综 述

https://www.cjos.ac.cn

AI 驱动相变信息存储半导体关键原子机制探索*

陈念科†, 王柏骞, 赵天宇, 李贤斌†

吉林大学电子科学与工程学院 (集成电路学院) 集成光电子全国重点实验室, 长春 130012

摘要: 相变存储器具有非易失、速度快、可微缩等优点, 适用于存储级内存和存算一体技术, 从而为未来打破传统计算机算力壁垒带来新机会. 相变存储的基本原理是相变存储半导体可逆结构转变诱导光、电信号 (0/1) 转变. 因此, 解析其原子尺度的结构/电子结构转变机制是开发新材料、克服产业化难题的物理基础. 然而, 用于揭示原子尺度机制的传统第一性原理计算研究受到算法、算力的限制, 对相变存储器的计算与模拟进入了瓶颈期. 而近年来兴起的人工智能 (artificial intelligence, AI) 科学研究新范式 (AI for science) 则给该领域带来了新希望. 本文综述了最近基于 AI 方法探索相变存储半导体原子尺度材料物性和相变机制的新进展, 包括相变存储半导体构效关系分析、相变半导体性质优化探索、相变半导体新体系开发、相变半导体机器学习势模型及相变存储器的器件级原子模拟等. 最后, 本文总结了 AI 驱动相变存储技术研究中所面临的主要挑战, 并展望了未来使用 AI 方法辅助解决相变存储技术产业化过程中所面临问题的新思路和新方向.

关键词: 相变信息存储; 半导体材料物理; 人工智能赋能科学

DOI: 10.3724/cjos.2025.025 **CSTR:** 32479.14.cjos.2025.025

引用格式: 陈念科, 王柏骞, 赵天宇, 等. AI 驱动相变信息存储半导体关键原子机制探索. 半导体学报, 2026, 1(3): 031101

1 引言

人工智能 (artificial intelligence, AI) 即将为人类社会的发展带来新一轮的技术革命, 其应用领域遍布各行各业, 已惠及日常生活、医疗健康、科学研究、技术革新、工业制造、商业服务等诸多方面, 呈现百家争鸣、百花齐放的态势. 如今, AI 已经是全球技术竞赛的核心, 未来必将成为新质生产力的主力要素. 同时, AI 算法的高效运行和应用又要立足于高性能的信息硬件 (如计算和存储芯片), AI 技术的爆炸式发展对芯片的算力、能耗、存储、效率等基础指标提出了更高的要求, 因而计算和存储等技术也亟须革新[1−3], 如开发低功耗计算、高密度存储、存算一体、类脑计算等技术. 而这些承载 AI 应用的硬件的发展则是建立在坚实的半导体科学与技术的基础上, 在硬件研发过程中所遇到的复杂科学与技术问题又可以使用 AI 算法辅助攻克. 从而逐渐形成“AI for AI”的技术研发新范式 (图 1(a)).

相变存储 (phase-change memory, PCM) 既是实现高效计算/存储、存算一体和类脑计算技术的重要方案之一, 也是新兴非易失性存储技术中商业化进程最快的技术之一[2, 7−10]. 相变存储的基本原理如图 1(b) 所示, 相变存储材料在外场 (电场或光场) 的作用下可以在晶态和非晶态之间发生快速的可逆相变, 两相之间巨大的电阻对比或介电性质对比即可用于二进制 (0/1) 信息存储[11, 12]. 通过改变外场大小还可以调控有效相变区域的大小, 从而实现多级存储和模拟信号[13−15]. 其可逆相变分为非晶化和晶化两个关键物理过程: 非晶化由强而短的脉冲作用下发生材料熔化-快速冷却-无序态凝固而实现, 而晶化则是在温和的脉冲作用下将非晶态加热到晶化温度之上驱动重结晶来实现. 基于相变存储技术, 美国英特尔公司和美光科技公司联合研发了基于 3D-XPoint 的傲腾存储芯片和产品 (图 1(c)), 欧洲意法半导体公

收稿日期: 2025-09-02; 修回日期: 2025-12-01

*基金项目: 新一代人工智能国家科技重大专项 (2022ZD0117600), 国家自然科学基金面上项目 (12274172, 12274180), 长春市科技发展计划项目 (2024GZZ07)

†通信作者: 陈念科, chennianke@jlu.edu.cn; 李贤斌, lixianbin@jlu.edu.cn

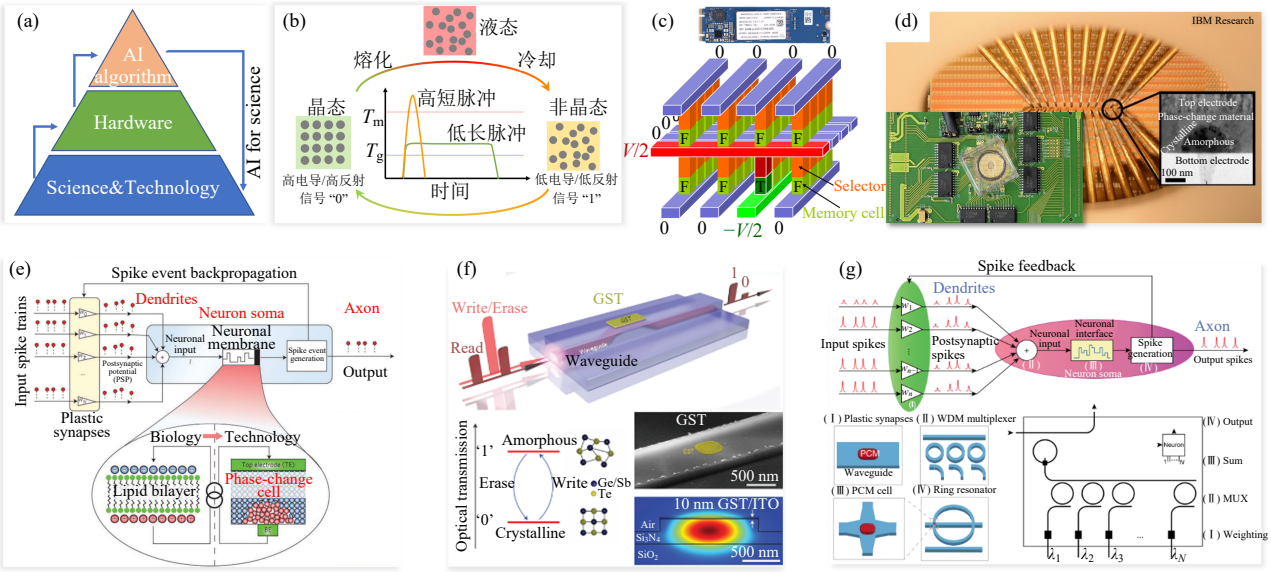


图 1 (a) AI 技术相互支撑示意图; (b) 相变存储基本原理示意图; (c) Intel Optane 相变存储芯片和 3D-XPoint 结构示意图; (d) IBM 研究院基于相变存储的存算一体芯片和类脑计算芯片; (e) 相变存储类脑芯片原理示意图[4]; (f) 全光相变存储芯片[5]; (g) 相变存储光学类脑芯片[6]

Fig. 1. (a) Schematic illustration of the mutual support among AI technologies; (b) Schematic diagram of the working principle of phase-change memory (PCM); (c) Intel Optane memory based on 3D-XPoint architecture and PCM; (d) PCM-based in-memory and neuromorphic computing chips developed by IBM Research; (e) Schematic illustration of PCM-based neuromorphic computing[4]; (f) All-optical PCM[5]; (g) Optical neuromorphic chip based on PCM[6].

司研制了车规级嵌入式芯片和产品, 我国新存科技 (武汉) 公司研发了 NM101 存储级内存模块等. 另外, IBM 研究院也基于相变存储技术研发出存算一体芯片和类脑芯片 (图 1(d) 和图 1(e))[4, 13]. 相变存储不仅可用作电子芯片, 亦可用于光子芯片. 基于相变存储开发的光学芯片已实现了全光非易失性存储 (图 1(f))、类脑计算 (图 1(g))、加速计算等多种功能[5, 6, 16].

在高密度三维相变存储芯片 (如傲腾存储芯片) 中, 除相变单元之外, 还需要选通单元来抑制漏电流 (图 1(c)). 实际上, 相变存储材料和奥氏阈值选通 (ovonic threshold switching, OTS) 材料都是硫族化合物, 其差别在于: 相变材料容易结晶, 工作过程中非晶态和晶态可以快速切换; 而选通材料则难结晶, 一直处于原子无序状态. 相变存储材料和选通材料具有很高的相似性, 它们的研究也一直属于同一领域. 因此, 本文的探讨不局限于相变存储材料, 亦涉及奥氏阈值选通材料的相关研究. 从相变存储的基本原理可知, 其原子机制研究不仅涉及晶体/非晶结构-电学/光学性质之间的复杂构效关系, 而且其相变也由复杂的动力学过程所主导. 更重要的是, 相变存储性能需要同时考虑存储速度、功耗、耐久度、数据保持力、信号对比度等综合因素; 选通材料亦需考虑阈值电压、阈值漂移、开关延迟、电流密度、开关次数等综合性能. 这些性能与材料微观原子机制之间极其复杂的对应关系给传统实验和基于第一性原理计算的理论研究带来了很大挑战. 近年来, AI 技术的迅速发展为基于大数据的材料开发和模拟开创出新的研究范式, 也为相变存储机制的研究带来新机会. 本文将主要介绍近年来在相变存储领域使用 AI 或类 AI 方法开展原子机制研究的几种代表性例子及其取得的成效, 期望为后续使用 AI 方法解决相变存储技术的相关难题提供参考, 亦希望能为其他半导体领域的研究带来一些有益的启发.

