nature computational science

Perspective

https://doi.org/10.1038/s43588-024-00632-5

Designing semiconductor materials and devices in the post-Moore era by tackling computational challenges with data-driven strategies

Received: 16 August 2023

Accepted: 18 April 2024

Published online: 23 May 2024

Check for updates

Jiahao Xie[1,3], Yansong Zhou[2,3], Muhammad Faizan[1], Zewei Li[1], Tianshu Li[1], Yuhao Fu[2], Xinjiang Wang[1] ✉ & Lijun Zhang[1] ✉

In the post-Moore's law era, the progress of electronics relies on discovering superior semiconductor materials and optimizing device fabrication. Computational methods, augmented by emerging data-driven strategies, offer a promising alternative to the traditional trial-and-error approach. In this Perspective, we highlight data-driven computational frameworks for enhancing semiconductor discovery and device development by elaborating on their advances in exploring the materials design space, predicting semiconductor properties and optimizing device fabrication, with a concluding discussion on the challenges and opportunities in these areas.

Moore's law, which predicts the continual shrinkage of silicon-based electronic devices, has encountered challenges due to quantum uncertainty and power-density issues[1,2]. Although emerging computing paradigms such as quantum[3] and neuromorphic computing[4] offer new architectural possibilities, they remain impractical for general applications. Consequently, the focus has shifted to advancing semiconductor materials to improve device performance in the post-Moore era. Transistors utilizing emerging materials such as carbon-based materials and two-dimensional (2D) transition metal dichalcogenides have shown substantial potential[5,6]. Recent progress in perovskite-based photovoltaics[7], novel luminescent materials for light-emitting diodes[8], and semiconductors with ferroelectric and ferromagnetic properties[9] have shown promising developments. However, practical applications have remained hindered by challenges in material growth and device fabrication[10,11]. Designing superior semiconductors and optimizing device fabrication processes are considered crucial for the future development of electronics. To address the well-known helium shortage, one potential solution is to find high-performance semiconductor alternatives to silicon that can maintain their performance at high temperatures.

Discovering semiconductors with desired properties essentially involves the inverse solution of the quantum mechanical Schrödinger equation, a task that remains a substantial challenge[12,13]. Material design research focuses on navigating potential and performance landscapes that involve the forward solutions of the Schrödinger equation. This approach aims to identify materials corresponding to the minima of the potential surface and maxima of the performance surface, as illustrated in Fig. 1a. New materials generation and material properties' prediction are two fundamental aspects that characterize current materials design research. New materials generation research aims to propose the most promising candidate materials for further validation through property-prediction methods. Combinatorial sampling as a traditional new materials generation method is widely used in high-throughput material searches[14]. Structural prototype-based elemental substitution strategies, such as homologous element substitution[15] and ion transmutation[16], are the fundamental methods for combinatorial sampling. However, these methods are limited by the need for extensive and costly computational validation and the inability to explore unknown material structures. Material performance-oriented inverse design is an emerging paradigm for new materials generation[13,17], which

[1]State Key Laboratory of Integrated Optoelectronics, Key Laboratory of Automobile Materials of MOE, Key Laboratory of Material Simulation Methods & Software of MOE, and School of Materials Science and Engineering, Jilin University, Changchun, China. [2]State Key Laboratory of Superhard Materials, International Center of Computational Method and Software, School of Physics, Jilin University, Changchun, China. [3]These authors contributed equally: Jiahao Xie, Yansong Zhou. ✉e-mail: xinjiang_wang@jlu.edu.cn; lijun_zhang@jlu.edu.cn

