



Views & Comments

用于材料发现的生成式人工智能——无需理解的设计

胡建军^a, 李琴^b, Nihang Fu^a^a Department of Computer Science and Engineering, University of South Carolina, Columbia, SC 29208, USA^b College of Big Data Statistics, Guizhou University of Finance and Economics, Guiyang 550025, China

1. 设计如同积木搭建

21 世纪, 人类社会正面临诸多根本性挑战, 包括气候变化、能源危机以及癌症和新冠病毒肺炎 (COVID-19) 等公共卫生危机。应对这些挑战的关键在于发现新型材料、分子、蛋白质和药物, 然而, 由于原子间相互作用的复杂性、稳定结构形成的物理/化学/几何约束, 以及结构与功能之间的复杂关系, 设计此类功能性原子结构极具挑战。

当前主流的材料设计范式与大多数其他工程设计活动一样, 采用理性设计方法, 强调对结构-功能关系的因果理解, 并依赖于专家知识和明确的设计规则。典型的设计过程从有限数量原型的拓扑设计开始, 最终以参数化设计结束。然而, 传统材料设计范式在设计能够有效满足需求的非凡功能材料时面临越来越多的挑战: 其有限的搜索能力通常导致在庞大的化学设计空间中产生次优解; 处理大量隐性知识和约束变得困难, 无法有效利用这些规则进行设计空间探索; 需要过多明确的设计规则; 以及难以设计高度受限的结构, 如周期性无机晶体。

本文旨在论证从理性材料设计到数据驱动的深度生成材料设计范式的变革性转变。在该范式中, 已知的材料数据被输入深度生成模型中, 以学习原子结构的显性和隐性知识, 并利用这些知识进行高效的生成。这一转变受

到了近期两项重大人工智能进展的启发。其一, 数据驱动的深度学习算法 AlphaFold2 表明, 深度学习模型可以通过学习已知蛋白质结构的隐性规则和约束, 在蛋白质结构空间中进行有效采样, 从而解决蛋白质结构预测中固有的“大海捞针”问题。其二, 基于深度学习的人工智能生成内容 (AIGC) 技术正加速生成逼真的图像、视频、文本、音乐和人声。如表 1 显示, 尽管数字作品与原子结构之间存在显著差异, 但图像和文本的设计与蛋白质、分子及材料的设计存在诸多共同特征, 不同层级的构建块被组装在一起, 形成满足各类语法、物理、化学或几何约束的特定稳定或有意义的结构。

与早期依赖明确定义的构建块、生成规则或语法的生成设计系统[1]相比, 深度生成设计范式采用神经网络来学习组装可合成和稳定结构的物理或化学规则。因此, 深度生成材料设计提供了一种新的方法和理念, 它从动态过程及其结果的角度看待材料, 其中神经网络不仅可以用于学习静态的原子间相互作用, 还可以学习自组装和自组织动态过程。正如自然界通过进化利用 DNA 的物理装置作为蛋白质合成和生物化学合成规则的信息载体一样, 神经网络同样也可以通过从已知材料或计算模拟中学习设计规则, 来实现自然界的材料设计方式。类似于雌性青蛙在不了解青蛙如何从受精卵发育而来的情况下即可产下青蛙, 深度生成设计以一种“设计-无需理解”的

表1 设计图像和文本与设计蛋白质、材料和分子的比较

Domain	Building block	Structures
Image/video	Pixel \rightarrow patches/strips	Objects/shapes
Language	Characters \rightarrow words \rightarrow sentences \rightarrow paragraphs	Text
Proteins	Amino acids (AAs) \rightarrow AA sequence \rightarrow secondary structures	Three-dimensional protein structures
Materials	Atoms \rightarrow bonds \rightarrow polyhedrons	Crystal structures
Molecules	Atoms \rightarrow bonds \rightarrow groups	Organic molecules