2 相变信息存储原子机制

2.1 基于机器学习辅助高通量计算的新型相变半导体开发

多年来, 相变存储半导体的优化和开发通常使用传统的实验试错法 (trial-and-error). 最常见的策略是以传统的相变存储半导体 Ge-Sb-Te (GST) 合金为母体材料, 根据经验从元素周期表中选取不同元素, 进行掺杂或组分替换. 例如, In 掺杂 GST 实现高速、高稳定、低功耗相变存储器[17], C 掺杂 GST 将材料

的十年数据保持温度从 85 °C 提高到 145 °C[18], Cu 掺杂 GeTe 将十年数据保持力提高到 130 °C[19, 20], Ti 掺杂 Sb_2Te_3 将晶化速度提升一个数量级 (~ 6 ns)[21], Sc 掺杂 Sb_2Te_3 进一步将存储速度提高到亚纳秒尺度 (~ 0.7 ns)[22] 等. 通常, 这种实验试错法的成本高、周期长、效率低. 基于材料基因工程的高通量筛选和计算方法能显著降低成本、缩短周期并提高效率, 从而为相变存储半导体的开发提供新思路. 例如, Liu 等[23] 根据相变存储材料的特性制定了一系列可表征材料性质的物理描述符, 用以开展高通量筛选研究. 具体描述符包括: (1) 主要元素组分, 根据传统相变材料的元素组分并考虑非晶半导体的稳定性, 选择含 IV、V、VI 主族元素的化合物作为主要元素组分, 但是排除非晶稳定性过高而不利于结晶的氧化物; (2) 带隙, 相变存储材料通常为窄带隙半导体, 窄带隙可以保证材料晶态电阻较低且容易产生足够的焦耳热, 因此选择带隙值小于 1 eV 的半导体; (3) 原子局域构型, 选择八面体构型作为原子局域构型, 因为 p 电子成键和超价键 (metavalent bond or hyperbond) 被认为是相变存储材料提供足够信号比和快速晶化能力的内在原因; (4) 能量凸包 (energy above hull, ΔE_{hull}), 该能量可用于判断材料分解为其他稳定物质的难易程度, 值越小则材料越稳定, 因此, 选择 ΔE_{hull} < 0.1 eV/atom 来保证材料的热力学稳定或亚稳定性. 根据以上物理判据, 从 Materials Project 数据库十几万种晶体结构中筛选出 158 种可能的相变存储材料 (图 2(a)); 再根据内聚能 (可表征材料熔点)、波恩有效电荷 (可表征材料信号对比)、键角偏离度 (即键角偏离 90°或 180°的程度, 可表征 p 电子成键特征) 筛选出 52 种与传统 GST 非常相似的材料, 包括 $TlBiTe_2$、$TlSbTe_2$、$CdPb_3Se_4$ 等 20 余种此前未被该领域考虑过的新组分体系. 虽然相变存储材料工作时具有晶体和非晶两种相, 但是该方法通过制定多种物理判据, 只从晶态结构出发即实现了对材料信号对比和相变能力的初步筛选. 后续的第一性原理计算验证了包括 $Ge_2Bi_2Te_5$、$CdPb_3Se_4$、$MnBi_2Te_4$ 和 $TlBiTe_2$ 在内的几种所筛选材料的晶态-非晶态信号对比. 进一步地, 第一性原理分子动力学模拟以 Cd-Pb-Se 体系为例验证了所筛选材料具有在 300 ps 内晶体形核的能力. 其中, 筛选结果中的 K-Bi-Se 体系的实验研究证明其可以成为有效的相变存储材料[28].

除了直接从数据库中筛选之外, 还可以使用高通量计算方法直接构建新材料或对传统材料进行掺杂改性. Hu 等[24] 使用基于第一性原理高通量计算和筛选方法, 系统研究了过渡金属元素掺杂相变材料 Sb_2Te_3 的性质, 从晶格失配 (≤ 5%)、晶格扭曲 (≤ 0.2%)、带隙值 (≥ 0.4 eV)、形成能 (≤ 0.5 eV) 等方面筛选出 Sc、Y 和 Hg 为最优的几种掺杂元素 (图 2(b)). Sc 元素可以降低材料热导率, 有利于提高热效率. 同时 Sc-Sb-Te 合金的实验发现其可以有效提高晶化速度[22]. 后续的实验研究证明, Y-Sb-Te 体系具有低功耗和高速度的优点[29], 通过调控组分比例, 还可以实现相变存储器更高的热稳定性与更低的阻值漂移[30]. Xu 等[31] 亦使用第一性原理高通量计算方法, 以 $Ge_1Sb_2Te_4$ 和 Sb_2Te_3 为母体, 构建了多种岩盐矿结构的硫族二元 (V_2VI_3) 和三元 ($IV_1V_2VI_4$) 化合物, 并通过结构弛豫和分子动力学模拟筛选出结构稳定的 9 种二元和 38 种三元硫族化合物, 这些材料中元素和空位的无序分布还可能带来安德森局域化, 为调控硫族化合物电子结构用于相变存储提供理论参考.

与相变存储材料相比, 奥氏阈值选通材料的筛选更为困难. 其原因在于奥氏阈值选通材料均是非晶态, 不仅缺少相应的材料数据库, 而且通过从头算分子动力学大规模构建非晶模型需要大量的计算资源. 事实上, 也可以采用先从经典理论出发制定非晶性质的描述符进行高通量初筛, 再对筛选出的非晶材料进行第一性原理计算建模的方式进行探索. Clima 等[32] 使用高通量计算方法筛选不含 As 和 Se 元素 (环境友好) 的奥氏阈值选通材料. 首先, 以无毒元素 B、C、N、Al、Si、P、S、Zn、Ga、Ge、In、Sn、Sb、Te 按照不同比例 (比例变化间隔为 10%) 组合出 13 104 种可能的材料; 然后, 参考已有典型奥氏阈值选通材料 (如 SiGeAsTe、SiGeAsSe、GeAsSe 等) 筛选出每原子平均价电子数为 5 左右的组分; 再使用 Lankhorst 提出的半经验模型评估材料的玻璃化转变温度[33], 由于奥氏阈值选通材料一般需要较高的非晶稳定性, 因此选取玻璃化转变温度大于 600 K 的材料作为奥氏阈值选通材料的候选体系; 进一步对初步筛选获得的材料进行非晶建模并计算形成能和电子结构, 选择形成能小于零 (热力学稳定) 和具有较小陷阱能级带隙 (导带和价带尾态之间的能量间隔) 的材料; 最后, 提出一种衡量开启状态电流的标准: 注入电荷诱导材料波恩有效电荷改变量 (OTS gauge > 17); 最终, 选出 35 种可能的二元奥氏阈值选通材料. 基于

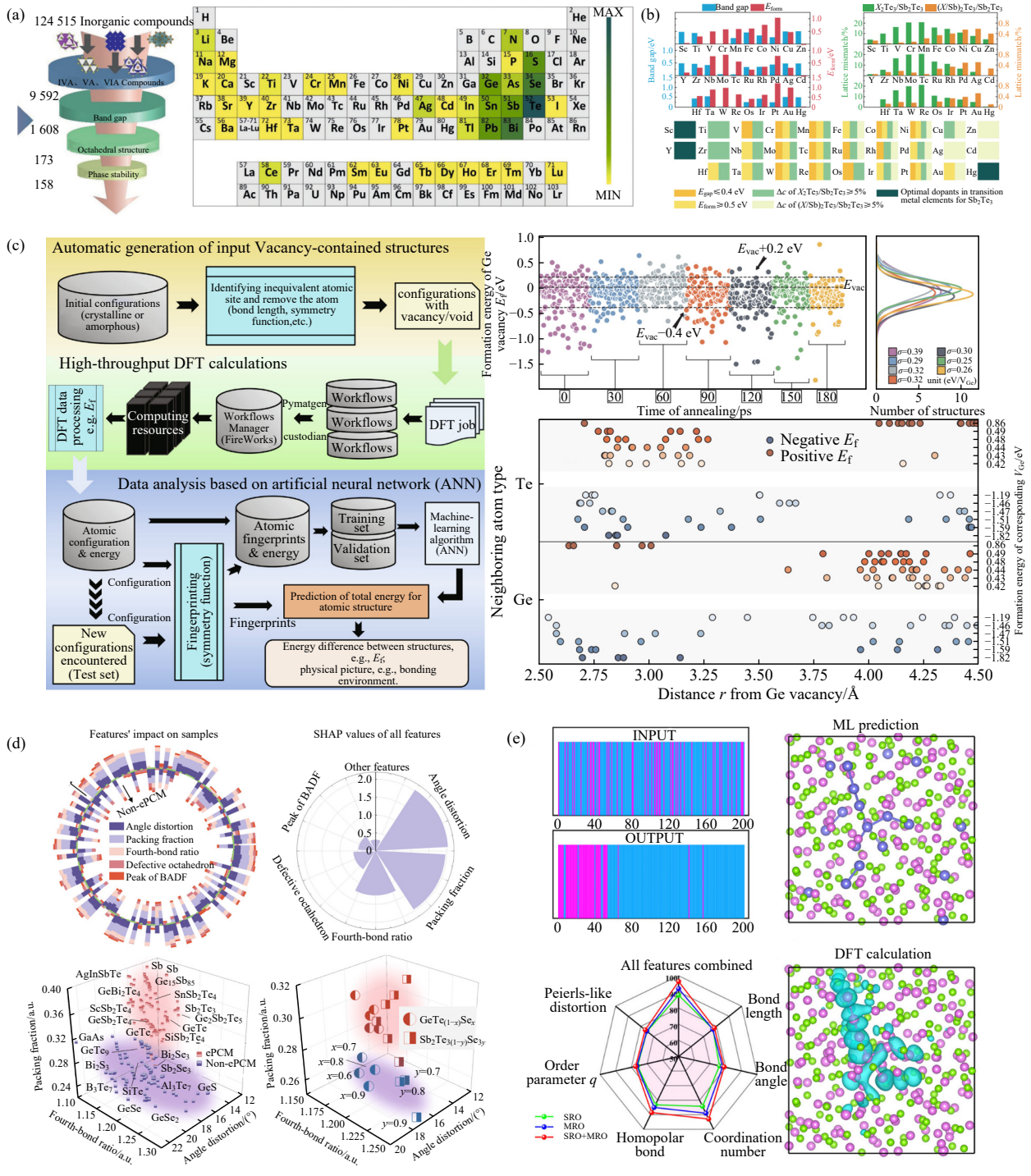


图 2 (a) 相变存储材料的高通量筛选[23]; (b) 过渡金属掺杂 Sb_2Te_3 的高通量筛选[24]; (c) 全自动的机器学习辅助 GeTe 空位缺陷性质的高通量计算[25]; (d) 机器学习方法区分相变存储材料和奥氏阈值选通材料[26]; (e) 机器学习方法预测奥氏阈值选通材料中的缺陷态构效关系[27]

Fig. 2. (a) High-throughput screening of PCM materials[23]; (b) High-throughput screening of transition-metal-doped Sb_2Te_3[24]; (c) Fully automated machine-learning-assisted high-throughput calculations of vacancy-defect properties in GeTe[25]; (d) Machine-learning-based classification of PCM materials and Ovonic threshold switching (OTS) materials[26]; (e) Machine-learning-based prediction of structure–property relationships of defects in OTS materials[27].