https://doi.org/10.1038/s43588-024-00632-5

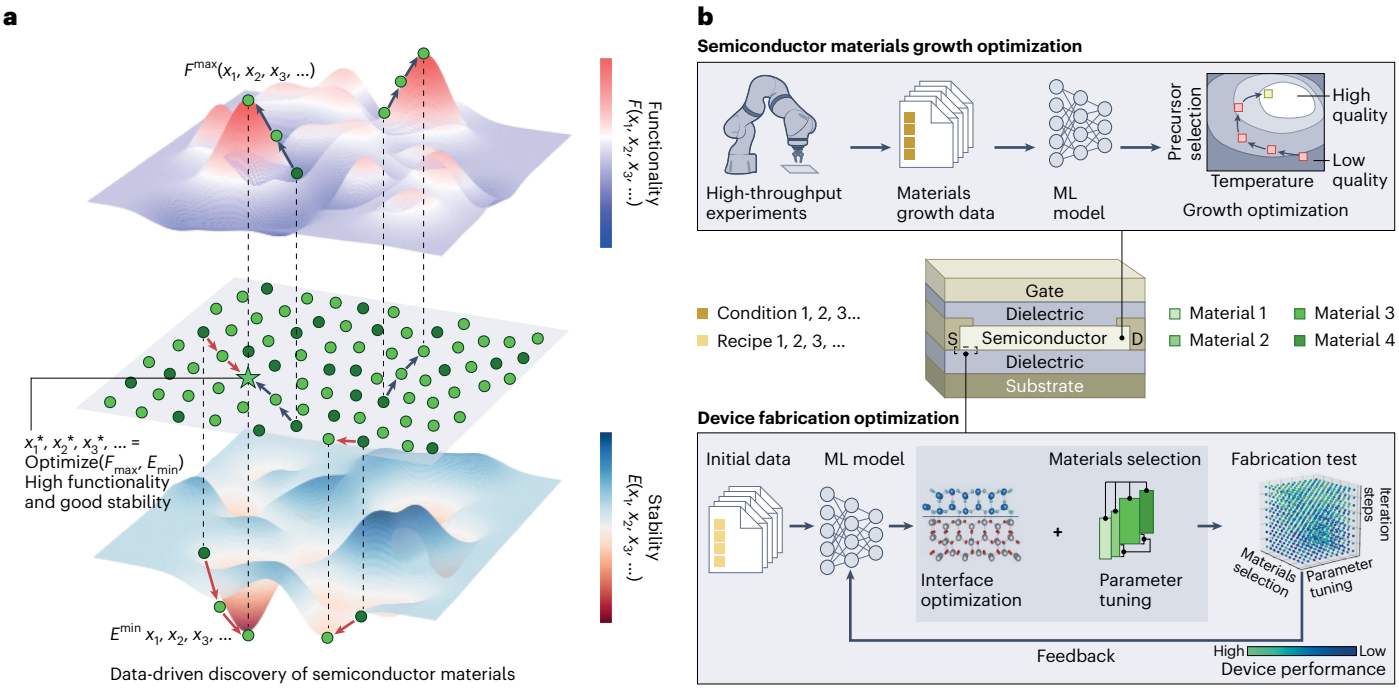

Fig. 1 | Schematics showing semiconductor material design and electronic device optimization. a, Schematic of data-driven discovery of semiconductor materials. The central plane represents the material design space, where potential materials are identified by coordinates $(x_1, x_2, x_3, ...)$, with each x_n reflecting a degree of freedom of the material. The top surface depicts the material's functionality surface defined by $F(x_1, x_2, x_3, ...)$, marked by a peak at $F_{max}(x_1, x_2, x_3, ...)$ indicating the optimal functionality. The bottom surface outlines the material's potential energy surface defined by $E(x_1, x_2, x_3, ...)$, with its lowest point at $E_{min}(x_1, x_2, x_3, ...)$, indicating the optimal stability. Designing semiconductor materials involves optimizing both functionality and stability, that is, performing Optimize (F_{max}, E_{min}) to identify the optimal material $(x_1^*, x_2^*, x_3^*, ...)$ that combines high performance and good stability within the design space. Blue represents poor functionality or stability, while red indicates good functionality or stability. **b**, Schematics highlighting electronic device performance optimization through a data-driven approach, namely, semiconductor material growth optimization and device fabrication optimization. S, source; D, drain.

helps to explore unknown material structures and reduce the computational validation required to discover optimal new materials. Exploring the material's potential energy surface or performance surface is key for inverse design, which can be achieved by an optimization algorithms-guided search[18] or by physical rules-guided sampling[19]. Despite the continuous evolution of such methods, material property prediction based on first-principles quantum mechanical calculations, such as density functional theory (DFT) calculations, remain as the efficiency bottleneck for large-scale new material discovery research, which requires the exploration of vast materials space with more than 10^{10} materials[20]. In addition, the limited spatial and temporal simulation scales of DFT calculations pose challenges in the simulation of complex structures and processes in device fabrication.

