v\mathbf{F}_v, \forall v. \tag{3}$$

Here $\mathbf{P}_v$ and $\mathbf{F}_v$ indicates the v-th view projection matrix and anchor set in the embedded spaces, respectively. The tri-factorization framework is defined as follows:

$$\begin{aligned} \min_{Z,\,\mathbf{F}_v,\,\mathbf{P}_v} \quad & \textstyle\sum_{v=1}^{V} \left(\|\mathbf{H}_v - \mathbf{P}_v\mathbf{F}_v\mathbf{Z}\|_F^2 + \lambda\|\mathbf{Z}\|_F^2 \right), \\ \text{s.t.} \quad & \forall v\ \mathbf{P}_v^{\mathrm{T}}\mathbf{P}^v = \mathbf{I}, \mathbf{F}_v^{\mathrm{T}}\mathbf{F}^v = \mathbf{I}, 1\mathbf{Z} = 1. \end{aligned} . \tag{4}$$

In specific, the column orthogonal constrains are used to preserve the independence among different $\mathbf{P}_v$ and $\mathbf{F}_v$ in the embedding space. Besides, the simplex constraint ensures that $\mathbf{Z}$ contains the statistic characteristics, while λ is a hyper-parameter to balance the effect of regularized term $\|\mathbf{Z}\|^2$. Notably, the proposed tri-factorization scheme possesses inherent properties, as guaranteed by the key lemma in what follows.

Lemma 1. *Assume $\mathbf{P}_v^{\mathrm{T}}\mathbf{P}_v = \mathrm{I}, \forall v$ is satisfied, then the first term in Eq. (4) could be factorized into two components, which leads to the following equation:*

$$\begin{aligned} &\sum\nolimits_{v=1}^{V} \|\mathbf{H}_v - \mathbf{P}_v\mathbf{F}_v\mathbf{Z}\|_F^2 \\ &\sum\nolimits_{v=1}^{V} \left\|\mathbf{H}_v - \mathbf{P}_v\mathbf{P}_v^{\mathrm{T}}\mathbf{H}_v\right\|_F^2 + \sum\nolimits_{v=1}^{V} \left\|\mathbf{P}_v^{\mathrm{T}}\mathbf{H}_v - \mathbf{F}_v\mathbf{Z}\right\|_F^2. \end{aligned} \tag{5}$$

Proof. To begin with, the left-hand side of Eq. (5) could be reformulated as follows:

$$\begin{aligned} &\sum\nolimits_{v=1}^{V} \|\mathbf{H}_v - \mathbf{P}_v\mathbf{F}_v\mathbf{Z}\|_F^2 \\ &= \sum\nolimits_{v=1}^{V} \|\mathbf{H}_v\|_F^2 - 2\sum\nolimits_{v=1}^{V} \mathrm{Tr}\left\{\mathbf{H}_v^{\mathrm{T}}\mathbf{P}_v\mathbf{F}_v\mathbf{Z}\right\} + \sum\nolimits_{v=1}^{V} \|\mathbf{P}_v\mathbf{F}_v\mathbf{Z}\|_F^2. \end{aligned} \tag{6}$$

Given that $\mathbf{P}_v^{\mathrm{T}}\mathbf{P}_v = \mathrm{I}$ is hold for all views, the first term on the right-hand side of Eq. (5) can be reformulated as follows:

$$\begin{aligned} &\textstyle\sum_{v=1}^{V} \left\|\mathbf{H}_v - \mathbf{P}_v\mathbf{P}_v^{\mathrm{T}}\mathbf{H}_v\right\|_F^2 \\ &= \textstyle\sum_{v=1}^{V} \|\mathbf{H}_v\|_F^2 - 2\sum_{v=1}^{V} \mathrm{Tr}\left\{\mathbf{H}_v^{\mathrm{T}}\mathbf{P}_v\mathbf{P}_v^{\mathrm{T}}\mathbf{H}_v\right\} + \sum_{v=1}^{V} \left\|\mathbf{P}_v\mathbf{P}_v^{\mathrm{T}}\mathbf{H}_v\right\|_F^2 \\ &= \textstyle\sum_{v=1}^{V} \|\mathbf{H}_v\|_F^2 - \sum_{v=1}^{V} \left\|\mathbf{P}_v\mathbf{P}_v^{\mathrm{T}}\mathbf{H}_v\right\|_F^2. \end{aligned} \tag{7}$$

Because of the quality $\sum_v \|\mathbf{F}_v\mathbf{Z}\|_F^2 = \sum_v \|\mathbf{P}_v\mathbf{F}_v\mathbf{Z}\|_F^2$ *, we can equivalently expressed the last term of Eq. (5) as follows:*

$$\begin{aligned}
&\sum_{v=1}^{V}\left\|\mathbf{P}_v^{\mathrm{T}}\mathbf{H}_v - \mathbf{F}_v\mathbf{Z}\right\|_F^2 \\
&= \sum_{v=1}^{V}\left\|\mathbf{P}_v^{\mathrm{T}}\mathbf{H}_v\right\|_F^2 - 2\sum_v \mathrm{Tr}\left\{\mathbf{H}_v^{\mathrm{T}}\mathbf{P}_v\mathbf{F}_v\mathbf{Z}\right\} + \sum_{v=1}^{V}\|\mathbf{F}_v\mathbf{Z}\|_F^2 \\
&= \sum_{v=1}^{V}\left\|\mathbf{P}_v^{\mathrm{T}}\mathbf{H}_v\right\|_F^2 - 2\sum_{v=1}^{V} \mathrm{Tr}\left\{\mathbf{H}_v^{\mathrm{T}}\mathbf{P}_v\mathbf{F}_v\mathbf{Z}\right\} + \sum_{v=1}^{V}\|\mathbf{P}_v\mathbf{F}_v\mathbf{Z}\|_F^2.
\end{aligned} \tag{8}$$

Clearly, the sum of the right-hand sides of Eq. (7) and Eq. (8) corresponds to the right-hand side of Eq. (6). Consequently, this sum is also equivalent to the left side of Eq. (5).

Notably, based on this Lemma, the target model in Eq. (4) could be further formulated as follows.

$$\begin{aligned}
&\min_{Z,\,\mathbf{F}_v,\,\mathbf{P}_v} \sum_{v=1}^{V}\left\|\mathbf{H}_v - \mathbf{P}_v\mathbf{P}_v^{\mathrm{T}}\mathbf{H}_v\right\|_F^2 + \sum_{v=1}^{V}\left\|\mathbf{P}_v^{\mathrm{T}}\mathbf{H}_v - \mathbf{F}_v\mathbf{Z}\right\|_F^2 + \lambda\|\mathbf{Z}\|_F^2 \\
&\text{s.t.}\ \forall v\ \mathbf{P}_v^{\mathrm{T}}\mathbf{P}_v = \mathbf{I},\ \mathbf{F}_v^{\mathrm{T}}\mathbf{F}_v = \mathbf{I},\ 1\mathbf{Z} = 1.
\end{aligned} \tag{9}$$

As can be seen, there are two parts in Eq. (10) that favors the different aspects of our target formulation. In specially, the first part refers to the *dimension reduction* term while the second part refers to the *scalable subspace clustering* term. Specifically, the $\mathbf{P}_v^{\mathrm{T}}\mathbf{H}_v$ term (i.e., $\mathbf{P}_v^{\mathrm{T}}\mathbf{X}\mathbf{W}\cdots\mathbf{W}$) refers to the simple graph conventional operations via v-order filters, among which the projection matrix $\mathbf{P}_v$ is actually a linear encoder in the v-th view. Therefore, these two mutually facilitative tasks in unsupervised learning, namely dimension reduction and scalable subspace clustering are jointly performed within the multi-view linear GCN paradigm.