模式运作。该方法无需完全理解潜在机制，即可实现设计目标。

2. 设计领域的生成式人工智能

生成式人工智能的思想可追溯至20世纪50年代，由克劳德·香农提出的用于语言生成的马尔可夫链拉开了序幕。20世纪80年代，研究人员开始探索用于图像和音乐生成的霍普菲尔德网络和玻尔兹曼机。随后，在20世纪90年代，概率图模型，如隐马尔可夫模型和高斯混合模型，成为生成式AI的主流方法。然而，生成式AI的真正突破发生在2014年，随着生成对抗网络（GAN）[2]的问世，人工智能才真正能够创造出逼真的图像、视频、文本

和音频。通过对大量现有数据样本进行训练，现代深度神经网络模型能够生成惊人逼真的数字作品，这些作品目前被ChatGPT或其他软件广泛应用于人工智能生成内容（AIGC）领域。这些模型通过其网络学习已知样本中构建块之间微妙而复杂的模式、节奏、风格、几何约束或相互依赖关系，并利用这些隐性知识来有效和高效地生成新样本。受到这些进展的启发，研究人员开始探索深度生成模型在更广泛领域的应用。

深度生成模型已被日益广泛地应用于DNA和蛋白质（序列和结构）、分子（组成和构象）[3]、材料（组成和结构）以及工程设计[4]等领域。不同研究中所使用的标记（tokens）或构建模块（buildingblocks）有所差异，但多数研究工作都采用了一组通用的生成模型架构（图1）。