类似的方法, Sung 等[34] 使用高通量计算方法筛选适用于单选通管存储器 (selector-only-memory, SOM) 的奥氏阈值选通材料, 这种单选通管存储器避免了相变单元和奥氏阈值选通单元的串联, 不仅简化工

艺, 而且消除了相变单元的不良影响, 使得存储功耗更低、循环性更好. 单选通管存储器的基本原理是利用了 OTS 器件阈值电压 (V_{th}) 的可调性; 选通管的机制导致其开启状态是易失性的, 利用其高、低阻态无法实现非易失的信号对比. 然而, 在某些 OTS 器件中, 施加不同方向的电脉冲会使得器件的 V_{th} 发生非易失性的改变. 因此, 利用不同的 V_{th} 作为信息存储的状态, 亦可实现非易失性存储. SOM 材料具体筛选策略包括: (1) 元素组分, 由 Al、Si、P、Ga、Ge、As、In、Sn、Sb 和硫族 S、Se、Te 组成 3 888 种不同组分配比的三元化合物; (2) 成键特性, 计算构建的化合物的电子杂化性和离子性, 筛选出杂化性较高 (1.7−2.7)、离子性较弱 (小于 0.4) 的材料; (3) 热稳定性, 使用内聚能和原子化焓 (atomization enthalpy) 衡量材料热稳定性, 筛选内聚能 (E_c > 2.9 eV/atom) 和原子化焓 (H_A > 275 kJ/mol) 较大的材料; (4) 电子结构, 计算材料态密度并统计带隙和陷阱能级密度, 再使用 Poole-Frenkel 模型评估阈值电压和阈值电压窗口; (5) 可靠性, 使用分子动力学模拟评估材料结构弛豫导致的带隙变化, 选择结构弛豫时带隙变化较小的材料. 最终, 筛选获得 18 种可能适用于 SOM 的奥氏阈值选通材料.

从以上讨论中可以发现, 相变存储材料和奥氏阈值选通材料的判据非常复杂, 过去主要通过物理上的理解和经验制定判据. 在进行高通量计算和构效关系分析过程中, 如果引入机器学习方法, 可以有效提高效率甚至发现常规经验之外的构效关系. Cheng 等[25] 使用基于人工神经网络 (artificial neural network, ANN) 的机器学习算法辅助相变存储材料 GeTe 中 Ge 空位缺陷的高通量计算 (图 2(c)), Ge 空位是相变存储材料 GeTe 中的常见缺陷, 对材料的载流子浓度等电子性质有显著影响. 晶态中的空位类型可以认为基本一致, 但是对于没有结构周期性的非晶态来说, 空位 (或者由多个空位组成的 void) 的类型有许多种, 且几乎不可能通过第一性原理方法进行遍历计算. 而使用基于 ANN 的机器学习方法可以用有限数量的第一性原理计算结果拟合出非晶中空位结构与材料总能的关系, 从而直接预测非晶态中空位缺陷的形成能, 实现高效率的形成能高通量计算. Xu 等[26] 使用机器学习辅助高通量计算方法, 完成对相变存储材料和奥氏阈值选通材料的智能区分 (图 2(d)). 如前文所述, 相变存储材料和奥氏阈值选通材料非常类似, 均是硫族化合物, 实验上对其进行区分需要大量复杂的材料制备和测试研究. 如果可以通过元素组分和结构直接判断材料类型, 将大幅提高相变材料和选通材料的筛选效率. 首先, 大量计算产生已知相变存储材料和奥氏阈值选通材料的非晶结构模型. 然后, 对这些非晶模型进行对分布函数 (pair correlation function, PCF)、配位数 (coordination number, CN)、键角分布函数 (bond angle distribution, BAD) 等结构分析, 构建一系列包括对分布函数次近邻最近邻距离差 ($r_{次近邻} - r_{最近邻}$)、第四近邻键长比 ($l_4/[(l_1 + l_2 + l_3)/3]$)、平均键角、键角偏离度 (偏离 90°的程度)、八面体结构参数、四面体结构参数、堆积比率 (packing fraction)、电负性差 (包括平均值、最大差和最小差) 等在内的多类型特征. 再使用典型机器学习算法 XGBoost 对相变材料和奥氏阈值选通材料进行分类学习, 在测试集上, 学习准确率达到 97%. 更值得关注的是, 该模型在学习集之外的材料体系 $GeTe_{(1-x)}Se_x$ 和 $Sb_2Te_{3(1-y)}Se_{3y}$ 中也起到很好的预测作用, 分别在这两类材料中预测出在组分变化过程中 x = 0.5 和 y = 0.6 是相变材料和奥氏阈值选通材料转变的临界值. 尽管现在的机器学习方法仍然具有不可解释的缺点, 但是研究者利用机器学习的分析、预测结果可以反过来推理背后的物理机制. 在该研究中[26], 对机器学习模型中的各个特征的权重进行分析发现键角偏离度、堆积比率和第四近邻键长比这三个特征起到最主要的作用, 其中键角偏离度的贡献最大. 这样便可以根据这些特征反过来挖掘相变存储材料和奥氏阈值选通材料分类背后起关键作用的物理图像. 从以上结果可以总结出, 相变存储材料的局域 Peierls 扭曲小, 键角仍然接近 90°, 键长偏差也较小; 而奥氏阈值选通材料的局域 Peierls 扭曲大, 键角偏离 90°更多, 键长偏差也更大. 其背后的电子结构机制在于, Peierls 扭曲可以一定程度上表征 p 电子成键的偏离程度, Peierls 扭曲越小, p 电子成键越强. 也就是说, p 电子成键强的硫系材料适合做相变存储材料, 而 p 电子成键偏离程度高的硫系材料适合做奥氏阈值选通材料. 该结论亦可与领域中的普遍认知相互印证[23]. 这个工作说明, 机器学习可以在非常复杂的构效关系分析中 (如在分析非晶态时) 抓取到核心特征, 并能反过来帮助研究者更好地理解微观原子机制. 另外, 机器学习方法不仅可以提取材料的特征结构, 还有可能建立微观

表 1 相变存储材料和奥氏阈值选通材料主要经验判据总结与对比[23, 26, 27, 34]

Table 1. Summary and comparison of the key empirical descriptors for PCM and OTS materials[23, 26, 27, 34].

经验判据	PCM材料	OTS材料
元素组分	硫族化合物 (Te基为主)	硫族化合物 (Se基为主)
热稳定性	相对低	高
原子局域环境	倾向八面体构型	倾向四面体构型
成键特性	Metavalent/hyper bonding	共价性更强
价电子特征	p电子成键为主	电子杂化性强
键角偏离度(偏离90°)	相对小	相对大
第四近邻键长比	相对小	相对大
带间缺陷态	不重要	重要

结构和电子性质的关联. Xu 等[27] 使用机器学习方法研究奥氏阈值选通材料带间陷阱态 (mid-gap states, MGS)的原子机制 (图 2(e)). 首先, 以典型奥氏阈值选通材料 GeSe 为例, 使用第一性原理分子动力学方法计算产生 100 多个非晶模型, 并根据局域态密度来表示原子对 MGS 的贡献. 然后, 对非晶中原子的局域环境进行分析, 制定包括键长、键角、配位数、异质键、结构序参数、Peierls 扭曲等在内的多个特征, 以上述特征开展机器学习训练获得的模型对 MGS 预测的识别率达到 95%. 之前的多个研究曾认为 Ge-Ge 键的形成和 Ge 的过配位是非晶奥氏阈值选通中产生 MGS 的结构起源, 该研究中机器学习模型识别的 MGS 中亦有 Ge-Ge 键的贡献. 然而, 事实上非晶态中还有很多其他 Ge-Ge 键并没有贡献 MGS; 为了探究根本原因, 使用深度学习模型同时考虑近程序和中程序, 通过对特征量的分数评估, 发现配位数和同质键确实是关键因素, 特别是 Ge-Ge 同质键构成的原子链, 以及原子链周围的 3 个配位 Se 起关键作用. 因此, 研究者借鉴该发现来进行逆向思考, 提出了局域原子环境违反了八电子规则 (8-N rule) 导致产生 MGS 的机制. 该研究表明, 机器学习可以从复杂非晶结构中识别出电子结构的微观原子结构起源, 并抓取关键特征, 从而助力揭示出其中的原子机制.

表 1 总结对比了上述相变存储材料和奥氏阈值选通材料搜索过程中常用到的主要经验判据. 其中, 区分相变材料和奥氏阈值选通材料的最主要判据是材料的热稳定性. 相变存储材料要求能够快速晶化, 其非晶态的热稳定性偏低; 而奥氏阈值选通材料是易失性开关, 必须保证材料始终是非晶态, 其热稳定性越高越好. 另外, 在这些经验判据中, 仍有许多物理原理有待发掘, 而判据的有效性也有待进一步的检验. 例如, 在奥氏阈值选通材料的成键特征中共价键是否起主导作用、金属共价键 (metavalent bond) 或超价键 (hyper bond) 是否起关键作用仍待澄清; 奥氏阈值选通材料的带间缺陷态是否是必要的、带间缺陷态的密度需要达到多少仍需进一步探索; 用作单选通管存储器的 OTS 材料既需要较大的 V_{th} 调控窗口又要保证较小的 V_{th} 漂移幅度, 两者常常矛盾, 是应该选择结构弛豫时带隙变化小的材料还是应该选择结构弛豫导致带隙变化大的材料仍然需要进一步的检验.