The diversity among multiple views makes it challenging to learn a unified anchor set that fits all of them. Moreover, even with a shared anchor set, it remains difficult to effectively capture the rich and varied structural information embedded in different views. In this paper, to mitigate the instability caused by using a single anchor set across all views, we construct multiple different anchor sets for each view. Some previous studies [11] have suggested that moderate diversity can be beneficial to an ensemble system. Specifically, we provide R anchor selections for each view. Accordingly, the corresponding anchor matrices are denoted as $\{\{\mathbf{A}_v^r\}_{v=1}^V\}_{r=1}^R$, where $\mathbf{A}_v^r \in \mathbb{R}^{d\times m_r}$, and m_r denotes the number of anchors in the r-th selection. The corresponding bipartite graph representation with diverse sizes are denoted as $\{\mathbf{Z}^r\}_{r=1}^R$, where $\mathbf{Z}^r \in \mathbb{R}^{m_r\times n}$. The objective function is then formulated as follows:

$$\begin{aligned}
&\min_{\mathbf{Z}^r,\mathbf{F}_v^r,\mathbf{P}_v^r} \sum_{v=1}^{V}\sum_{r=1}^{R}\left(\left\|\mathbf{H}_v - \mathbf{P}_v^r(\mathbf{P}_v^r)^{\mathrm{T}}\mathbf{H}_v\right\|_F^2 + \left\|(\mathbf{P}_v^r)^{\mathrm{T}}\mathbf{H}_v - \mathbf{F}_v^r\mathbf{Z}^r\right\|_F^2 + \lambda\|\mathbf{Z}^r\|_F^2\right), \\
&\text{s.t.}\ \forall v\ (\mathbf{P}_v^r)^{\mathrm{T}}\mathbf{P}_v^r = \mathbf{I},\ (\mathbf{F}_v^r)^{\mathrm{T}}\mathbf{F}_v^r = \mathbf{I},\ 1\mathbf{Z}^r = 1.
\end{aligned} \tag{10}$$

Here $\mathbf{P}_v^r \in \mathbb{R}^{d\times f_r}$ and $\mathbf{F}_v^r \in \mathbb{R}^{f_r\times m_r}$ indicates the r-th anchor choice on v-th view projection matrix and anchor set in the embedded spaces, respectively.

2.4 Multi-scale Anchor Graph Fusion

After obtaining multi-scale anchor graphs from multiple views, a critical challenge still remains: how can we effectively fuse anchor graphs with varying sizes?

Traditionally, based on landmark spectral clustering and anchor graph theory, the bipartite graph representation $\mathbf{Z}^r \in \mathbb{R}^{m_r \times n}$ can be used reconstruct the full $n \times n$ affinity matrix via Eq. (11). The resultant $\mathbf{S}^r \in \mathbb{R}^{n \times n}$ is a doubly-stochastic matrix, with each row and column summing to one.

$$\mathbf{S}^r = \mathbf{Z}^{r\mathrm{T}} \sum\nolimits^{-1} \mathbf{Z}^r, \tag{11}$$

where $\sum = \mathrm{diag}(\mathbf{Z}^r 1)$. By defining $\widehat{\mathbf{Z}}^r = \sum^{-\frac{1}{2}} \mathbf{Z}^r$, the full graph can be computed as $\mathbf{S}^r = \widehat{\mathbf{Z}}^{r\mathrm{T}} \widehat{\mathbf{Z}}^r$, and the fusion graph $\overline{\mathbf{S}} = \sum_{r=1}^{R} \mathbf{S}^r$. A common approach is to perform spectral clustering on $\overline{\mathbf{S}}$ and subsequently apply k-means to the spectral embeddings to derive the clustering results. However, performing singular value decomposition (SVD) on a $n \times n$ matrix incurs a computational cost of $\mathrm{O}(n^3)$. To address this, we propose an alternative approach to directly compute the k right singular vectors of the concatenated anchor graph $\overline{\mathbf{Z}}$ defined in Eq. (12).

$$\overline{\mathbf{Z}} = \left[\widehat{\mathbf{Z}}^1; \widehat{\mathbf{Z}}^2; \cdots; \widehat{\mathbf{Z}}^R \right]. \tag{12}$$

2.5 Efficient Optimization Framework

In this subsection, we adopt the alternating iterative algorithm to minimize the overall model of our MM-AGC. To be specific, the optimization steps mainly consist of three steps, where the correspondingly details are described as below.

Optimizing $\mathbf{F}_v^r$. By fixing the variables $\mathbf{P}_v^r$, $\forall v$ and $\mathbf{Z}^r$, then the update rule w.r.t. variable $\mathbf{F}_v$ is as follows:

$$\mathbf{F}_v^r = \mathbf{U}_v \mathbf{M}_v{}^{\mathrm{T}}, \tag{13}$$

where $\mathbf{U}_v$ and $\mathbf{M}_v$ could be obtained by performing the SVD on $\mathbf{P}_v^{\mathrm{T}} \mathbf{H}_v (\mathbf{Z}^r)^{\mathrm{T}}$.

Optimizing $\mathbf{P}_v^r$. By fixing the variables $\mathbf{F}_v^r$, $\forall v$ and $\mathbf{Z}^r$, then the update rule w.r.t. variable $\mathbf{P}_v^r$ is as follows:

$$\mathbf{P}_v^r = \mathbf{B}_v \mathbf{O}_v{}^{\mathrm{T}}. \tag{14}$$

Similarity, the $\mathbf{B}_v$ and $\mathbf{O}_v$ could be obtained by performing the SVD on $\mathbf{H}_v (\mathbf{Z}^r)^{\mathrm{T}} (\mathbf{F}_v^r)^{\mathrm{T}}$.

Optimizing $\mathbf{Z}^r$. By fixing the variables $\mathbf{P}, \forall v$ and $\mathbf{F}, \forall v$, then the update rule w.r.t. variable $\mathbf{Z}^r$ is as follows:

$$\begin{aligned} &\min_{\mathbf{Z}^r} \left\| \mathbf{Z}^r - \frac{1}{1+\lambda} \sum\nolimits_{v=1}^{V} (\mathbf{F}_v^r)^{\mathrm{T}} (\mathbf{P}_v^r)^{\mathrm{T}} \mathbf{H}_v \right\|^2, \\ &\text{s.t.} 1\mathbf{Z}^r = 1, \mathbf{Z}^r \geq 0. \end{aligned} \tag{15}$$

Since every column in Eq. (15) is decoupled from each other, then we solve each sub-problem independently by using the simplex projection approach in [12].