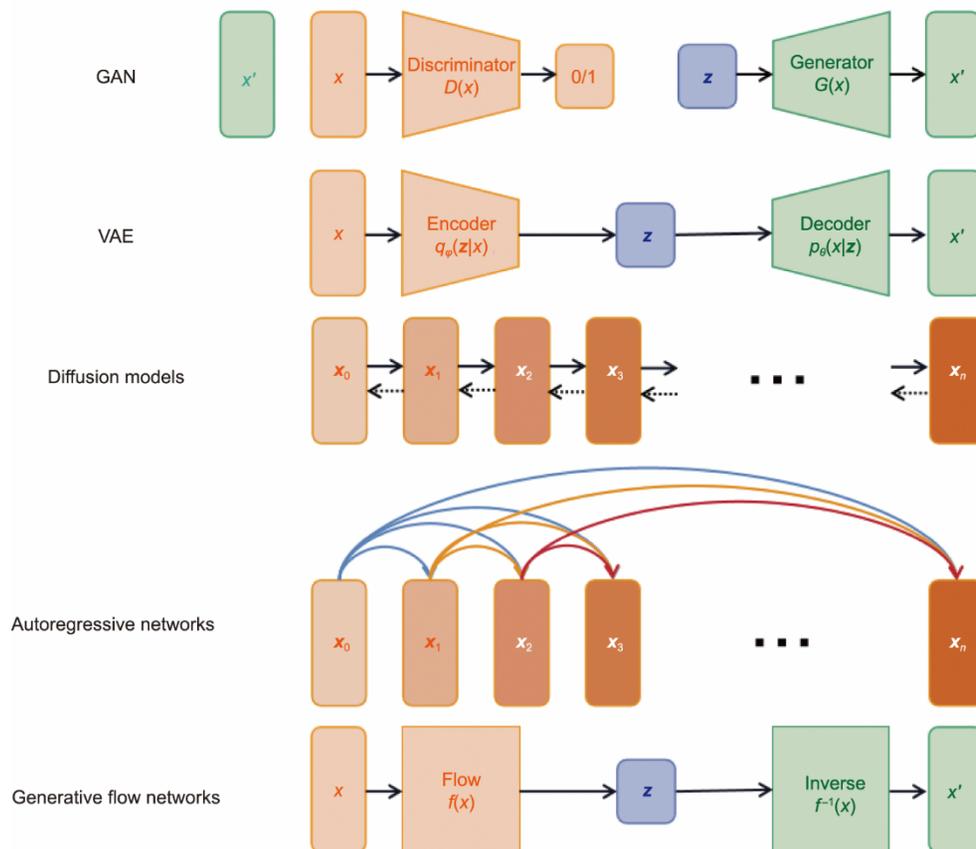


图1. 深度生成式神经网络模型。

变分自编码器 (VAE) 模型由编码器和解码器组成。编码器将原始输入信息 x 最大限度地压缩到低维潜在空间 z , 使解码器能够以最小的重构误差将其重构为 x' 。潜在空间的正则化则通过最小化潜在分布与标准高斯分布之间的 Kullback-Leibler 散度来实现。GAN 由生成器和判别器组成。生成器学习生成符合训练数据分布的新样本, 而判别器则学习区分真实样本和生成样本。这两个部分通过对抗训练的方式进行共同优化: 生成器不断提升生成样本的真实度, 而判别器则不断提高区分真假样本的能力。相较于 VAE 所依赖的重构损失, GAN 模型更擅长捕捉训练数据分布中的语义信息, 并生成更具多样性的新样本。

扩散模型 (diffusion model) 通过逐步向训练数据添加高斯噪声破坏原始数据, 并利用去噪神经网络学习逆向恢复数据的过程。训练完成后, 该模型可以通过将随机采样噪声输入学习到的去噪过程, 进而生成数据。扩散模型的一项关键优势在于, 它能够生成高度真实且多样化的图像, 而无需复杂的对抗训练。此外, 通过根据类别标签或语义嵌入等附加输入调节扩散过程, 扩散模型还可生成具有可控属性的图像。另一类重要的生成模型是自回归模型 (autoregressive models), 如 GPT 等常用于文本生成的语言模型。此类模型通过预测序列中前序词语的下一个词语来生成文本。自回归模型通常在大规模文本语料库上进行训练, 以学习每个词语在给定上下文中的概率分布。此外, 它们也可以用于图像生成, 以像素为单位逐步生成图像。

生成流网络模型 (GFlowNet) [5] 是一种概率模型,

其通过对可能的构建模块进行概率采样, 并依次添加新模块来构建目标对象。GFlowNet 以与奖励成比例的频率构建对象, 并通过学习代理模型 (surrogate model) 的奖励函数进行训练。训练完成后, 该生成模型可以利用 GFlowNet 在庞大的化学设计空间中进行智能采样, 然后使用代理模型对生成的结构进行优先级排序。因此, GFlowNet 适用于大规模化学设计问题, 能够有效提高搜索效率。

3. 生成式人工智能驱动的材料发现

在传统的材料设计中, 研究人员主要依赖试错法来测试新材料, 这一过程通常耗时且成本高昂。生成式材料设计旨在加速这一过程, 通过学习和利用化学、几何及物理约束, 高效生成符合特定标准的新材料, 如可合成性、稳定性、导电性或光学特性。随后, 这些材料可以在实验室中进行合成和测试, 从而有望带来新的科学发现。图2展示了用于材料设计的代表性生成模型。

3.1. 材料成分的生成式设计

材料成分设计的目标是发现能够合成并形成稳定晶体结构的材料成分。这些成分可用于指导实验合成, 或将其输入晶体结构预测算法[6]以获取其可能的稳定结构。此外, 材料成分还可用于基于成分的材料性质机器学习模型 (如弹性常数或带隙预测) 进行成分筛选。然而, 材料成分的生成并非易事, 主要面临以下三个核心挑战。