Data-driven strategies are revolutionizing semiconductor design and device optimization by accelerating material explorations through machine learning (ML) prediction models[21,22], introducing new paradigms in design space exploration with generative models[23,24] and streamlining device manufacturing parameter optimization with experimental planning models[25,26] (Fig. 1b). Although still in the development process and not surpassing first-principles calculations, these methods hold the potential for groundbreaking advancements in electronic device performance. From this perspective, we first examine the two critical aspects of semiconductor material design: the strategies for exploring design space and the methods for predicting semiconductor performance. We then discuss data-driven methods for optimizing electronic devices in terms of semiconductor growth and device fabrication. In each area, we highlight how data-driven approaches can enhance existing methods or forge new paths. We conclude with an outlook on potential solutions and opportunities for addressing computational challenges with data-driven strategies.

Exploring materials space for new semiconductor design

High-throughput hierarchical material screening

High-throughput hierarchical screening consists of an iterative enumeration method. In each iteration, a filter with increasing cost is used to narrow down the material pool, ultimately selecting materials that meet all criteria[27,28], as illustrated in Fig. 2a. Key aspects of such an approach include the accuracy and efficiency of materials performance prediction methods, along with the construction of an initial materials design space, which is often based on existing structural prototypes and element-related variations. This direct approach has led to the guidance of experimental research[29–31], such as the finding of non-toxic double halide perovskites for optoelectronic applications[16,32,33], the computational exfoliation of known bulk materials to find potential 2D materials including semiconductors[34,35] and the design of efficient molecular organic light-emitting diodes by extensive virtual screening[36]. However, using DFT-based stability and performance evaluations remains impractical for searching vast material spaces exceeding 10^{10} materials[20].

ML prediction models offer the potential for accelerating high-throughput screening. The challenge lies in developing models with strong generalization abilities to effectively expedite the discovery of new materials, ensuring that the efficiency gains from model predictions outweigh the training costs. Solutions include training the initial model based on existing computational datasets and employing adaptive strategies such as active learning and transfer learning to continuously enhance model performance. Recent studies by Merchant et al.[37] and Schmidt et al.[20] demonstrated the utilization of ML-based energy prediction models for large-scale material screening. Merchant et al. iteratively trained ML models for energy prediction using an active