For clarity, Algorithm 1 outlines the optimization procedure.

Algorithm 1 MM-AGC

Input: Input feature matrix $\mathbf{X}$ with adjacency matrix $\mathbf{A}$, the number of anchors k, the embedding dimension f, and the number of views V.

1: $\forall v$ Compute $\mathbf{H}_v$ by using Eq. (1);
2: $\forall v$ Initialize $\mathbf{P}_v^r$ with a randomized PCA on $\mathbf{H}_v$;
3: Initialize $\mathbf{Z}^r$ and $\mathbf{F}^r$ through a k-means on $\sum_v (\mathbf{P}_v^r)^\top \mathbf{H}_v$;
4: **While** the convergence condition is not meet **do**
5: $\quad \forall v$ Update $\mathbf{F}_v^r$ by using Eq. (13);
6: $\quad \forall v$ Update $\mathbf{P}_v^r$ by using Eq. (14);
7: $\quad$ Update $\mathbf{Z}^r$ by using Eq. (15);
8: end while
9: Construct $\overline{\mathbf{Z}}$ according to Eq. (12)

Output: The common anchor representation $\overline{\mathbf{Z}}$.

3 Experiments

3.1 Datasets Description

To validate our proposed approach, comprehensive experiments are conducted on five widely used public benchmarks, namely Cora, Citeseer, Pubmed, Wiki and Amazon Computers. The detailed descriptions w.r.t. these benchmarks are provided in Table 1.

Table 1. Dataset statistics.

Dataset	Nodes	Edges	Features	Classes
Cora [13]	2708	5429	1433	7
Citeseer [13]	3327	4732	3703	6
Pubmed [13]	19717	44338	500	3
Wiki [14]	2405	17981	4973	17
Amazon Computers [15]	13381	259159	767	10

3.2 Experimental Settings

In particular, we use three commonly-used metrics in our experiments: Accuracy (ACC), Normalized Mutual Information (NMI), and Adjusted Rand Index (ARI) [16]. For the three metrics, larger values indicate the more promising clustering results.

To ensure fairness, all experiments were conducted on the same device. For our MM-AGC approach, we set the maximum number of iterations to 10 and the reduced

dimension f to $r+1$, and the number of anchor selections R is simply set to 4, which means that the number of anchors varies from k to 4 k, where k denotes the number of target clusters. Moreover, the hyper-parameter values λ in our proposed model, is adjusted with the ranges of $\lambda \in \{10^{-3}, 10^{-1}, 10^{1}, 10^{3}, 10^{5}\}$ by reporting the best results. Note that the average results are reported within the 10 runs for each experiment.

3.3 Experimental Results and Analyses

To evaluate the clustering performance of our proposed approach, we compared it with eleven state-of-the-art approaches using multiple evaluation metrics. Specifically, Table 2 presents the experimental results, with the best performance highlighted in bold. As can be seen, baseline approaches such as AGC, DAEGC, and SDCN significantly outperform traditional clustering algorithms, including k-means and spectral clustering. This observation indicates that concurrently utilizing the double sematic information in attributed graph data, is ability of achieving more promising clustering performance. Additionally, the experimental results indicate that the proposed method consistently outperforms eleven baselines in most comparisons. Take the Citeseer dataset as an instance, our MM-AGC approach has obtained the improvements of 1.4%, 1.7%, and 3.5% over ACC, NMI, and ARI metrics, when comparing to the second best competitors.

Besides, to illustrate the efficiency of the devised integration scheme, we count its execution time. Due to space limitations, Fig. 1 presents the running time on five datasets under the optimal parameter settings. The reported runtime values exclude the costs for input (loading datasets) and output (saving clustering results).

Overall, these analysis confirm the effectiveness of the proposed approach, which integrating the tri-factorization scheme with linear multi-order GCN framework.

3.4 Visualization Analysis

Moving forward, we carry out the visual experiments by using the t-distributed Stochastic Neighbor Embedding (t-SNE) [23] algorithm. As depicted in Fig. 2, by comparing with the three baseline approaches (i.e., RAW, S^2GC, and SAGSC), our proposed approach has shown clearer and more compact clustering structures. This highlights the superior capability of MM-AGC approach in learning more accurate and distinct node representations.

Furthermore, we visualize the similarity heatmap derived from $\overline{\mathbf{S}} = \overline{\mathbf{Z}}^{\mathrm{T}}\overline{\mathbf{Z}}$, providing an intuitive representation of the high-quality fusion structure learned by our method.

The visualization result on the Cora dataset is presented in Figs. 3. As shown in the figure, the affinity matrix produced by our proposed method generally exhibits clearer block-diagonal structures compared to those generated by other competing approaches. The visualization results align well with the clustering results in Table 2, further validating the effectiveness of the proposed method.

Table 2. Clustering performance on the Cora, Citeseer, Pubmed, Wiki, and Computer datasets. The best results are marked in bold and underlined, respectively.

Method	Input	Cora			Citeseer			Pubmed			Wiki			Computer		
		Acc	NMI	ARI	Acc	NMI	ARI	Acc	NMI	ARI	Acc	NMI	ARI	Acc	NMI	ARI
k-means [17]	X	34.7	16.7	6.3	38.5	17.0	11.8	57.3	29.1	25.1	33.4	30.2	3.5	26.2	12.4	6.6
Spectral-f [18]	X	36.3	15.1	7.3	46.2	21.2	21.6	59.9	32.6	21.8	41.3	44.0	5.2	30.5	15.0	7.1
Spectral-g [18]	A	34.2	19.5	14.6	25.9	11.8	1.4	43.2	7.9	3.9	23.6	19.3	3.5	26.3	12.7	6.8
AGC [19]	A&X	65.0	52.3	41.6	67.0	41.1	41.6	N/A	N/A	N/A	46.8	43.7	14.1	59.5	46.1	27.1
DAEGC [5]	A&X	67.7	51.9	40.5	66.8	38.9	39.7	N/A	N/A	N/A	43.1	38.3	11.5	47.1	41.2	19.5
SDCN [6]	A&X	65.5	47.1	39.2	65.7	38.8	40.0	64.8	28.6	25.7	41.5	37.9	10.9	45.2	39.8	16.6
GIC [4]	A&X	72.5	53.7	50.8	68.3	44.5	46.0	65.1	26.4	24.6	44.2	44.7	28.3	46.8	47.5	31.3
S^2GC [20]	A&X	67.2	53.3	43.4	68.2	43.5	43.1	70.7	32.5	33.3	50.3	47.0	26.6	64.6	55.8	47.8
GCC [10]	A&X	74.0	58.9	50.8	69.4	45.0	45.4	70.8	32.3	33.2	53.8	**55.1**	33.3	67.2	55.8	46.1
IAGC [21]	A&X	72.4	55.9	49.3	69.2	43.2	44.5	N/A	N/A	N/A	51.6	46.8	24.2	61.5	50.3	36.8
SAGSC [22]	A&X	71.8	56.3	48.4	67.6	42.5	43.4	**71.1**	32.9	**34.1**	52.8	51.6	30.6	68.2	**57.9**	46.6
MM-AGC	A&X	**74.9**	**59.3**	**53.4**	**70.8**	**46.7**	**47.9**	71.0	**33.3**	33.7	**56.4**	53.5	**33.6**	**68.6**	54.7	**48.1**