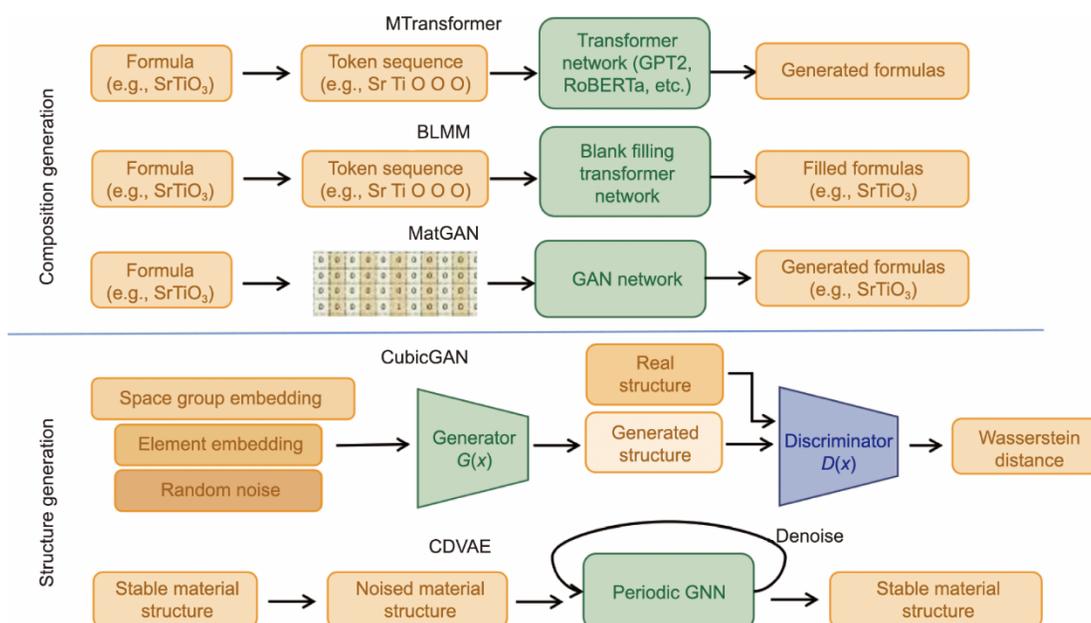


图2. 深度生成式材料设计框架。

(1) “大海捞针”问题[7]: 对于三元材料, 可能的成分组合超过 109 种, 而四元材料的组合数更是超过 1012 种, 其中大多数组合甚至不满足基本化学规则。例如, 假设每种元素的原子数不超过 8, 仅有 6.7% (四元) 和 7.8% (五元) 的组合满足电荷中和与电负性平衡的约束。

(2) 成分-结构关系的复杂性: 成分与其可合成性及稳定结构形成能力之间的关系极为复杂, 受多种化学和几何约束的影响, 难以直接解析。

(3) 质量评估的挑战: 在缺少结构信息的情况下, 很难评估生成成分的质量, 例如, 判断其可合成性或结构稳定性, 从而筛选出有潜力的候选材料。

早期的生成式设计研究尝试使用实数向量编码材料成分中各元素的原子数量, 并借助条件变分自编码器 (Conditional VAE) 和条件生成对抗网络 (Conditional GAN) 进行生成[7]。然而, 由于原子数量采用实数值表示, 导致生成的大部分成分不符合化学规则, 难以通过电荷中和与电负性平衡等基本化学约束进行有效筛选。为了解决这一问题, 我们认识到离散化编码对元素原子数表示的重要性, 并提出了 MATGAN 材料成分生成器[8]。MATGAN 基于 GAN, 并采用独热 (one-hot) 二进制矩阵来表示材料成分。这种编码方式极大地增强了卷积神经网络 (CNN) 在 GAN 训练过程中对 ICSD 和 Material Project 数据库中已知材料的复杂化学规则和模式的学习能力。实验结果表明, 当 GAN 在 ICSD 数据库中以符合化学规则的样本进行训练时, 即使 GAN 模型中并未显式施加这些化学规则, MATGAN 生成的成分中化学有效 (电荷中和且电负性平衡) 的比例可达 84.5%。这表明 MATGAN 具备学习隐式化学成分规则的能力, 从而能够在生成过程中利用这些隐式约束, 提高生成的成分形成稳定可合成化合物的可能性。