2.2 基于机器学习势的相变动力学机制模拟

相变存储器的性能主要由相变性质决定. 例如, 晶化速度决定器件擦写速度, 熔化温度决定器件功耗, 晶化温度决定数据保持力, 相变过程中的分相能力决定器件的耐久度, 非晶结构弛豫决定器件阻值漂移等. 因此, 研究相变存储材料工作过程中的相变机制是领域核心的研究内容. 然而, 传统基于密度泛函 (density functional theory, DFT) 理论的分子动力学模拟方法由于求解薛定谔方程的计算量通常只能模拟数百原子在 1 ns 内的情况, 如果考虑器件级别的模拟, 加上电极或衬底界面作用, 基本只能模拟低维相变存储器. 例如, Wang 等[35] 基于密度泛函理论分子动力学模拟了包含 SiO_2 或 h-BN 衬底/盖层的二维相变存储器件, 揭示相变材料和衬底/盖层间界面化学键钉扎作用是导致相变存储材料微缩至亚 10 nm 后晶化失效的原因, 提出使用单层 Sb_2Te_3 作为核心相变材料, 并以具有弱范德华作用的二维材料 (如单层 h-BN) 作为衬底/盖层来设计二维相变存储器的方案 (图 3(a)). 该模拟全部使用基于量子力学的第一性原理计算方法, 因此, 模拟的原子数只有数百个, 难以拓展到更大尺度的纳米相变存储器件. 为了能模拟更大空间和时间尺度的相变动力学过程, 必须在兼顾模拟精度的情况下提高分子动力学模拟的效率. 其中一个办法是采用经典力场方法, 此方法无需求解薛定谔方程, 计算速度远远大于第一性原理分子动力学. 但传统的力场方法需要构建准确的势函数, 对于这种多元合金来说难度大、门槛高. 而机器学习方法则为构建经典势函数带来新的机会, 不仅能获得量子力学精度的原子间作用势, 而且也

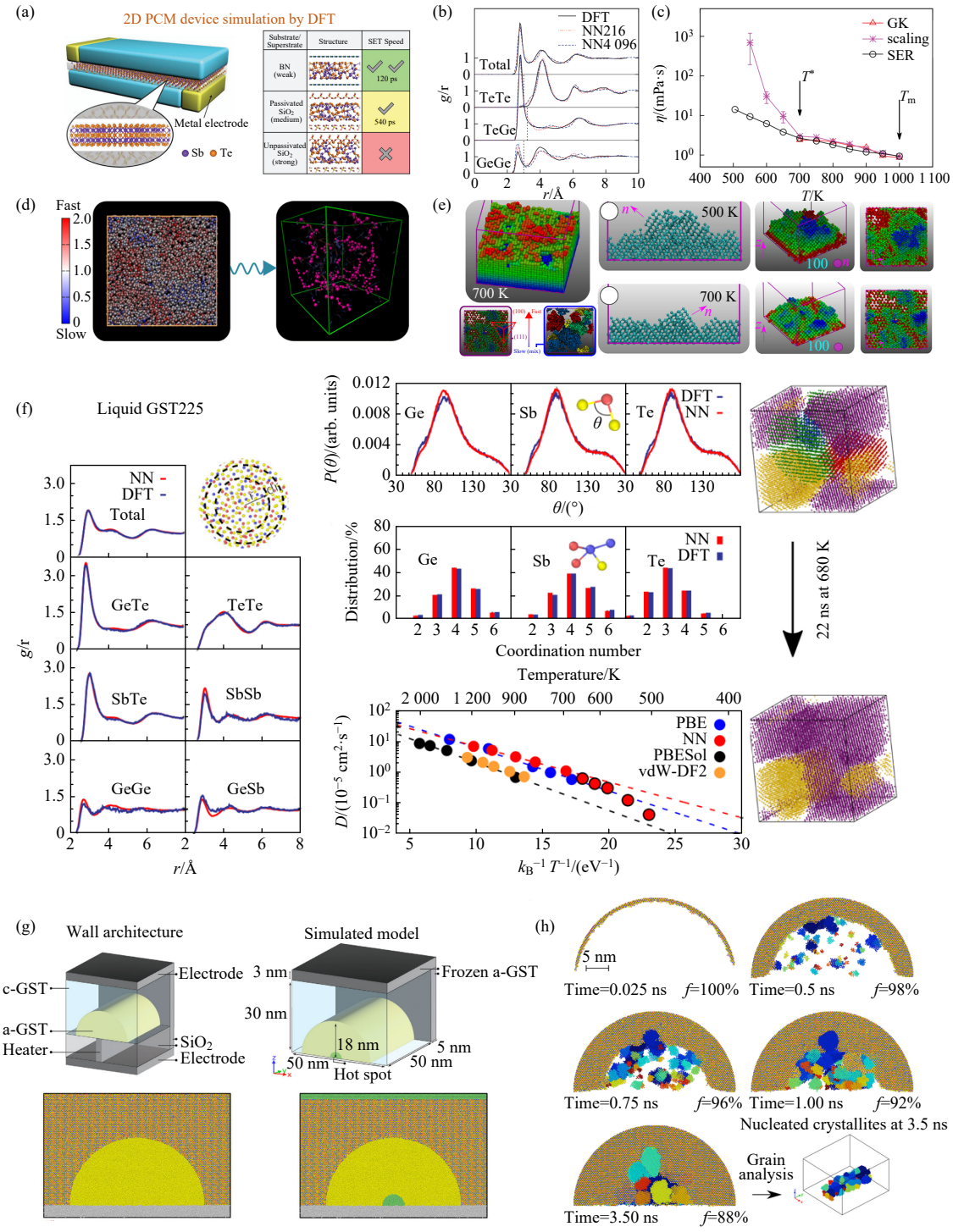


图 3 (a) 基于第一性原理的二维相变存储器模拟[35]; (b) 基于神经网络的 GeTe 机器学习势描述结构对关联函数时达到 DFT 量子力学精度[36]; (c) 机器学习势分子动力学揭示 GeTe 中斯托克斯-爱因斯坦关系的失效[37]; (d) GeTe 熔化过程中的非均匀原子迁移和结构分布[38]; (e) GeTe 不同晶面的晶化动力学过程模拟[39]; (f) 基于 DeepMD 的 GST 机器学习势[40]; (g) 和 (h) 基于 GST 机器学习势的意法半导体公司刀片电极架构的相变存储器模拟[41]

Fig. 3. (a) First-principles simulations of two-dimensional PCM devices[35]; (b) Neural-network-based machine-learning potentials (MLP) for GeTe achieving DFT-level accuracy in describing pair correlation functions[36]; (c) Molecular dynamics (MD) simulations based on MLP revealing the breakdown of the Stokes-Einstein relation in GeTe[37]; (d) Non-uniform atomic migration and structures during the melting process of GeTe[38]; (e) MD simulations of the crystallization dynamics on different crystallographic facets of GeTe[39]; (f) DeepMD-based MLP for GST[40]; (g, h) MD simulations of a PCM device with the blade-electrode architecture developed by STMicroelectronics based on MLP[41].

将大幅度降低构造势函数的门槛.

2012 年, Sosso 等[36] 构建了典型相变存储材料 GeTe 的神经网络 (neural network, NN) 势 (图 3(b)), 该势可以将模拟尺度扩大到数千个原子时仍能同时较好地描述晶态、非晶态和液态, 因此可以用于 GeTe 相变过程的大尺度模拟. 例如, Sosso 等[37] 发现, 相变材料 GeTe 在降温到玻璃化转变温度 (~700 K) 附近时成为过冷液体, 其黏度值较高, 但其原子的扩散能力依然很强, 扩散系数 (D)、温度 (T) 和黏度 (η) 之间的关系打破了斯托克斯-爱因斯坦关系 (Stokes-Einstein relation, $D \propto T/\eta$) (图 3(c)), 正是这种性质使得相变存储材料可以在晶化温度附近快速结晶. 该研究还为曾在相变存储材料 GST 的实验研究中观察到过的类似现象提供了理论支撑[42]. Sosso 等[43] 还使用 NN 势模拟了 GeTe 的晶化过程, 进一步证明 GeTe 在 500−700 K 温度区间是脆性过冷液体, 原子扩散很快, 打破了斯托克斯-爱因斯坦关系, 使得 GeTe 晶体的生长过程非常快, 整体晶化速度不取决于形核率. Sosso 等[38] 还揭示出 GeTe 过冷液态中的非均匀动力学, 即过冷液态 GeTe 中众多原子的运动在空间和时间上并不是完全均匀的 (图 3(d)). 不仅在时间上有快过程和慢过程, 对应到空间上也有运动快的区域和运动慢的区域. 其中富含 Ge-Ge 链的区域原子运动快, 被认为是非均匀运动的原因, 也是斯托克斯-爱因斯坦关系失效的原因. Sosso 等[39] 又使用 GeTe 的 NN 势进行了包含数万原子数的大尺度固液界面处的晶化动力学模拟 (图 3(e)). 研究发现温度较低时 (500 K) 过冷液态很容易形核, 而温度较高时 (700 K) 主要是通过固液界面进行生长, 这与传统“低温形核, 高温生长”的认知也是一致的. 在生长过程中, 立方相 GeTe 的 (001) 面比 (111) 面的生长速度快很多, 其原因可能是 (001) 面上同时有 Ge 原子和 Te 原子作为生长点位, 而 (111) 面的最表面要么是一层 Ge 原子, 要么是一层 Te 原子, 特别是当 Te 原子层在最外面的时候, 结构相对稳定, 生长速度下降. Gabardi 等[44] 使用 NN 势模拟了直径达到 9 nm 的 GeTe 纳米线的相变过程, 发现纳米线的熔点比体材料低 100 K, 晶化速度也降低了约两倍.