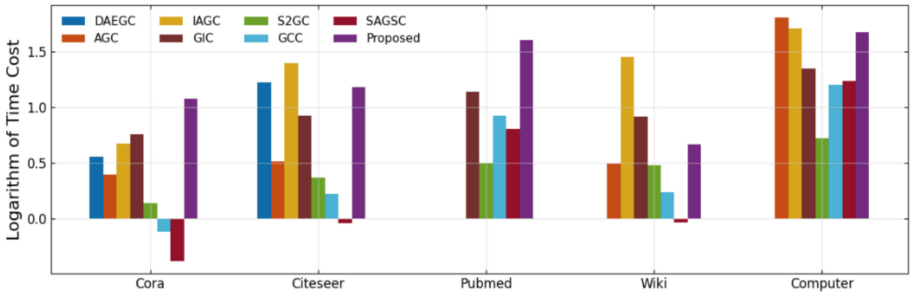


Fig. 1. Comparison of the relative logarithm running time of eight methods across five datasets. The y-axis is log-scaled to reduce the disparity among methods.

3.5 Convergence and Sensitivity

In this section, we investigate the convergence behavior of the proposed MM-AGC approach on two benchmark datasets. As illustrated in Fig. 4, the objective function values of MM-AGC decrease monotonically and converge rapidly as the number of iterations increases, demonstrating the desirable convergence property of the proposed method.

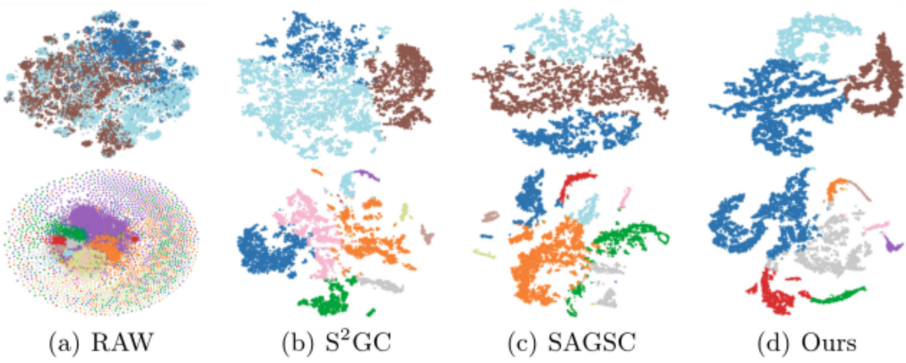


Fig. 2. 2D t-SNE visualizations of three methods on Pubmed and Computers.

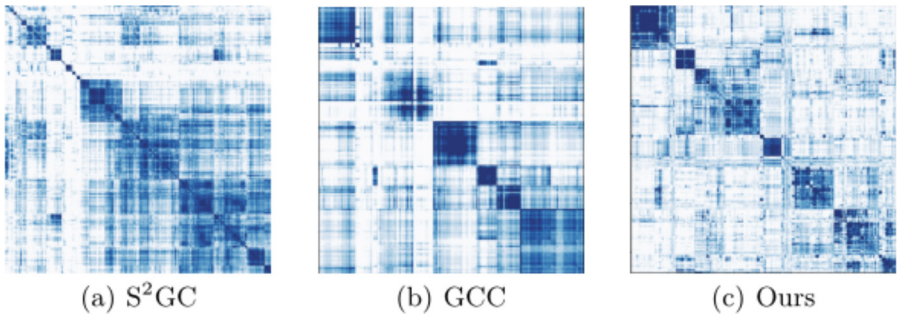


Fig. 3. Affinity matrix visualization of three methods on Cora.

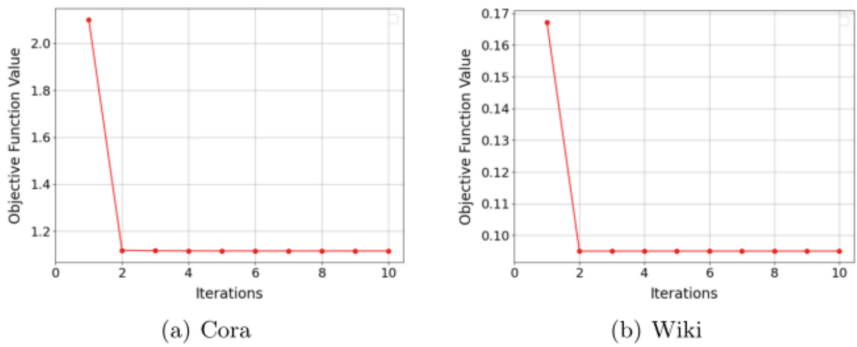


Fig. 4. Convergence of the objective function value of MM-AGC with increasing iterations.

To evaluate the parameter sensitivity of MM-AGC, we analyze the influence of the regularization parameter λ and the number of anchors R across three benchmark datasets. Specifically, Fig. 5 presents the clustering performance of MM-AGC under varying values of λ and R. The performance remains relatively stable within a reasonable range of these parameters, as observed from the results.

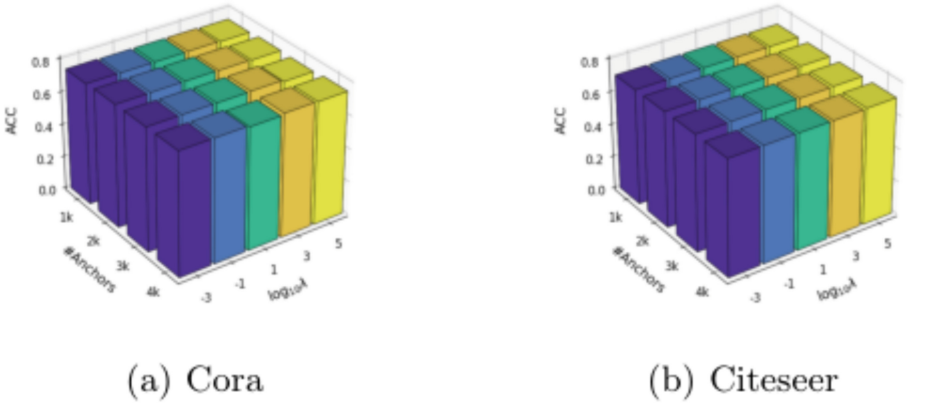


Fig. 5. ACC-based parameter sensitivity of MM-AGC on Cora and Citeseer.

3.6 Ablation Study

In this section, we conduct experiments to evaluate the effectiveness of two key modules in the MM-AGC framework: the Multi-order Convolutional module and the Multi-scale Anchor Graph Learning module. As presented in Table 3, the complete MM-AGC approach consistently achieves the best results across all four datasets. Notably, removing the Multi-order module causes the most significant performance drop on Cora, Citeseer, and Pubmed, confirming its essential role in capturing high-order neighborhood semantics. In contrast, excluding the Multi-scale module results in relatively smaller declines on these datasets, but causes a sharper drop on Wiki, demonstrating its importance in

Table 3. Ablation study on MM-AGC. The bold values represent the best results.