像 SrTiO_3 这样的材料成分可以自然地表示为元素符号序列, 如 SrTiOOO。这一特点启发我们利用现代生成式语言模型, 如 GPT 和 BERT, 来构建材料成分生成模型。这些语言模型在文本生成、分子生成和蛋白质序列生成等任务中已经取得了显著成果。在研究[9]中, 我们开发并系统评估了七种主流语言模型 (包括 GPT、GPT-2、GPT-Neo、GPT-J、BLMM、BART 和 RoBERTa), 并将其作为 Material Transformer 算法用于材料成分生成。实验使用了 ICSD、OQMD 和 Materials Project 数据库中的六种不同数据集, 这些数据集包含或不包含非电荷中性或电负性不平衡的样本。研究结果表明, 基于因果语言模型 (如 GPT) 的材料变换器 (transformer) 生成的材料成分具有较高的化学有效性, 其中电荷中性比例最高可达 97.54%, 电负

性平衡比例达 91.40%, 相比于伪随机采样算法, 生成结果的化学有效性提高了六倍以上。这一结果表明, 语言模型具备学习并捕捉隐含化学规律和约束的能力, 从而能够有效生成符合化学规则的材料成分。此外, 为了提升模型的可解释性, 我们引入填空式概率语言模型 (blank-filling probabilistic language model), 并应用于材料成分生成任务[10]。基于此, 我们提出了 Crystal Transformer 算法, 在电荷中性和电负性平衡等指标上均实现了最佳生成性能。同时, 该模型还允许研究人员对给定的材料成分进行定向调整, 结合模型学习到的材料化学规律, 进一步探索材料设计空间, 这对材料掺杂研究具有重要应用价值。

在材料成分生成式设计中, 一个关键的决策是如何评估生成模型的性能, 特别是在缺乏结构信息的情况下。虽然电荷中性、电负性平衡以及预测的形成能可以作为初步的性能衡量标准, 但我们还可以从新颖性、独特性和召回率等角度对生成模型进行评估。其中, 召回率尤为重要, 如果一个生成模型能够重新发现大多数被排除的但曾经被成功合成的材料成分, 那么这强烈表明其具备较强的材料生成能力。

3.2. 晶体结构的生成式设计

从头设计并生成具有良好合成性和稳定性的新型晶体材料极具挑战, 其难点在于材料组成与稳定结构之间存在高度复杂的关联。相较于有机分子等其他结构, 晶体材料倾向于呈现高对称性的周期性排布, 从而导致其设计空间呈现高度约束且多峰的复杂特性。此外, 构成晶体材料的元素种类繁多 (>85 种), 以及原子间复杂的相互作用, 都进一步增加了设计难度。

近年来, 数据驱动的晶体结构生成设计领域蓬勃发展, 涌现出一系列基于深度生成模型和多种晶体编码策略的方法。其中, iMatGen 是早期尝试晶体结构生成的代表性算法。该方法采用变分自编码器 (VAE) 架构, 并以 V_xO_y 族化合物的结构数据进行训练, 成功发现了 40 种相对稳定的结构, 其 $E_{\text{min}} E_{\text{max}} < 80 \text{ meV/原子}$ 。Kim 等[11]的研究表明, 基于 GAN 的生成模型, 结合点云数据作为输入, 能够有效生成稳定的 Mg-Mn-O 三元化合物。Court 等[12]则基于 iMatGen 的体素化晶体表示方法, 训练了一种条件深度特征一致性 VAE, 旨在生成超越特定化学体系的新型晶体。然而, 该模型需要依赖 UNet 分割网络将预测的电子密度映射到原子位点, 这在一定程度上制约了其性能。