https://doi.org/10.1038/s43588-024-00632-5

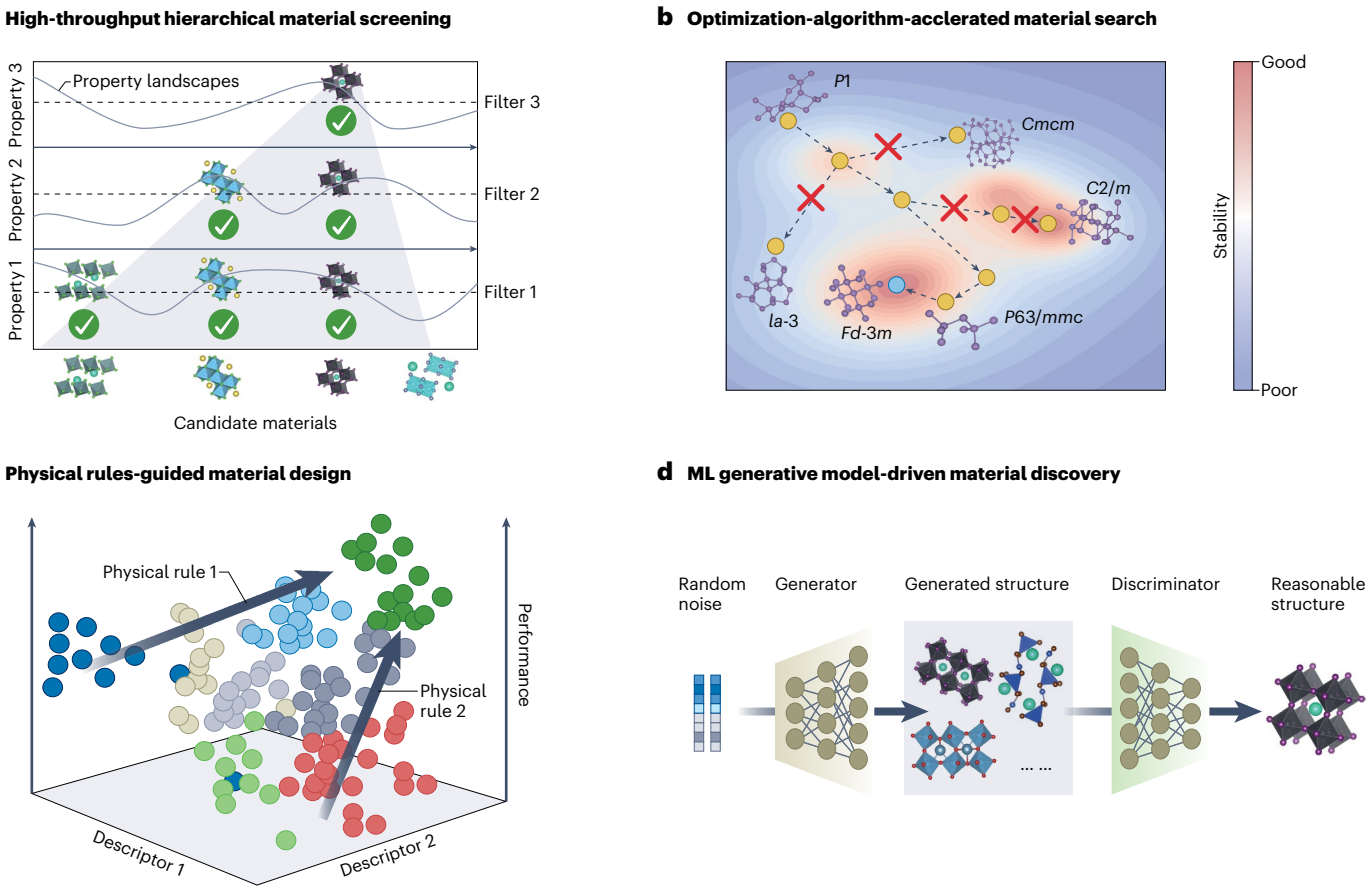

Fig. 2 | Summary of data-driven strategies for material design space exploration. a, Schematic of search based on filter screenings. Multiple filters are used to screen the candidate materials, with each selecting materials corresponding to the higher portion of the property landscape, ultimately yielding materials that satisfy all property criteria. **b**, Schematic of search based on the optimization algorithm. The background contour map represents the material potential surface. The black dashed lines symbolize the material selection process, directed by the optimization algorithm. Space groups are labeled near the example material structures. **c**, Schematic of materials design guided by physical rules. Circles with different colors represent different kinds of materials, the horizontal coordinates are the different descriptors describing the material and the vertical coordinates are the material performance. The blue arrows are the physical rules found in this descriptor space, representing the mapping relationships from simple descriptors to performance. Guided by the mapping relationships, new high-performance materials can be designed from simple material descriptors. **d**, Simplified schematic of a GAN network illustrating how ML generative models can accelerate the discovery of new materials. Panel **c** adapted with permission from ref. 150, American Chemical Society.

learning strategy and expanded the training dataset from 10^4 to approximately 10^7. This iterative cycle of ML prediction and DFT validation enhanced both the accuracy and the generalizability of the models, leading to the discovery of new materials. Schmidt et al. adopted a larger-scale, stepwise transfer learning approach to progressively enhance the chemical and structural diversity of the training set. They explored a material design space with up to 1 billion compounds. These studies demonstrated the concept of continually refining ML prediction models to discover new materials. However, these studies are often based on existing structural prototypes, making it difficult to explore unknown material structures.

Optimization algorithms for accelerating materials search

Optimization algorithms, including heuristic algorithms such as genetic algorithms and particle swarm optimization, along with data-driven methods exemplified by Bayesian optimization, can effectively advance material structures toward enhanced stability[38] or performance[18,39,40], as illustrated in Fig. 2b. Given that single-objective optimizations of material performance can often lead to instability in new materials, a multi-objective optimization approach is necessary to simultaneously enhance stability and performance, aiming for a higher Pareto front[40,41]. The essentially increasing number of structures on the high-dimensional Pareto front also facilitates the selection of feasible solutions rather than the exhaustive search for all dominated or non-dominated solutions[41]. However, the main challenge of searching based on optimization algorithms stems from the large amounts of required material property evaluations, mainly based on DFT calculations, making this process resource intensive and inefficient.