Dataset	Model	ACC(%)	NMI(%)	ARI(%)	F1(%)
Cora	(w/o) Multi-scale	73.8	58.0	51.5	69.7
	(w/o) Multi-order	61.5	49.4	38.3	63.4
	Ours	**74.9**	**59.3**	**53.4**	**70.6**
Citeseer	(w/o) Multi-scale	70.2	45.7	47.0	65.6
	(w/o) Multi-order	66.9	41.7	41.2	62.5
	Ours	**70.8**	**46.7**	**47.9**	**65.6**
Pubmed	(w/o) Multi-scale	70.9	33.2	33.6	70.2
	(w/o) Multi-order	61.8	33.2	29.2	61.2
	Ours	**71.0**	**33.3**	**33.7**	**70.3**
Wiki	(w/o) Multi-scale	55.2	53.4	33.0	46.0
	(w/o) Multi-order	52.8	**53.7**	32.7	45.5
	Ours	**56.4**	53.5	**33.6**	**46.6**

modeling anchor-based global structures. This indicates that combining the two modules benefits the clustering performance.

4 Conclusion

This paper proposes a lightweight AGC approach with promising effectiveness. Unlike traditional methods that rely heavily on deep networks or overlook high-order structures, our proposed approach produces multiple high-order propagation matrices as multiple views to capture richer and more diverse topological information. Furthermore, we develop a tri-factorization framework guided by linear graph filtering and introduce a multi-scale anchor graph learning strategy to effectively model both local and global relationships. In this way, both the dimension reduction and scalable subspace clustering are performed in a mutual reinforcement manner. Extensive experiments on the five benchmarks have validated the superiority of the proposed MM-AGC over several state-of-the-art methods.

Acknowledgments. This work was supported in part by the National Natural Science Foundation of China (62206099, 62276277, 62202176), the Natural Science Foundation of Guangdong Province of China (2023A1515012885), and the Science and Technology Program of Guangzhou, China (2024A04J4451).

References

1. Kim, M., Leskovec, J.: Modeling social networks with node attributes using the multiplicative attribute graph model. arXiv preprint arXiv:1106.5053 (2011)
2. Hamilton, W., Ying, Z., Leskovec, J.: Inductive representation learning on large graphs. Advances in neural information processing systems 30 (2017)
3. Karim, M.R., Beyan, O., Zappa, A., Costa, I.G., Rebholz-Schuhmann, D., Cochez, M., Decker, S.: Deep learning-based clustering approaches for bioinformatics. Brief-
4. Mavromatis, C., Karypis, G.: Graph infoclust: maximizing coarse-grain mutual information in graphs. In: Pacific-Asia Conference on Knowledge Discovery and Data Mining, pp. 541–553. Springer (2021)
5. Wang, C., Pan, S., Hu, R., Long, G., Jiang, J., Zhang, C.: Attributed graph clustering: A deep attentional embedding approach. arXiv preprint arXiv:1906.06532 (2019)
6. Bo, D., Wang, X., Shi, C., Zhu, M., Lu, E., Cui, P.: Structural deep clustering network. In: Proceedings of the Web Conference 2020, pp. 1400–1410 (2020)
7. Fettal, C., Labiod, L., Nadif, M.: Simultaneous linear multi-view attributed graph representation learning and clustering. In: Proceedings of the Sixteenth ACM International Conference on Web Search and Data Mining, pp. 303–311 (2023)
8. Zhang, G.Y., Guan, C.B., Huang, D., Wen, Z., Wang, C.D., Xiao, L.: Scalable tri-factorization guided multi-view subspace clustering. Knowledge-Based Systems p. 113119 (2025)
9. Chen, M.S., Wang, C.D., Huang, D., Lai, J.H., Yu, P.S.: Concept factorization based multiview clustering for large-scale data. IEEE Trans. Knowl. Data Eng. (2024)
10. Fettal, C., Labiod, L., Nadif, M.: Efficient graph convolution for joint node representation learning and clustering. In: Proceedings of the Fifteenth ACM International Conference on Web Search and Data Mining, pp. 289–297 (2022)

11. Hadjitodorov, S.T., Kuncheva, L.I., Todorova, L.P.: Moderate diversity for better cluster ensembles. Information Fusion **7**(3), 264–275 (2006)
12. Nie, F., Wang, X., Huang, H.: Clustering and projected clustering with adaptive neighbors. In: Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 977–986 (2014)
13. Sen, P., Namata, G., Bilgic, M., Getoor, L., Galligher, B., Eliassi-Rad, T.: Collective classification in network data. AI Mag. **29**(3), 93 (2008)
14. Yang, C., Liu, Z., Zhao, D., Sun, M., Chang, E.Y.: Network representation learning with rich text information. In: IJCAI, vol. 2015, pp. 2111–2117 (2015)
15. Shchur, O., Mumme, M., Bojchevski, A., Günnemann, S.: Pitfalls of graph neural network evaluation. arXiv preprint arXiv:1811.05868 (2018)
16. Hubert, L., Arabie, P.: Comparing partitions. Journal of classification **2**, 193–218 (1985)
17. Hartigan, J.A., Wong, M.A.: Algorithm as 136: a k-means clustering algorithm. Journal of the royal statistical society. series c (applied statistics) **28**(1), 100–108 (1979)
18. Ng, A., Jordan, M., Weiss, Y.: On spectral clustering: analysis and an algorithm. Advances in neural information processing systems 14 (2001)
19. Zhang, X., Liu, H., Li, Q., Wu, X.M.: Attributed graph clustering via adaptive graph convolution. arXiv preprint arXiv:1906.01210 (2019)
20. Zhu, H., Koniusz, P.: Simple spectral graph convolution. In: International conference on learning representations (2021)
21. Zhang, X., Liu, H., Li, Q., Wu, X.M., Zhang, X.: Adaptive graph convolution methods for attributed graph clustering. IEEE Trans. Knowl. Data Eng. **35**(12), 12384–12399 (2023)
22. Fettal, C., Labiod, L., Nadif, M.: Scalable attributed-graph subspace clustering. In: Proceedings of the AAAI Conference on Artificial Intelligence, vol. 37, pp. 7559–7567 (2023)
23. Van der Maaten, L., Hinton, G.: Visualizing data using t-sne. J. Mach. Learn. Res. **9**(11) (2008)