针对晶体结构的对称性特点, 我们团队开发了 CubicGAN [13], 一种基于 GAN 的通用立方晶系晶体结构生成

模型。经验证，该模型成功生成了506种新型的、具有潜在稳定性的假设材料，如 $\text{Li}_6\text{N}_6\text{Cl}$ 和 CaCO_6 ，并通过声子谱计算进行了验证。我们进一步研究发现，将额外的对称性原理和基于物理的约束条件融入生成模型，能够显著提升生成模型的性能。正如我们提出的PGCGM算法[14]所示，该算法能够生成超过30个空间群的晶体结构。在生成式晶体材料设计领域，另一项重要进展是基于扩散模型的CDVAE [15]。该方法训练了一个可以在扩散过程中逐步生成材料的解码器。这种基于神经网络的扩散模型能够引导原子坐标向能量更低的状态演化，并通过选择合适的原子类型来满足相邻原子之间的成键偏好。此外，CDVAE还能够模拟跨周期性边界的相互作用，并保证置换、平移、旋转和周期性等操作的不变性，从而进一步提升其性能。该模型已被成功应用于发现数千种具有潜在应用价值的二维材料。然而，其在三维材料（特别是高对称性材料）生成方面的能力仍有待提高[14]，这一问题已在其最新的模型Mattergen [16]中得到初步改善。

值得注意的是，尽管前文对生成式材料设计的讨论主要集中在无机体系，但其核心原理与模型亦被广泛应用于蛋白质[17]、有机材料[3,18]以及结构材料的设计中。尤其是在结构材料领域，语言模型与扩散模型已成为正向设计与逆向设计的有力工具[19]。

4. 挑战与机遇

生成式材料设计目前仍处于发展初期，面临诸多严峻挑战，同时也蕴含着巨大的发展机遇。这些亟待解决的挑战涵盖了从算法模型构建、训练数据集构建，到设计目标的设定等诸多方面。

挑战 1：多目标函数和复杂约束条件下的可控生成设计。如何实现具有多目标函数和复杂物理化学/几何约束条件下的可控生成设计是当前面临的主要挑战之一。更为棘手的是，现有的大部分性能模拟代码不具备可微性，导致难以将这些性能目标直接纳入损失函数中，从而有效指导模型的训练和材料结构的生成。

挑战 2：大型语言模型（LLM）与生成模型的融合。尽管深度语言模型已成功应用于蛋白质[20]和基于SMILES编码的分子设计[18]，但由于晶体材料固有的高对称性约束以及难以寻觅稳定、可合成的晶体结构，将LLM应用于晶体材料设计依然极具挑战。

挑战 3：高效的候选材料筛选与验证。许多材料设计约束（包括显性和隐性约束）难以直接嵌入生成模型，因此必须在生成样本后进行验证。然而，可合成性、力学稳

定性以及热力学稳定性等关键筛选标准的计算成本高昂。若进一步考虑可制造性，问题将更为复杂。如何提高生成在物理和化学上可行的材料的效率，并训练快速、精确的评估模型，是当前亟待解决的核心问题。

挑战 4：突破现有框架，实现创造性生成。材料设计的最终目标是发现具有卓越性能的全新材料。然而，现有生成模型往往倾向于生成与训练集高度相似的样本，难以跳出现有框架。此外，神经网络通常擅长插值计算，但在外推预测方面表现不足，导致超出训练分布范围的样本性能预测精度下降，进而误导生成模型产生无效的设计方案。

挑战 5：小样本数据集下的生成式设计。受限于实验成本高昂和DFT计算资源有限，当前材料数据库中关于宏观性能（如热导率和压电性）的数据相对匮乏。然而，与此同时，存在着大量的未标记材料结构信息。如何利用先进的机器学习技术，如预训练模型和物理信息神经网络，充分挖掘这些未标记数据，将是解决小样本问题的关键。此外，探索基于混合精度数据集训练的代理模型，并研究具有卓越泛化性能的分外机器学习模型，将为新材料的发现提供有力支撑。

随着OpenAI推出的Sora等更先进的深度生成AI技术的出现，材料与结构的生成设计研究将在未来几年迎来重大变革。我们有理由相信，这一变革将为应对全球气候变化、能源问题及健康挑战提供重要支持。