Dragoni 等[45] 使用单质 Sb 的神经网络机器学习势研究了 3 nm 厚 Sb 薄膜的相变行为, 发现体材料容易晶化的非晶 Sb 在 3 nm 几何限域条件下仍然稳定, 与实验观察一致[46]. 因此, 微缩的 Sb 提高了非晶态的热稳定性, 使其更适合用于非易失性存储技术. 动力学分析发现主导限域效应下 Sb 非晶稳定性的动力学过程是局域结构的 β 弛豫, 为研究相变材料在限域效应下的动力学提供了重要思路. Perego 等[47] 使用基于 NN 的 GeTe 机器学习势模拟了 GeTe 在磁控溅射条件下的生长过程, 发现溅射获得的非晶态与熔化冷却获得的非晶态相比含有更多的 Ge 四面体构型和 Ge-Ge 键. 但是该差别对材料的晶化性质没有带来很大的影响, 因此推测实验上溅射获得的非晶态与熔化冷却获得的非晶态之所以晶化动力学不同是因为材料中含有的初始晶核不一样.

Abou El Kheir 等[40] 基于深度神经网络构建了典型相变存储材料 GST 的势 (图 3(f)), 使用 DeepMD 框架并在图形处理器 (graphics processing unit, GPU) 加速条件下, 计算效率大幅提高, 可以进行数万原子、数百纳秒的分子动力学模拟, 从而可以在真实器件操作的时间尺度上系统研究多种相变或弛豫过程. 例如, 在晶化模拟过程中, 不仅能再现晶粒粗化现象, 而且揭示出 β 弛豫只在玻璃化转变温度附近存在, 在 500−640 K 温度区间的晶化主要由原子的扩散能力主导, 其晶化速度难以直接使用经验的 Wilson-Frenkel 公式描述, 而需要使用改进后的 Wilson-Frenkel 公式来分析. 这种基于 DeepMD 的机器学习势使得相变存储器中的晶界相互作用、分相行为等长时间、大尺度现象的模拟成为可能. 随后, Abou El Kheir 等[41] 使用此 GST 的 NN 势开展数十万原子模型的器件级模拟 (图 3(g) 和图 3(h)), 以意法半导体公司开发的相变存储芯片架构为原型, 模拟了纳米刀片电极附近蘑菇型相变区的 SET 操作 (晶化过程). 在 600 K 热浴条件下, 相变区形成大量晶核, 晶核的生长和晶体-非晶界面处的生长呈竞争关系. 模拟中发现从晶体-非晶界面处的生长速度略快于从晶核处的生长. 这种竞争关系受温度影响较大, 温度较高时形核率会显著下降. 在以上模拟中, 相当于晶化释放的相变潜热是被所有原子的热浴移除掉; 如果考虑相变潜热只是由上电极耗散, 则体系的整体温度会上升, 使得在晶体-非晶界面处的晶化速度比内部晶核的晶化快很多, 也与“低温形核高温生长”的经验相符. 另外, 这种条件下晶化过程中的形核数变少, 晶核尺度增大. 如果热源集中在内部形成热点且采用顶电极散热, 则内部温度高, 而晶体-非晶界面温

度低, 其结果与之前的情况又不一样. 晶体-非晶界面处的生长速度和内部晶化速度相差不大. 这些研究说明相变存储器在 SET 过程中的动力学过程及相变效果与热源分布和散热条件息息相关, 为了解决相变存储器工作过程中晶化随机性带来的器件性能不一致的问题, 理论上应该提高相变温度, 减少形核数, 增大晶核尺寸.

另外一种在该领域使用较多的是高斯近似势 (Gaussian approximation potential, GAP). 2018 年, Mocanu 等[48] 使用 GAP 框架构建了典型相变存储材料 GST 的机器学习势. 使用该势, 不仅可以模拟液态、非晶结构和晶化过程, 而且模型的增大使得非晶的"中程有序"结构研究变成可能. 例如, 使用 7200 个原子的模型再现了非晶结构因子中的第一尖锐衍射峰 (first sharp diffraction peak, FSDP), 这个实验中探测到的 1.1 $Å^{-1}$ 处幅值较小的 FSDP, 使用较小的模型 (如几百个原子) 是无法重现的 (图 4(a)). 对大尺度非晶结构分析则揭示出非晶 GST 中的 FSDP 来源是以阳离子为中心的关联结构, 其关联长度可达 20 Å. 另外, 该势不仅可以研究 GST, 还适用于 GeTe 和 Sb_2Te_3 的研究.

随后, Konstantinou 等[51] 使用 GST 的 GAP 势完成了大量非晶模型的构建, 针对非晶模拟随机性问题, 开展了统计研究. 研究发现熔化冷却的非晶模型中只有 37% 是有带间缺陷态的, 5 配位的 Ge 原子是缺陷态的主要来源, 其所在的微结构片段与晶体结构较像. 非晶中局部原子应力分布不均匀, 缺陷态相关的原子局部应力反而很小. 以上研究对揭示相变材料的电阻漂移机理具有重要意义, 因为非晶阻值逐渐增大通常来源于缺陷态的消除. 然而, 由于缺陷态的浓度实际很低, 传统第一性原理分子动力学模拟生成的非晶结构通常难以自发采样到带隙中的局域缺陷态, 因此, 采用 GAP 势提高计算效率, 有效解决了这一问题. Mocanu 等[52] 用 GAP 势研究了尺寸效应和降温速率对非晶 GST 的影响. 研究尺寸效应时, 对比模型的原子数从 300 到 24 300 (约 10 nm, 使用 15 K/ps 速率降温), 发现模型尺寸对晶化的影响作用较小. 研究降温速率时, 对比速率从 100 到 5 K/ps (这里使用 900 原子的模型), 发现降温速率会影响获得非晶结构的能量, 该关系可以用幂指数函数或对数 (log) 的幂指数函数拟合. 另外, 研究发现 1 K/ps 速率下会发生部分结晶, 因此 1 K/ps 可以作为非晶化的最低降温速率. Wang 等[49] 使用 GAP 势研究了 GeTe/Sb_2Te_3 超晶格和传统立方相 GST 的熔化行为. 当使用较小的模型时 (如数百原子), 超晶格模型包含的范德华空位层数量较少, 又受限于模型的有限尺寸效应, 其熔化行为与立方相 GST 非常相似, 达到熔点后整个模型迅速熔化. 但是当模型扩大到数千原子后, 明显观察到超晶格 GST 的熔化起始于反位缺陷导致阳离子进入范德华空位层, 破坏了稳定的层间作用, 形成初始液核, 并逐渐生长扩大. 而立方相 GST 由于空位分布均匀, 达到熔点后每一个空位都可以是熔化的形核位点, 整个模型迅速均匀熔化. 即超晶格 GST 的熔化表现出一种局部熔化逐渐生长的特性, 而传统立方相 GST 的熔化是由高密度空位主导的均匀形核熔化 (图 4(b) 和图 4(c)). Zhou 等[50] 优化了机器学习的数据集, 增加了更多复杂的结构场景, 涵盖了更多化学计量比的组分 (包括 $Ge_8Sb_2Te_{11}$、$Ge_4Sb_2Te_7$、$Ge_1Sb_2Te_4$、$Ge_2Sb_2Te_5$ 等), 最终构建出新版的相变存储材料 GAP 势 (GST-GAP-22), 并首次完成了 GST 的器件级原子模拟 (图 4(d) 和图 4(e)). 包括多脉冲作用下的 SET 操作 (晶化行为)、非等温条件下的典型"蘑菇型"RESET 操作 (非晶化行为) 以及以 3D-XPoint 存储架构为参考原型模拟的 40 nm × 20 nm × 20 nm (原子数超过五十万) 限制条件下的纳米相变存储器的 RESET 操作. 该势函数已经开源, 研究者可以在此基础上添加训练集, 构建更多复杂场景的新势.

与基于 DeepMD 的 NN 势相比, GAP 势的精度虽然较高, 但是其计算速度相对较慢, 在进行器件级原子模拟时消耗的计算和时间资源更多. 为了进一步提高计算效率, du Toit 和 Zhou 等[53, 54] 构建了基于原子簇展开 (atomic-cluster-expansion, ACE) 框架的 GST 合金机器学习势. ACE 势亦可使用 GPU 加速, 且与 GAP 势相比, 计算效率大幅提高, 实现了 3D-XPoint 架构相变存储器的完整擦写过程的分子动力学模拟. 并且, 以蘑菇型器件结构为例, 模拟了相变存储材料在存算一体情景下工作的过程. 与 GAP 势和 NN 势相比, ACE 势进一步提高了计算效率, 为实现真实尺寸的器件模拟提供了新的机会.

除了基于 NN 和 GAP 框架的机器学习势之外, Wang 等[55] 发展了一套图卷积 (graph convolutional) 机器学习势的生成框架并形成自动化训练软件 (PotentialMind), 以相变存储材料 Sb-Te 合金为例构