附件：ICIC 为中国计算机学会 CCF 推荐的 C 类会议论文的佐证

1.中国计算机学会推荐国际学术会议和期刊目录正式发布

https://www.ccf.org.cn/Academic_Evaluation/By_category/2023-03-08/787209.shtml

价的依据。如果由于将此《目录》作为学术评价的依据而引起纠纷，本学会不承担任何责任。

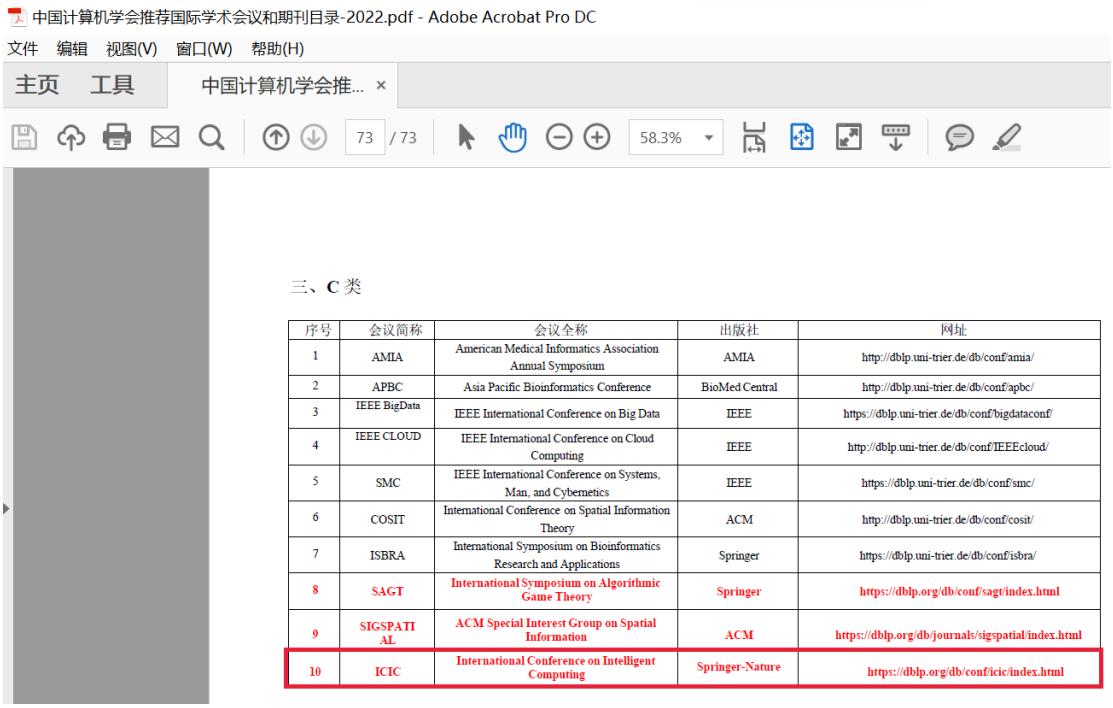
此次《目录》修订工作是CCF会员、各专业委员会以及参加初审和终审的专家们共同努力的结果，期间得到国内外众多专家学者的支持，许多专家提出了很多宝贵的建议并提供了翔实的数据或证据，在此致以诚挚的谢意。

中国计算机学会

2023年3月8日

中国计算机学会推荐国际学术会议和期刊目录-2022

2. 在《中国计算机学会推荐国际学术会议和期刊目录》（CCF 推荐目录）2022 年版本中，第 73 页将 ICIC 列为 C 类会议。



三、C 类

序号	会议简称	会议全称	出版社	网址
1	AMIA	American Medical Informatics Association Annual Symposium	AMIA	http://dblp.uni-trier.de/db/conf/amia/
2	APBC	Asia Pacific Bioinformatics Conference	BioMed Central	http://dblp.uni-trier.de/db/conf/apbc/
3	IEEE BigData	IEEE International Conference on Big Data	IEEE	https://dblp.uni-trier.de/db/conf/bigdataconf/
4	IEEE CLOUD	IEEE International Conference on Cloud Computing	IEEE	http://dblp.uni-trier.de/db/conf/IEEEcloud/
5	SMC	IEEE International Conference on Systems, Man, and Cybernetics	IEEE	https://dblp.uni-trier.de/db/conf/smc/
6	COSIT	International Conference on Spatial Information Theory	ACM	http://dblp.uni-trier.de/db/conf/cosit/
7	ISBRA	International Symposium on Bioinformatics Research and Applications	Springer	https://dblp.uni-trier.de/db/conf/isbra/
8	SAGT	International Symposium on Algorithmic Game Theory	Springer	https://dblp.org/db/conf/sagt/index.html
9	SIGSPATIAL	ACM Special Interest Group on Spatial Information	ACM	https://dblp.org/db/journals/sigspatial/index.html
10	ICIC	International Conference on Intelligent Computing	Springer-Nature	https://dblp.org/db/conf/icic/index.html