ML material energy direct prediction models[42,43] and machine learning potentials (MLPs)[44–46] have been used to expedite crystal structure prediction (CSP) studies. Investigations by Cheng et al.[43] and Li et al.[42] combined ML energy prediction models, pre-trained on existing computational material databases, with optimization algorithms to accelerate CSP. MLPs can substantially enhance CSP by efficiently optimizing and evaluating structures, even for complex systems with hundreds of atoms[46], and can perform finite-temperature CSP by calculating the corresponding free energy[44,46]. However, training MLPs for CSP remains challenging, owing to the absence of prior crystal structure knowledge, leading to situations where high-energy structures in the CSP process exceed the accurate prediction range of MLPs. Hong et al. overcame this issue using the dynamical trajectories of liquid and quenched amorphous phases as the training set, which required no prior crystal structure information and could effectively sample diverse local orders that are potentially present in stable or metastable crystalline structures[47]. Over 180 unknown ternary metal oxide structures have been identified using this method, demonstrating its capability

of predicting multi-elemental compounds[45,48]. In addition, data-driven methods such as template-based CSP have emerged, which match similar chemical formulas to construct new materials, offering a promising route for the efficient discovery of new materials[49–51].

Choubisa et al. demonstrated the integration of ML property-prediction models with optimization algorithms in semiconductor design[52]. They trained ML models to predict bandgap and energy, which were used alongside evolutionary algorithms to search for targeted semiconductor development. This approach, which partially accomplished multi-objective optimization by combining bandgap and stability in a fitness function, successfully designed semiconductors with specific direct-bandgap characteristics. This method offers potential, particularly if further enhanced with real multi-objective optimization algorithms.

Physical rules-guided materials design

Leveraging physicochemical insights on structure–property relationships as a design rule to propose a candidate serves as a key traditional approach in semiconductor design[17,19], as shown in Fig. 2c. However, the discovery of such design rules is heavily dependent on the knowledge of the researcher, with a certain amount of chance, making this type of research inefficient. Data-driven methods assist in uncovering new material design rules that go beyond physical intuitions. Through regression or classification, ML models can rank descriptors based on their statistical correlation with target properties[53]. If the statistical correlation is sufficiently strong, then new materials aligned with these descriptors can be proposed as promising candidates. For example, Choubisa et al. extracted descriptors linked to bandgap characteristics from the data augmented by the evolutionary algorithm, and then the descriptors were successfully used to design new semiconductors with targeted ultraviolet/infrared light-emitting and direct-bandgap features[52]. In addition, symbolic regression techniques, represented by the sure independence screening and sparsifying operator (SISSO)[54], can be used to discover high-dimensional descriptors with enhanced accuracy. For instance, Bartel et al. utilized SISSO to identify a new descriptor for perovskite material synthesizability, moving beyond the traditional tolerance factor[55]. This method faces challenges in converting discovered rules into physical insights and leveraging these insights for high-performance material development. Symbolic regression, aimed at optimal feature combination, often produces complex formulas difficult to physically interpret, complicating the identification of high-performance materials owing to their non-reversible nature. An effective solution is conducting sensitivity analysis[56] on the formulas to isolate influential physical features. This process helps clarify understanding and guides the design of new, high-performance materials.

Generative model-driven material discovery

Generative models have emerged as a promising strategy for exploring the unknown materials design space[23,24], as shown in Fig. 2d. These models learn the intrinsic features of existing materials, acquire knowledge related to stable structures, and then sample and generate new materials based on learned patterns. Techniques including generative adversarial network (GAN)[57–59], variational autoencoders (VAE)[60,61], transformers[62] and diffusion models[61,63] have been used to construct generative models for materials. Intuitively, generative models require a material encoding method that is compliant with symmetry operations and is invertible. Although such effort has been made[60], the main representations implemented in current generative models are the lattice vectors, atomic numbers and site coordinates for encoding the unit cell shape, elemental species and positions of atoms, respectively[58,59,61,63]. This is considered the most natural and invertible method for encoding material structures, and symmetries contained in crystals are typically captured by special architectures in the models. In the physical-guided crystal generative model (PGCGM) developed by Zhao et al.[59], the inclusion of restrictions related to space group and atomic distances can substantially improve the generation performance of PGCGM, and 93.5% of generated materials can be optimized successfully by DFT. Equivariant graph neural networks (GNNs) were used in the crystal diffusion variational autoencoder (CDVAE) developed by Xie et al. to capture the crystal symmetries[61], and the model was found to be effective by multiple materials design works[64–66]. The recently developed diffusion model-based MatterGen adopted variance-preserving diffusion, generating diverse inorganic materials across the periodic table, with 75% below the 0.1 eV per atom threshold of the current convex hull[63]. Furthermore, this model can be fine-tuned by specific property constraints to generate materials with desired properties, such as a specific bandgap, representing an important forward step in the inverse design of materials. Despite the progress of model architectures, the training of generative models has not been sufficiently scaled up. The training set still lacks a variety of external conditions, such as temperature and pressure, hindering the model from generating stable structures under specific conditions. Continuously scaling up the model training, as was done with the predictive models in refs. 20,40, and moving toward universal material generation models is a foreseeable prospect.