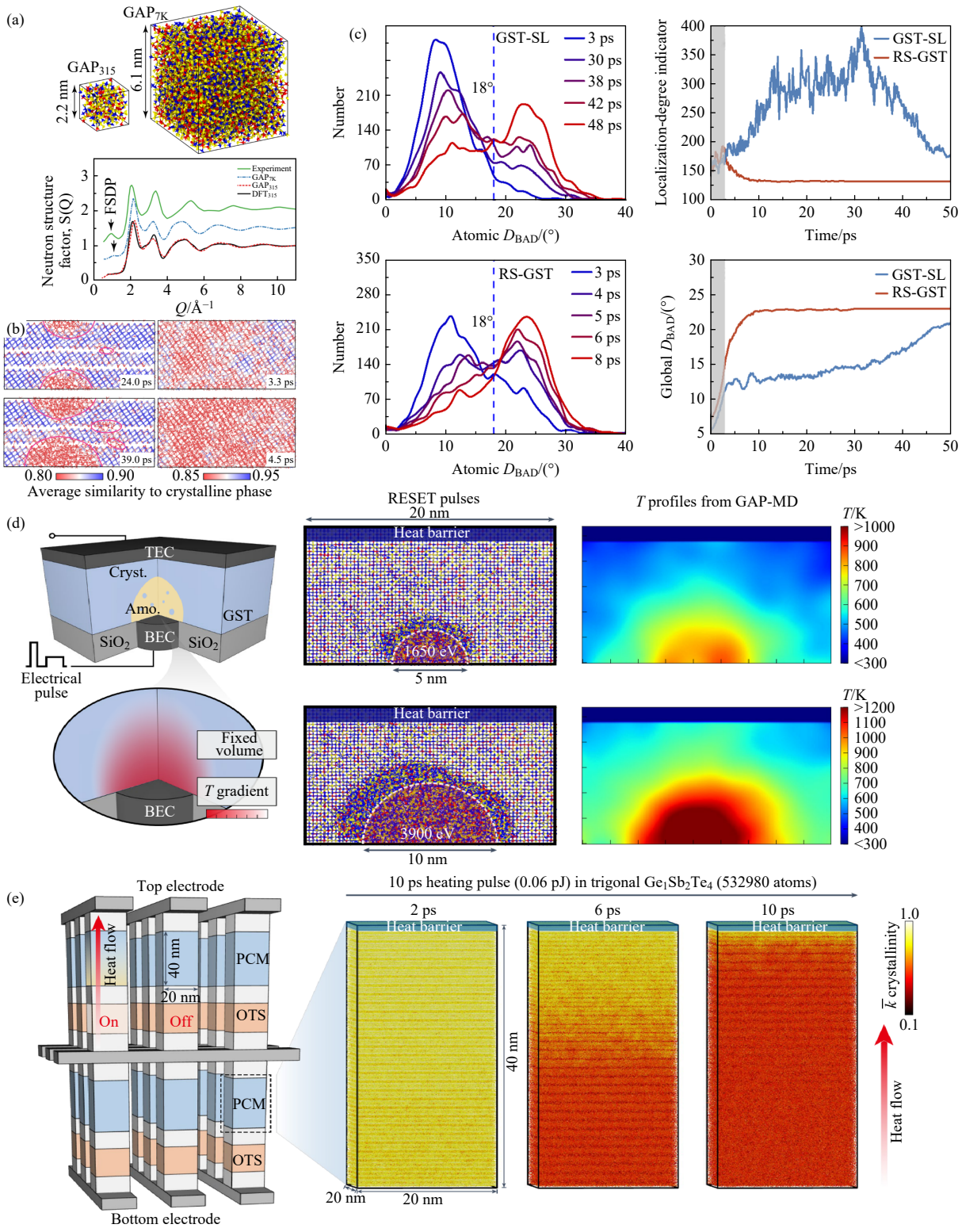


图 4 (a) GST-GAP 机器学习势分子动力学揭示出相变材料非晶态中第一尖锐衍射峰的微观起源[48]; (b-c) GST-GAP 机器学习势分子动力学揭示超晶格相变存储材料中独特的局部熔化-逐渐生长的熔化动力学过程[49]; (d) 基于 GST-GAP-22 机器学习势的传统蘑菇型相变存储器分子动力学模拟[50]; (e) 基于 GST-GAP-22 机器学习势的 3D-XPoint 架构相变存储器分子动力学模拟[50]

Fig. 4. (a) MD simulations based on GST-GAP MLP revealing the microscopic origin of the first sharp diffraction peak in the amorphous phase of phase-change materials[48]; (b, c) MD simulations based on GST-GAP MLP uncovering a unique local-melting dynamics in superlattice PCM materials[49]; (d) MD simulations of conventional mushroom-type PCM devices based on the GST-GAP-22 MLP[50]; (e) MD simulations of 3D-XPoint architecture PCM devices based on the GST-GAP-22 MLP[50].

建了可以描述多种 Sb-Te 组分比例的机器学习势. Wang 等[56] 使用神经演化势 (neuro evolution potential, NEP) 框架构建了 GeTe/Sb_2Te_3 超晶格的机器学习势, 并研究了超晶格的热输运性质, 较好地模拟出体系的声子和热导率. Achar 等[57] 基于矩张量势 (moment tensor potential, MTP) 框架构建了奥氏阈值选通材料 GeSe 和 Ti 界面的机器学习势, 研究了 Ti 电极与 GeSe 接触界面的扩散行为. Wang 等[58] 基于 MTP 框架构建了 $Sc_{0.2}Sb_2Te_3$ 的机器学习势, 研究了该材料相变过程中的声子及热导行为. Xu 等[59] 使用 DeepMD 构建了多种金属 (Al、Sc、Cd)-Sb 化物的 NN 势, 通过分析数纳秒尺度的晶化机制, 预测 ScSb、CdSb、YSb 和 ZrSb 可能是较好的相变存储材料. Yu 等[60] 使用 DeepMD 构建了 $Ge_3Sb_6Te_5$ 合金的 NN 势, 研究了该材料的晶化动力学过程. Acharya 等[61] 构建了 GeTe/$TiTe_2$ 超晶格的 NN 势, 研究了 GeTe 在纳米限域条件下的晶化行为. Choi 等[62] 使用 DeepMD 和配合随机原子系统生成方法构建训练集, 生成了 GeTe 的 NN 势并研究了晶化过程中的声子硬化现象. Choi 等[63] 使用神经等变原子间势配合随机原子系统生成方法构建了 GST 的机器学习势并研究了 GST 中的局域声子现象. 总之, 近年来涌现了大量的基于机器学习势的相变存储半导体原子机制的研究. 关于大数据、机器学习驱动相变材料研究和机器学习势方法的更详细探讨, 可以参考 Wang、Zhang、Sosso 等[64–66] 的综述论文.

3 结论与展望

综上, 目前机器学习驱动相变存储半导体原子机制的研究主要以两种方式开展. 第一种是利用机器学习方法揭示相变存储半导体中复杂的构效关系. 构效关系可以用来制定恰当的描述符以加速高通量筛选, 也可以直接用于材料结构和性质的预测, 还可以用于反推思考, 即通过获得的构效关系逆向分析其中的原子机制. 第二种则是基于机器学习势的大尺度分子动力学模拟. 与传统第一性原理分子动力学相比, 机器学习势分子动力学在结构模拟时, 在基本保证量子力学精度的情况下, 将模拟尺度提高几个数量级, 极大地扩大了模拟的空间尺度和时间尺度, 使得过去难以模拟研究的界面效应、几何限域效应、非均匀相变、长时间动力学过程以及真实纳米器件级的原子仿真等成为可能, 为多角度系统研究相变存储技术中的动力学机制带来新的机会. 值得一提的是, 我国科研工作者在以上研究中贡献了多个关键创新, 已经处于国际前沿水平. 可以预见, 未来基于机器学习的研究方法和新范式不仅能推动相变存储技术的发展, 也会在其他方向大放光彩, 为我国半导体领域的科学研究和技术开发提供重要助力.

从当前的发展来看, 基于机器学习势的研究比基于构效关系分析的研究发展更快、效果更好、方法更成熟. 主要原因在于相变存储 (奥氏阈值选通) 材料结构性质复杂, 又涉及变组分合金和非晶态, 基于量子力学方法的大规模样本生成不仅成本昂贵而且耗时久, 导致构效关系学习样本不够, 学习效果不佳. 另一个原因在于适用于构效关系研究的方法还在持续发展中, 尚未形成一套成熟的研究范式; 而机器学习势的训练已经形成一些成熟有效的方法和软件, 如 DeepMD[67].

然而, 目前基于机器学习势的相变存储器模拟虽然已经达到了器件级, 但是器件的激励方式仍然是热效应, 而真实器件使用的是电脉冲或光脉冲. 电脉冲的作用不仅在于对材料中的原子施加库仑力, 更重要的是器件中的电压和电流分布会导致非均匀的热效应, 直接影响相变区和相变过程. 但是目前基于机器学习势的模拟很难直接包含电流/电压效应, 一个可能的解决方案是与有限元方法结合. 例如, 可以用有限元模拟相变存储器中的电场/电流分布及其诱导的焦耳热分布 (图 5(a))[68], 再将有限元模拟结果对应到机器学习势分子动力学模拟的热分布中, 甚至未来使用 AI 方法替代有限元评估焦耳热效应, 实现场耦合的器件模拟. 与电场作用不同, 光脉冲会直接改变相变材料的原子间相互作用势能面, 即基态的势函数无法描述激发态的原子间相互作用 (图 5(b)). 因此, 研究光脉冲作用下的相变机制, 需要生成激发态的数据集, 训练适用于光激发的新势函数. 目前, 前沿的激发态分子动力学模拟方法是基于含时密度泛函理论的分子动力学模拟, 其模拟尺度仍在 1 ps 以内[70, 69]. 因此, 激发态机器学习势[71] 将对光激发诱导相变存储机制的研究带来潜在的革新方案. 另外, 当前在对相变存储半导体性质和相变存储器模拟的过程中, 一个重要的缺失是材料的电子结构, 传统的第一性原理计算方法无法计算含数十万原子数

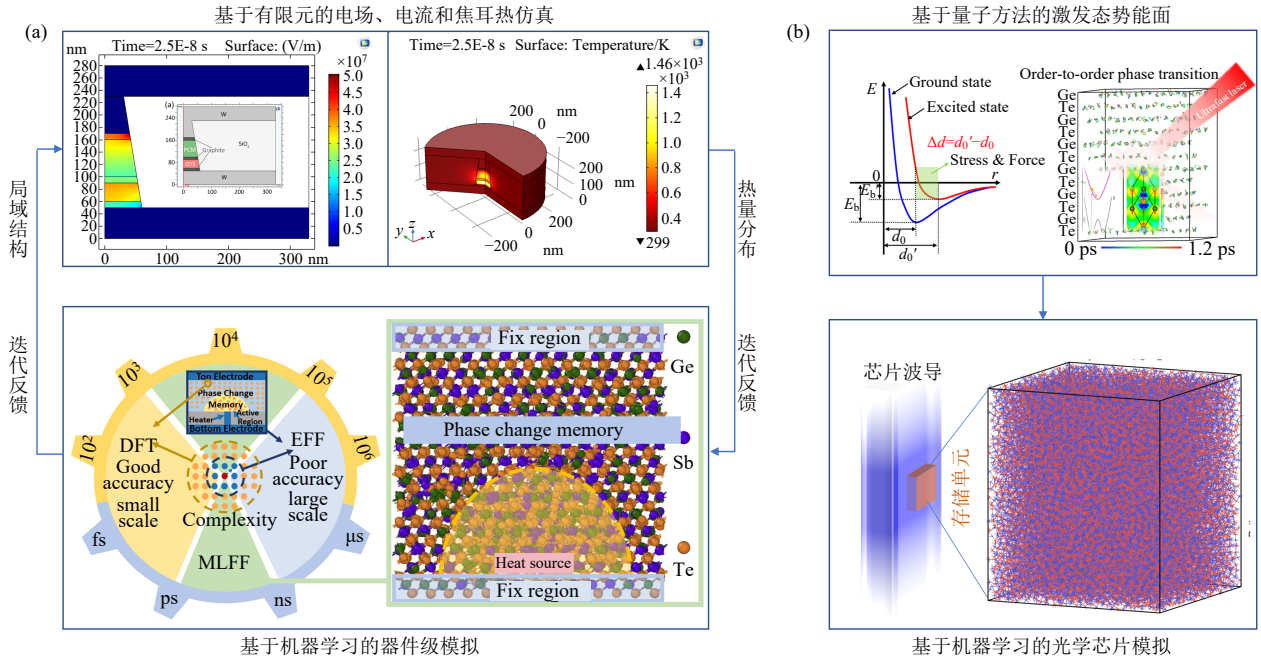


图 5 (a) 机器学习势-有限元模拟相结合的相变存储器仿真研究示意图[68]; (b) 激发态机器学习势有望为光学相变存储芯片的模拟提供新的可能[69]

Fig. 5. (a) Schematic illustration of PCM simulations combining machine-learning interatomic potentials with finite-element analysis[68]; (b) Excited-state MLP offering new opportunities for the simulation of optical PCM chips[69].