蓝桥杯大赛

获奖证书

华南农业大学张广煜：

指导温鑫荣获第十五届蓝桥杯全国软件和信息技术专业人才大赛全国总决赛C/C++程序设计大学A组三等奖，被评为优秀指导教师。

特发此证，以资鼓励。

证书编号：011525134

证件号码：440106199105151814

工业和信息化部
人才交流中心

蓝桥杯大赛组委会

2024年6月2日

蓝桥杯大赛

获奖证书

华南农业大学张广煜：

指导郑汝酬荣获第十五届蓝桥杯全国软件和信息技术专业人才大赛广东赛区C/C++程序设计大学B组三等奖，被评为优秀指导教师。

特发此证，以资鼓励。

证书编号：021549428

证件号码：440106199105151814

工业和信息化部
人才交流中心

蓝桥杯大赛组委会

2024年4月29日

蓝桥杯大赛

获奖证书

华南农业大学张广煜：

指导温鑫荣获第十五届蓝桥杯全国软件和信息技术专业人才大赛广东赛区C/C++程序设计大学A组一等奖，被评为优秀指导教师。

特发此证，以资鼓励。

证书编号：011514351

证件号码：440106199105151814

工业和信息化部
人才交流中心

蓝桥杯大赛组委会

2024年4月29日

蓝桥杯大赛

获奖证书

华南农业大学张广煜：

指导温鑫荣获第十六届蓝桥杯全国软件和信息技术专业人才大赛全国总决赛C/C++程序设计大学B组二等奖，被评为优秀指导教师。

特发此证，以资鼓励。

证书编号：1602114802

证件号码：440106199105151814

工业和信息化部
人才交流中心

蓝桥杯大赛组委会

2025年6月23日

蓝桥杯大赛

获奖证书

华南农业大学张广煜：

指导李泽扬荣获第十六届蓝桥杯全国软件和信息技术专业人才大赛全国总决赛Python程序设计大学B组三等奖，被评为优秀指导教师。

特发此证，以资鼓励。

证书编号：1633031579

证件号码：440106199105151814

工业和信息化部
人才交流中心

蓝桥杯大赛组委会

2025年6月23日

蓝桥杯大赛

获奖证书

华南农业大学张广煜：

指导杨生源荣获第十六届蓝桥杯全国软件和信息技术专业人才大赛全国总决赛Python程序设计大学B组三等奖，被评为优秀指导教师。

特发此证，以资鼓励。

证书编号：1633031664

证件号码：440106199105151814

工业和信息化部
人才交流中心

蓝桥杯大赛组委会

2025年6月23日

蓝桥杯大赛

获奖证书

华南农业大学张广煜：

指导曾志华荣获第十六届蓝桥杯全国软件和信息技术专业人才大赛广东赛区C/C++程序设计大学B组三等奖，被评为优秀指导教师。

特发此证，以资鼓励。

证书编号：1602057667

证件号码：440106199105151814

工业和信息化部
人才交流中心

蓝桥杯大赛组委会

2025年5月26日

蓝桥杯大赛

获奖证书

华南农业大学张广煜：

指导温鑫荣获第十六届蓝桥杯全国软件和信息技术专业人才大赛广东赛区C/C++程序设计大学B组一等奖，被评为优秀指导教师。

特发此证，以资鼓励。

证书编号：1602058020

证件号码：440106199105151814

工业和信息化部
人才交流中心

蓝桥杯大赛组委会

2025年5月26日

蓝桥杯大赛

获奖证书

华南农业大学张广煜：

指导卢信博荣获第十六届蓝桥杯全国软件和信息技术专业人才大赛广东赛区Python程序设计大学B组二等奖，被评为优秀指导教师。

特发此证，以资鼓励。

证书编号：1633016052

证件号码：440106199105151814

工业和信息化部
人才交流中心

蓝桥杯大赛组委会

2025年5月26日

蓝桥杯大赛

获奖证书

华南农业大学张广煜：

指导周源彬荣获第十六届蓝桥杯全国软件和信息技术专业人才大赛广东赛区Python程序设计大学B组二等奖，被评为优秀指导教师。

特发此证，以资鼓励。

证书编号：1633016173

证件号码：440106199105151814

工业和信息化部
人才交流中心

蓝桥杯大赛组委会

2025年5月26日

蓝桥杯大赛

获奖证书

华南农业大学张广煜：

指导李泽扬荣获第十六届蓝桥杯全国软件和信息技术专业人才大赛广东赛区Python程序设计大学B组一等奖，被评为优秀指导教师。

特发此证，以资鼓励。

证书编号：1633016174

证件号码：440106199105151814

工业和信息化部
人才交流中心

蓝桥杯大赛组委会

2025年5月26日

蓝桥杯大赛

获奖证书

华南农业大学张广煜：

指导杨生源荣获第十六届蓝桥杯全国软件和信息技术专业人才大赛广东赛区Python程序设计大学B组一等奖，被评为优秀指导教师。

特发此证，以资鼓励。

证书编号：1633016031

证件号码：440106199105151814

工业和信息化部
人才交流中心

蓝桥杯大赛组委会

2025年5月26日

张广煜获得广东省第二届计算机科学青年学术秀“三等奖”

1. 中国计算机学会网站报道主要内容截图

https://www.ccf.org.cn/Member_Activities/2021-10-12/745144.shtml

您的位置：首页 > 会员活动

青年学者风采飞扬，领域专家金语点评：广东省第二届计算机科学青年学者学术秀

阅读量:426　2021-10-12　收藏本文

2021年10月10日，由中国计算机学会（CCF）主办、CCF广州会员活动中心和CCF YOCSEF 广州学术委员会联合承办的“广东省第二届计算机科学青年学者学术秀”在华南理工大学大学城中心酒店南华厅顺利举办。华南理工大学朱敏副校长，欧洲科学院外籍院士、清华大学朱文武教授，广东省计算机学会理事长、华南理工大学韩国强教授，CCF专委工作委员会副主任、长江学者、国家杰青、中山大学肖侬教授，CCF理事、华南师范大学汤庸教授，国家杰青、暨南大学翁健副校长，长江学者、复旦大学计算机科学技术学院院长姜育刚教授，CCF广州主席、华南理工大学许勇教授，广州计算机学会理事长、华南师范大学范冰冰教授，暨南大学珠海校区管委会副主任方俊彬教授，CCF广州秘书长、华南师范大学郝天永教授，CCF YOCSEF广州现任主席、新华三集团吴一冰总监，CCF YOCSEF广州往届主席谭台哲教授、黄书强教授、王昌栋副教授，CCF广州委员陈平华教授、黄琼教授、李红洲经理，CCF YOCSEF广州 AC委员黄栋副教授、谢光强教授、曾衍瀚副教授、蓝连涛博士，华南理工大学詹志辉教授，广东科技报社刘雷副社长等莅临参加。此次活动吸引了来自香港科技大学（广州）、香港中文大学（深圳）、复旦大学、中山大学、华南理工大学、暨南大学、华南师范大学、广东工业大学、华南农业大学、广东外语外贸大学、广东财经大学、深圳大学、汕头大学等省内二十多所知名高校的80余人欢聚一堂、座无虚席，广州市恒联计算机科技有限公司提供赞助。

三等奖颁奖现场

附获奖名单:

一等奖

周池，深圳大学

陈颖聪，香港科技大学(广州)

胡建芳，中山大学

二等奖

樊继聪，香港中文大学(深圳)

刘竹琳，华南理工大学

甘文生，暨南大学

杨振国，广东工业大学

王海燕，华南农业大学

三等奖

孟献兵，华南理工大学

白文，华南师范大学

李佩雅，暨南大学

赵志丹，汕头大学

李诩琦，华南师范大学

张广煜，华南农业大学

姜思羽，广东外语外贸大学

刘宁，广东外语外贸大学

2. 颁发的三等奖牌匾

荣誉证书

张广煜 同志：

在2022年度 科技 工作中成绩突出，授予“十佳工作者”称号。特发此证，以资鼓励。

华南农业大学数学与信息学院 软件学院

2022年12月

蓝桥杯大赛

获奖证书

华南农业大学张广煜：

指导温鑫荣获第十五届蓝桥杯全国软件和信息技术专业人才大赛全国总决赛C/C++程序设计大学A组三等奖，被评为优秀指导教师。

特发此证，以资鼓励。

证书编号：011525134

证件号码：440106199105151814

工业和信息化部
人才交流中心

蓝桥杯大赛组委会

2024年6月2日

蓝桥杯大赛

获奖证书

华南农业大学张广煜：

指导郑汝酬荣获第十五届蓝桥杯全国软件和信息技术专业人才大赛广东赛区C/C++程序设计大学B组三等奖，被评为优秀指导教师。

特发此证，以资鼓励。

证书编号：021549428

证件号码：440106199105151814

工业和信息化部
人才交流中心

蓝桥杯大赛组委会

2024年4月29日

蓝桥杯大赛

获奖证书

华南农业大学张广煜：

指导温鑫荣获第十五届蓝桥杯全国软件和信息技术专业人才大赛广东赛区C/C++程序设计大学A组一等奖，被评为优秀指导教师。

特发此证，以资鼓励。

证书编号：011514351

证件号码：440106199105151814

工业和信息化部
人才交流中心

蓝桥杯大赛组委会

2024年4月29日

数学与信息学院 软件学院文件

关于公示数学与信息学院　软件学院教职工2024年年度考核结果的通知

各系（中心、室）、各教职工：

根据学校《关于做好2024年年度考核工作的通知》《数学与信息学院　软件学院2024年年度考核工作方案》等文件精神，学院业已完成教职工2024年年度考核工作。现将《数学与信息学院　软件学院教职工2024年年度考核结果》（详见附件）予以公示。

在公示时间内，个人和单位均可通过来信、来电、来访等形式，向学院反映公示对象的情况和问题。以个人名义反映的，提倡签署或自报真实姓名；以单位名义反映的，须加盖单位印章。反映公示对象的情况和问题，应坚持实事求是原则，不得借机诽谤、诬告。