References

- [1] McCormack J, Dorin A, Innocent T. Generative design: a paradigm for design research. In: Redmond J, Durling D, de Bono A, editors. Proceedings of Futureground, Design Research Society International Conference; 2004 Nov 17–21; Melbourne, VIC, Australia. Clayton: Monash University Publishing; 2004.
- [2] Goodfellow I, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, et al. Generative adversarial networks. *Commun ACM* 2020;63(11):139–44.
- [3] Bilodeau C, Jin W, Jaakkola T, Barzilay R, Jensen KF. Generative models for molecular discovery: recent advances and challenges. *Wiley Interdiscip Rev Comput Mol Sci* 2022;12(5):e1608.
- [4] Regenwetter L, Nobari AH, Ahmed F. Deep generative models in engineering design: a review. *J Mech Des* 2022;144(7):071704.
- [5] Bengio E, Jain M, Korablyov M, Precup D, Bengio Y. Flow network based generative models for non-iterative diverse candidate generation. In: Ranzato M, Beygelzimer A, Dauphin Y, Liang PS, Wortman Vaughan J, editors. *Advances in neural information processing systems* 34. Red Hook: Curran Associates, Inc.; 2021. p. 27381–94.
- [6] Davies DW, Butler KT, Jackson AJ, Morris A, Frost JM, Skelton JM, et al. Computational screening of all stoichiometric inorganic materials. *Chem* 2016; 1(4):617–27.
- [7] Sawada Y, Morikawa K, Fujii M. Study of deep generative models for inorganic chemical compositions. 2019. arXiv:1910.11499.
- [8] Dan Y, Zhao Y, Li X, Li S, Hu M, Hu J. Generative adversarial networks (GAN) based efficient sampling of chemical composition space for inverse design of inorganic materials. *npj Comput Mater* 2020;6(1):84.
- [9] Fu N, Wei L, Song Y, Li Q, Xin R, Omees SS, et al. Material transformers: deep learning language models for generative materials design. *Mach Learn Sci*

- Technol 2023;4(1):015001.
- [10] Wei L, Li Q, Song Y, Stefanov S, Siriwardane EMD, Chen F, et al. Crystal transformer: self-learning neural language model for generative and tinkering design of materials. 2022. arXiv:2204.11953.
- [11] Kim S, Noh J, Gu GH, Aspuru-Guzik A, Jung Y. Generative adversarial networks for crystal structure prediction. ACS Cent Sci 2020;6(8):1412–20.
- [12] Court CJ, Yildirim B, Jain A, Cole JM. 3-D inorganic crystal structure generation and property prediction via representation learning. J Chem Inf Model 2020;60(10):4518–35.
- [13] Zhao Y, Al-Fahdi M, Hu M, Siriwardane EMD, Song Y, Nasiri A, et al. High throughput discovery of novel cubic crystal materials using deep generative neural networks. Adv Sci 2021;8(20):2100566.
- [14] Zhao Y, Siriwardane EMD, Wu Z, Fu N, Al-Fahdi M, Hu M, et al. Physics guided deep learning for generative design of crystal materials with symmetry constraints. npj Comput Mater 2023;9(1):38.
- [15] Xie T, Fu X, Ganea OE, Barzilay R, Jaakkola T. Crystal diffusion variational autoencoder for periodic material generation. 2021. arXiv:2110.06197.
- [16] Zeni C, Pinsler R, Zügner D, Fowler A, Horton M, Fu X, et al. MatterGen: a generative model for inorganic materials design. 2023. arXiv:2312.03687.
- [17] Buehler MJ. Generative pretrained autoregressive transformer graph neural network applied to the analysis and discovery of novel proteins. J Appl Phys 2023;134(8):084902.
- [18] Luu RK, Wysokowski M, Buehler MJ. Generative discovery of *de novo* chemical designs using diffusion modeling and transformer deep neural networks with application to deep eutectic solvents. Appl Phys Lett 2023;122(23):234103.
- [19] Lew AJ, Buehler MJ. Single-shot forward and inverse hierarchical architected materials design for nonlinear mechanical properties using an attention diffusion model. Mater Today 2023;64:10–20.
- [20] Ni B, Kaplan DL, Buehler MJ. ForceGen: end-to-end *de novo* protein generation based on nonlinear mechanical unfolding responses using a language diffusion model. Sci Adv 2024;10(6):eadl4000.