的模型的电子结构. 未来可以尝试机器学习辅助紧束缚模型方法, 如 DeepH[72] 等大尺度电子结构计算方法. 另外, AI 技术在训练过程中需要用到大量的数据集, 而不同的研究团队常常独立生成重复的数据, 造成计算资源的浪费. 因此, 相关研究团队甚至是相应的半导体公司如果能够对 AI 数据集进行标准统一和资源整合, 并逐渐形成有组织、有体系的合作研究范式和平台, 必将进一步提升研究效率.

除了提升和拓展 AI 技术在科学研究中的能力和范围之外, AI 在应用中的核心优势和重点目标亦值得继续思考. 除了基本的相变半导体物性研究之外, 应该发扬 AI 处理大尺度复杂问题的能力, 聚焦相变存储芯片工业化过程中遇到的关键工艺问题, 从基础科学上为解决真正的工业难题提供深入的理解和理论支持.

致谢

感谢华中科技大学徐明教授、缪向水教授, 西安交通大学张伟教授, 北京航空航天大学孙志梅教授、肖利民教授、祝令刚副研究员、王良副教授, 中国科学院上海微系统与信息技术研究所宋文雄副研究员、宋志棠研究员, 上海交通大学朱敏教授等在相变存储材料 AI 新范式相关研究中给予的宝贵建议和讨论.

参考文献

[1] Editorial. Big data needs a hardware revolution. Nature, 2018, 554: 145

[2] Li X B, Chen N K, Wang X P, et al. Phase-change superlattice materials toward low power consumption and high density data storage: Microscopic picture, working principles, and optimization. Adv Funct Mater, 2018, 28(44): 1803380

[3] Li W Z, Song Y X, Wang S Y. Intelligent computing, computational power, computational power networks and technology ecosystems. 2024 IEEE 14th International Conference on Electronics Information and Emergency Communication (ICEIEC), Beijing, China, 2024: 1

[4] Sebastian A, Tuma T, Papandreou N, et al. Temporal correlation detection using computational phase-change memory. Nat Commun, 2017, 8: 1115

[5] Ríos C, Stegmaier M, Hosseini P, et al. Integrated all-photonic non-volatile multi-level memory. Nature Photon, 2015, 9(11): 725

[6] Feldmann J, Youngblood N, Wright C D, et al. All-optical spiking neurosynaptic networks with self-learning capabilities. Nature, 2019, 569(7755): 208

[7] Burr G W, Kurdi B N, Scott J C, et al. Overview of

candidate device technologies for storage-class memory. IBM J Res & Dev, 2008, 52(4.5): 449

[8] Burr G W, Shelby R M, Sebastian A, et al. Neuromorphic computing using non-volatile memory. Adv Phys X, 2017, 2(1): 89

[9] Yang J J, Strukov D B, Stewart D R. Memristive devices for computing. Nature Nanotech, 2013, 8(1): 13

[10] Zhang W, Mazzarello R, Wuttig M, et al. Designing crystallization in phase-change materials for universal memory and neuro-inspired computing. Nat Rev Mater, 2019, 4(3): 150

[11] Wuttig M. Towards a universal memory? Nature Mater, 2005, 4(4): 265

[12] Wuttig M, Yamada N. Phase-change materials for rewriteable data storage. Nature Mater, 2007, 6(11): 824

[13] Tuma T, Pantazi A, Le Gallo M, et al. Stochastic phase-change neurons. Nature Nanotech, 2016, 11(8): 693

[14] Wright C D, Hosseini P, Diosdado J A V. Beyond von-Neumann computing with nanoscale phase-change memory devices. Adv Funct Mater, 2013, 23(18): 2248

[15] Wright C D, Liu Y W, Kohary K I, et al. Arithmetic and biologically-inspired computing using phase-change materials. Adv Mater, 2011, 23(30): 3408

[16] Feldmann J, Youngblood N, Karpov M, et al. Parallel convolutional processing using an integrated photonic tensor core. Nature, 2021, 589(7840): 52

[17] Wang R B, Song Z T, Song W X, et al. Phase-change memory based on matched Ge-Te, Sb-Te, and In-Te octahedrons: Improved electrical performances and robust thermal stability. InfoMat, 2021, 3(9): 1008

[18] Zhou W Y, Wu L C, Zhou X L, et al. High thermal stability and low density variation of carbon-doped $Ge_2Sb_2Te_5$ for phase-change memory application. Appl Phys Lett, 2014, 105(24): 243113

[19] Chen N K, Li X B, Wang X P, et al. Origin of high thermal stability of amorphous Ge_1Cu_2Te3 alloy: A significant Cu-bonding reconfiguration modulated by Te lone-pair electrons for crystallization. Acta Mater, 2015, 90: 88

[20] Sutou Y, Kamada T, Sumiya M, et al. Crystallization process and thermal stability of $Ge_1Cu_2Te_3$ amorphous thin films for use as phase change materials. Acta Mater, 2012, 60(3): 872

[21] Zhu M, Xia M J, Rao F, et al. One order of magnitude faster phase change at reduced power in Ti-Sb-Te. Nat Commun, 2014, 5: 4086

[22] Rao F, Ding K Y, Zhou Y X, et al. Reducing the stochasticity of crystal nucleation to enable subnanosecond memory writing. Science, 2017, 358(6369): 1423

[23] Liu Y T, Li X B, Zheng H, et al. High-throughput screening for phase-change memory materials. Adv Funct Mater, 2021, 31(21): 2009803

[24] Hu S W, Liu B, Li Z, et al. Identifying optimal dopants for Sb_2Te_3 phase-change material by high-throughput *ab initio* calculations with experiments. Comput Mater Sci, 2019, 165: 51

[25] Cheng Y X, Zhu L G, Wang G J, et al. Vacancy formation energy and its connection with bonding environment in solid: A high-throughput calculation and machine learning study. Comput Mater Sci, 2020, 183: 109803

[26] Xu Q D, Xu M, Tang S Q, et al. Machine learning for discrimination of phase-change chalcogenide glasses. InfoMat, 2025, 7(4): e70006

[27] Xu M, Xu M, Miao X S. Deep machine learning unravels the structural origin of mid-gap states in chalcogenide glass for high-density memory integration. InfoMat, 2022, 4(6): e12315

[28] Islam S M, Sangwan V K, Buchholz D B, et al. Amorphous to crystal phase change memory effect with two-fold bandgap difference in semiconducting $K_2Bi_8Se_{13}$. J Am Chem Soc, 2021, 143(16): 6221

[29] Liu B, Liu W L, Li Z, et al. Y-doped Sb_2Te_3 phase-change materials: Toward a universal memory. ACS Appl Mater Interfaces, 2020, 12(18): 20672

[30] Liu B, Li K Q, Zhou J, et al. Ultrahigh overall-performance phase-change memory by yttrium dragging. J Mater Chem C, 2023, 11(4): 1360

[31] Xu Y Z, Wang X D, Zhang W, et al. Materials screening for disorder-controlled chalcogenide crystals for phase-change memory applications. Adv Mater, 2021, 33(9): 2006221

[32] Clima S, Matsubayashi D, Ravsher T, et al. In silico screening for As/Se-free ovonic threshold switching materials. npj Comput Mater, 2023, 9: 96

[33] Lankhorst M H R. Modelling glass transition temperatures of chalcogenide glasses. Applied to phase-change optical recording materials. J Non Cryst Solids, 2002, 297(2/3): 210

[34] Sung H J, Choi M, Kang Y, et al. Ab-initio screening of amorphous chalcogenides for selector-only memory (SOM) through electrical properties and device reliability. 2024 IEEE International Electron Devices Meeting (IEDM), San Francisco, CA, USA, 2024: 1

[35] Wang X P, Li X B, Chen N K, et al. Phase-change-memory process at the limit: A proposal for utilizing monolayer Sb_2Te_3. Adv Sci, 2021, 8(13): 2004185

[36] Sosso G C, Miceli G, Caravati S, et al. Neural network interatomic potential for the phase change material GeTe. Phys Rev B, 2012, 85(17): 174103

[37] Sosso G C, Behler J, Bernasconi M. Breakdown of Stokes–Einstein relation in the supercooled liquid state of phase change materials. Phys Status Solidi B, 2012, 249(10): 1880

[38] Sosso G C, Colombo J, Behler J, et al. Dynamical heterogeneity in the supercooled liquid state of the phase change material GeTe. J Phys Chem B, 2014, 118(47): 13621

[39] Sosso G C, Salvalaglio M, Behler J, et al. Heterogeneous crystallization of the phase change material GeTe *via* atomistic simulations. J Phys Chem C, 2015, 119(11): 6428

[40] Abou El Kheir O, Bonati L, Parrinello M, et al. Unraveling the crystallization kinetics of the $Ge_2Sb_2Te_5$ phase change compound with a machine-learned interatomic potential. npj