公示时间：2025年2月24日至26日

受理单位：华南农业大学数学与信息学院　软件学院

地　　址：华南农业大学数学与信息学院　软件学院党政办公室

邮政编码：510642

联系电话：020-85285393

电子邮箱：xxxy@scau.edu.cn

数学与信息学院 软件学院

2025年2月24日

附件

数学与信息学院　软件学院教职工 2024 年年度考核结果

1. 优秀等次人员（30 人）：

陈　羽、史晨阳、江雪萍、周　燕、朱艳科、杨德贵、刘木伙、李泽华、刘曼莉、魏　佳、张胜祥、王金凤、邱少健、黄沛杰、高月芳、涂淑琴、陈湘骥、徐禹洪、王海燕、杨　磊、张广煜、王春桃、张素敏、熊俊涛、韦婷婷、张建桃、张勇威、韩方珍、左晓丽、温奕航

2. 合格等次人员（169 人）：

张伟峰、杨志程、王石安、肖　莉、李海绸、王雪琴、聂笃宪、袁利国、雷春林、张　昕、金玲玉、方　平、徐小红、曹　静、梁茹冰、李　朗、郑佳悦、周裕中、蔡贤资、曾庆茂、邱　华、危苏婷、梁艳影、李　倩、刘　丹、夏英俊、王　霞、毛卫华、林利云、姚焕城、胡巧怡、卢建平、廖　彬、朱玲湘、丁仕虹、利小玲、刘鹏飞、朱艳丽、陈银辉、张连宽、李娇娇、黄小虎、杨光锐、彭泓毅、谌秋辉、张　娜、李　凤、罗志坚、赵　峰、何　婧、郭　军、郭子君、田秀蓉、陆　琪、马立华、尤雪莲、刘文琰、黄一德、梁　云、尹　令、肖德琴、崔金荣、黄立峰、马　莎、彭超达、周子涵、李西明、杨秋妹、冯健昭、张慧玲、郭玉彬、王　栋、张丽霞、殷建军、张足生、王国华、林毅申、彭红星、吴理华、李康顺、王美华、张　猜、成明月、林旭东、

祝胜林、周　敏、刘　心、钟德祥、罗浩宇、肖　磊、钟灿琨、李双娟、边　山、司徒浩臻、彭利民、石　方、邝颖杰、黄　玲、万　华、姚金涛、朱梅阶、綦　羽、刘汉兴、陈　琰、蓝连涛、唐德玉、梁早清、张义青、毛宜军、郭艾侠、项　灿、林壮壁、宋鸿陟、司国东、张春玲、田绪红、李宏博、陈　霓、王　琴、方凤美、孙微微、彭明明、朱　凯、邢仲璟、陈浩磊、万军洲、张大斌、黄文玲、邓　金、张建军、胡焕玲、宋　歌、张　凯、古万荣、方永美、黄丽清、陈炜颖、杨振刚、杨　春、毛小娟、骆　威、刘昌余、余平祥、吴春胤、宋世领、刘伟章、曹　咏、邓成剑、肖媚燕、文　斌、曹　维、周运华、岑冠军、肖克辉、欧阳庆、李舜鹏、喻财铃、陈　思、余丹华、陆剑华、张佳午、叶增柯、严　颖、徐　妤、张俊韬、刘雨丝、袁蓓缅、杨　莹、张秋燕

3.博士后合格等次人员（2人）：

刘又夫、温自豪

蓝桥杯大赛

获奖证书

华南农业大学张广煜：

指导李泽扬荣获第十六届蓝桥杯全国软件和信息技术专业人才大赛全国总决赛Python程序设计大学B组三等奖，被评为优秀指导教师。

特发此证，以资鼓励。

证书编号：1633031579

证件号码：440106199105151814

工业和信息化部
人才交流中心

蓝桥杯大赛组委会

2025年6月23日

蓝桥杯大赛

获奖证书

华南农业大学张广煜：

指导温鑫荣获第十六届蓝桥杯全国软件和信息技术专业人才大赛全国总决赛C/C++程序设计大学B组二等奖，被评为优秀指导教师。

特发此证，以资鼓励。

证书编号：1602114802

证件号码：440106199105151814

工业和信息化部
人才交流中心

蓝桥杯大赛组委会

2025年6月23日

蓝桥杯大赛

获奖证书

华南农业大学张广煜：

指导杨生源荣获第十六届蓝桥杯全国软件和信息技术专业人才大赛全国总决赛Python程序设计大学B组三等奖，被评为优秀指导教师。

特发此证，以资鼓励。

证书编号：1633031664

证件号码：440106199105151814

工业和信息化部
人才交流中心

蓝桥杯大赛组委会

2025年6月23日

蓝桥杯大赛

获奖证书

华南农业大学张广煜：

指导曾志华荣获第十六届蓝桥杯全国软件和信息技术专业人才大赛广东赛区C/C++程序设计大学B组三等奖，被评为优秀指导教师。

特发此证，以资鼓励。

证书编号：1602057667

证件号码：440106199105151814

工业和信息化部
人才交流中心

蓝桥杯大赛组委会

2025年5月26日

蓝桥杯大赛

获奖证书

华南农业大学张广煜：

指导李泽扬荣获第十六届蓝桥杯全国软件和信息技术专业人才大赛广东赛区Python程序设计大学B组一等奖，被评为优秀指导教师。

特发此证，以资鼓励。

证书编号：1633016174

证件号码：440106199105151814

工业和信息化部
人才交流中心

蓝桥杯大赛组委会

2025年5月26日

蓝桥杯大赛

获奖证书

华南农业大学张广煜：

指导卢信博荣获第十六届蓝桥杯全国软件和信息技术专业人才大赛广东赛区Python程序设计大学B组二等奖，被评为优秀指导教师。

特发此证，以资鼓励。

证书编号：1633016052

证件号码：440106199105151814

工业和信息化部
人才交流中心

蓝桥杯大赛组委会

2025年5月26日

蓝桥杯大赛

获奖证书

华南农业大学张广煜：

指导温鑫荣获第十六届蓝桥杯全国软件和信息技术专业人才大赛广东赛区C/C++程序设计大学B组一等奖，被评为优秀指导教师。

特发此证，以资鼓励。

证书编号：1602058020

证件号码：440106199105151814

工业和信息化部
人才交流中心

蓝桥杯大赛组委会

2025年5月26日

蓝桥杯大赛

获奖证书

华南农业大学张广煜：

指导杨生源荣获第十六届蓝桥杯全国软件和信息技术专业人才大赛广东赛区Python程序设计大学B组一等奖，被评为优秀指导教师。

特发此证，以资鼓励。

证书编号：1633016031

证件号码：440106199105151814

工业和信息化部
人才交流中心

蓝桥杯大赛组委会

2025年5月26日

蓝桥杯大赛

获奖证书

华南农业大学张广煜：

指导周源彬荣获第十六届蓝桥杯全国软件和信息技术专业人才大赛广东赛区Python程序设计大学B组二等奖，被评为优秀指导教师。

特发此证，以资鼓励。

证书编号：1633016173

证件号码：440106199105151814

工业和信息化部
人才交流中心

蓝桥杯大赛组委会

2025年5月26日