These generative models have also been used to instruct multiple materials design research, including the design of spintronic materials[67] guided by the CubicGAN model[58], and the design of low-dimensional materials[64,65] and superconductors[66] guided by the CDVAE model. The main research paradigm involves training the generative model with initial stable materials of interest and using the constructed model to generate candidates for further DFT or surrogate ML model validation. The future development of generative model-derived material discovery should be based on a pre-trained stable general material generative model, which can be fine-tuned for different downstream operations, as practiced by Zeni et al.[63]. In addition, based on various computational material design methods, including material generation models and data-driven CSP, a large number of new materials have been proposed. These materials are currently being collected into hypothetical materials databases, such as the Carolina materials database, for more detailed theoretical or experimental validation in the future.

Optimizing performance by using ML's predictive capability

Evaluating the stability and performance of proposed candidate materials semiconductors is critical for new semiconductor design. Although DFT calculations offer precise results for ground-state charge density and wavefunctions, they must be connected to practical performance by hierarchical computational frameworks, as shown in Fig. 3. ML prediction models can be integrated with these frameworks by targeting different properties to enhance predictive efficiency, as shown in Fig. 3. According to the level of target property, data-driven strategies for semiconductor performance evaluation can be classified as directly predicting application-relevant properties or indirectly predicting intermediate variables in the physical models of these properties. The latter approach can effectively combine ML methods with physical models to improve the accuracy of property predictions and understanding of physical and material design principles. However, it typically demands more computational resources and faces challenges in generalization.

Representations for predicting semiconductor properties

Data representation is a foundational element in data-driven material property-prediction research, substantially affecting the performance and applicability of developed models. Predicting material properties from their structure fundamentally involves approximating quantum mechanical relationships that adhere to the symmetry constraints of the E(3) group, which includes rotation, translation and inversion in three-dimensional space. Scaler properties, such as energy, remain invariant under these transformations, whereas vector properties