Comput Mater, 2024, 10: 33

[41] Abou El Kheir O, Bernasconi M. Million-atom simulation of the set process in phase change memories at the real device scale. Adv Elect Materials, 2025, 11(13): e2500110

[42] Orava J, Greer A L, Gholipour B, et al. Characterization of supercooled liquid $Ge_2Sb_2Te_5$ and its crystallization by ultrafast-heating calorimetry. Nature Mater, 2012, 11(4): 279

[43] Sosso G C, Miceli G, Caravati S, et al. Fast crystallization of the phase change compound GeTe by large-scale molecular dynamics simulations. J Phys Chem Lett, 2013, 4(24): 4241

[44] Gabardi S, Baldi E, Bosoni E, et al. Atomistic simulations of the crystallization and aging of GeTe nanowires. J Phys Chem C, 2017, 121(42): 23827

[45] Dragoni D, Behler J, Bernasconi M. Mechanism of amorphous phase stabilization in ultrathin films of monoatomic phase change material. Nanoscale, 2021, 13(38): 16146

[46] Salinga M, Kersting B, Ronneberger I, et al. Monatomic phase change memory. Nature Mater, 2018, 17(8): 681

[47] Perego S, Dragoni D, Gabardi S, et al. Structure and crystallization kinetics of As-deposited films of the GeTe phase change compound from atomistic simulations. Physica Rapid Research Ltrs, 2023, 17(8): 2200433

[48] Mocanu F C, Konstantinou K, Lee T H, et al. Modeling the phase-change memory material, $Ge_2Sb_2Te_5$, with a machine-learned interatomic potential. J Phys Chem B, 2018, 122(38): 8998

[49] Wang B Q, Zhao T Y, Ding H R, et al. Partial melting nature of phase-change memory Ge-Sb-Te superlattice uncovered by large-scale machine learning interatomic potential molecular dynamics. Acta Mater, 2024, 276: 120123

[50] Zhou Y X, Zhang W, Ma E, et al. Device-scale atomistic modelling of phase-change memory materials. Nat Electron, 2023, 6(10): 746

[51] Konstantinou K, Mocanu F C, Lee T H, et al. Revealing the intrinsic nature of the mid-gap defects in amorphous $Ge_2Sb_2Te_5$. Nat Commun, 2019, 10: 3065

[52] Mocanu F C, Konstantinou K, Elliott S R. Quench-rate and size-dependent behaviour in glassy $Ge_2Sb_2Te_5$ models simulated with a machine-learned Gaussian approximation potential. J Phys D: Appl Phys, 2020, 53(24): 244002

[53] Thomas du Toit D F, Zhou Y X, Deringer V L. Hyperparameter optimization for atomic cluster expansion potentials. J Chem Theory Comput, 2024, 20(22): 10103

[54] Zhou Y X, Thomas du Toit D F, Elliott S R, et al. Full-cycle device-scale simulations of memory materials with a tailored atomic-cluster-expansion potential. Nat Commun, 2025, 16: 8688

[55] Wang G J, Sun Y Q, Zhou J, et al. PotentialMind: Graph convolutional machine learning potential for Sb–Te binary compounds of multiple stoichiometries. J Phys Chem C, 2023, 127(51): 24724

[56] Wang B, Li K Q, Zhang W M, et al. Thermal transport of $GeTe/Sb_2Te_3$ superlattice by large-scale molecular dynamics with machine-learned potential. J Phys Chem C, 2025, 129(13): 6386

[57] Achar S K, Schneider J, Stewart D A. Using machine learning potentials to explore interdiffusion at metal–chalcogenide interfaces. ACS Appl Mater Interfaces, 2022, 14(51): 56963

[58] Wang C, Chen Y. Thermal switching across the ultrafast amorphous to crystalline transition in $Sc_{0.2}Sb_2Te_3$. Phys Rev B, 2024, 110(21): 214202

[59] Xu Q D, Xu M, Gu R C, et al. Enhancing elemental phase-change chalcogenide glass through tailored alloying. Acta Mater, 2025, 289: 120896

[60] Yu W, Zhang Z F, Wan X H, et al. High-accuracy machine-learned interatomic potentials for the phase change material $Ge_3Sb_6Te_5$. Chem Mater, 2023, 35(17): 6651

[61] Acharya D, Abou El Kheir O, Campi D, et al. Crystallization kinetics of nanoconfined GeTe slabs in $GeTe/TiTe_2$-like superlattices for phase change memories. Sci Rep, 2024, 14: 3224

[62] Choi Y J, Jhi S H. Efficient training of machine learning potentials by a randomized atomic-system generator. J Phys Chem B, 2020, 124(39): 8704

[63] Choi Y J, Ghim M, Jhi S H. Study of vacancy ordering and the boson peak in metastable cubic Ge-Sb-Te using machine learning potentials. Phys Rev Materials, 2024, 8: 013606

[64] Wang G J, Wang C R, Zhang X G, et al. Machine learning interatomic potential: Bridge the gap between small-scale models and realistic device-scale simulations. iScience, 2024, 27(5): 109673

[65] Zhang X G, Li K Q, Zhou J, et al. Chalcogenide phase-change materials: Unveiling new horizons with big data and machine learning. J Mater Chem C, 2025, 13(28): 14168

[66] Sosso G C, Bernasconi M. Harnessing machine learning potentials to understand the functional properties of phase-change materials. MRS Bull, 2019, 44(9): 705

[67] Zhang L F, Han J Q, Wang H, et al. Deep potential molecular dynamics: A scalable model with the accuracy of quantum mechanics. Phys Rev Lett, 2018, 120(14): 143001

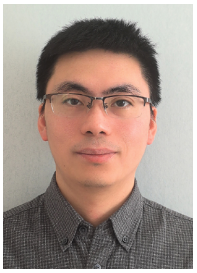
[68] Chen N K, Wang B Q, Niu M, et al. Intensive structural disorder induces electronic delocalization: Amorphous solid-liquid transition in ovonic threshold switching materials. Adv Funct Mater, 2024, 34(52): 2410622

[69] Chen N K, Li X B, Bang J, et al. Directional forces by momentumless excitation and order-to-order transition in Peierls-distorted solids: The case of GeTe. Phys Rev Lett, 2018, 120(18): 185701

[70] Chen N K, Bang J, Li X B, et al. Optical subpicosecond nonvolatile switching and electron-phonon coupling in ferroelectric materials. Phys Rev B, 2020, 102(18): 184115

[71] Linker T, Nomura K I, Aditya A, et al. Exploring far-from-equilibrium ultrafast polarization control in ferroelectric oxides with excited-state neural network quantum molecular dynamics. Sci Adv, 2022, 8(12): eabk2625

[72] Li H, Wang Z, Zou N L, et al. Deep-learning density functional theory Hamiltonian for efficient *ab initio* electronic-structure calculation. Nat Comput Sci, 2022, 2(6): 367



陈念科 2007−2018 年于吉林大学分别攻读并获得学士、硕士和博士学位, 目前为吉林大学电子科学与工程学院 (集成电路学院) 微电子科学与工程系副教授, 吉林大学唐敖庆青年学者, 主要研究方向为相变信息存储技术中的关键半导体物理问题研究. 相关成果以第一/通信作者发表于 *Physical Review Letters*、*Nature Communications*、*Nano Letters*、*Advanced Materials*、*Advanced Functional Materials* 等期刊. 曾获吉林大学优秀青年教师培养计划、吉林大学“金种子”人才计划、全国高等学校电子信息青年授课竞赛全国二等奖等荣誉.



李贤斌 2001−2010 年于吉林大学分别攻读并获得学士、硕士和博士学位, 国家级领军人才,目前为吉林大学电子科学与工程学院 (集成电路学院) 副院长, 微电子科学与工程系教授, 主要研究方向为信息存储器和光电器件的半导体物理问题研究. 任全国光学青年学术论坛副主席、全国电子信息青年科学家论坛 (第四届、第五届) 半导体基础物理方向召集人、第三届全国半导体物理青年论坛组织者等学术兼职.

Review

AI-driven explorations of the key atomic mechanism of phase-change memory semiconductors*

CHEN Nianke†, WANG Baiqian, ZHAO Tianyu, and LI Xianbin†

State Key Laboratory of Integrated Optoelectronics, College of Electronic Science and Engineering (College of Integrated Circuits), Jilin University, Changchun 130012, China

Abstract: Phase-change memory (PCM) is a promising candidate for storage-class memory and in-memory computing, thanks to its non-volatility, high speed, reliability, and scalability, thus offering the potential to circumvent the computational limitations of conventional computing architectures. The PCM working principle relies on reversible phase transition between crystalline phase (with low resistance or high reflectivity) and amorphous phase (with high resistance or low reflectivity) of PCM materials. Therefore, understanding structural and electronic transitions at the atomic scale is fundamental to developing new high-performance materials and addressing key challenges for PCM applications. However, conventional methods for uncovering atomic-scale mechanisms, such as first-principles calculations, are limited by computational cost, leading to a bottleneck in PCM device-level simulations. The recent emergence of the artificial-intelligence (AI) driven research paradigm, namely, AI for science, has brought new opportunities and renewed optimism to the research community. This paper reviews recent advances in applying AI methods to investigate PCM material properties and phase-change mechanisms, including explorations of structure-property relationships, optimization of existing PCM materials, design of new PCM materials, development of machine-learning potentials and PCM device-level simulations. The major challenges in AI-driven PCM research are then discussed. Finally, new ideas and future directions based on AI methods to address the key obstacles for PCM commercialization are presented.

Keywords: phase-change memory; semiconductor physics; AI for science

DOI: 10.3724/cjos.2025.025 **CSTR:** 32479.14.cjos.2025.025

Citation: Chen N K, Wang B Q, Zhao T Y, *et al*. AI-driven explorations of the key atomic mechanism of phase-change memory semiconductors. *Chin. J. Semicond.*, 2026, 1(3): 031101

*Project supported by the National Science and Technology Major Project (2022ZD0117600), the National Natural Science Foundation of China (12274172, 12274180), and the Science and Technology Development Plan Project of Changchun (2024GZZ07).

†Corresponding author: CHEN Nianke, chennianke@jlu.edu.cn; LI Xianbin, lixianbin@jlu.edu.cn