https://doi.org/10.1038/s43588-024-00632-5

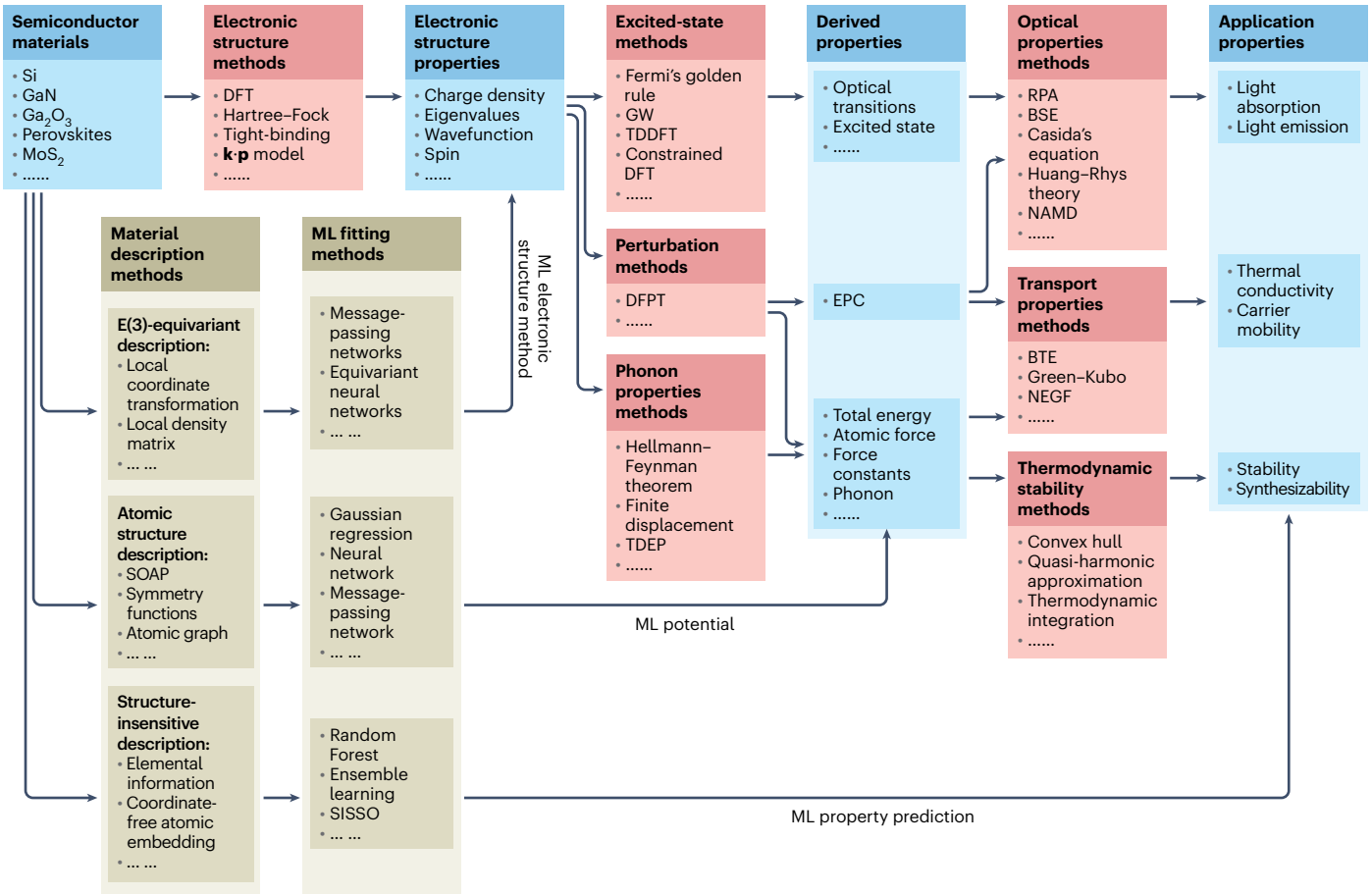

Fig. 3 | ML-assisted computational framework for semiconductor functionalities and electronic devices. A detailed representation of the ML-assisted hierarchical computational frameworks for stability, thermal conductivity, carrier conductivity and the optical properties of semiconductors for applications in electronic devices. The blue boxes depict semiconductor materials and their properties across different levels. The red boxes illustrate various theoretical and computational frameworks and the brown boxes highlight data-driven methods for predicting semiconductor material properties. Starting with semiconductor materials, fundamental electronic structure properties can be obtained through electronic structure calculation methods or ML prediction models. Subsequently derived properties can be calculated based on computational methods such as many-body perturbation theory based on Green's function (GW), time-dependent density-functional theory (TDDFT), density-functional perturbation theory (DFPT), and temperature dependent effective potentials (TDEP). Among the derived properties, ML predictions of total energy and atomic forces are well established. Properties determining device application can be obtained through further physical modeling using methods such as random phase approximation (RPA), Bethe–Salpeter equation (BSE), and non-equilibrium Green's function (NEGF). All these properties can be predicted using ML models trained on relevant datasets as well.

are invariant to translation but equivariant under rotation and inversion. Sole data-driven approaches are inapplicable for material design research owing to the complex relationship between raw material structure representation, depicted by spatial coordinates in an arbitrary Cartesian system, and material properties. Owing to the lack of symmetry invariance and length invariance, raw representations are difficult to process and learn from by sole data-driven approaches. While techniques such as training set augmentation can mitigate issues with invariance, they cannot substantially boost model performance. Therefore, developing effective representations for material structure and properties that align with fundamental physics to highlight the physical nature and unify the form of the dataset is crucial for leveraging data-driven methods in materials research.

Various symmetry-invariant material representations and ML frameworks have been developed to predict scaler properties, such as formation energy and bandgap. These methods include discrete descriptors[53], local structural representations[68–73] and graph representations[74–79]. Discrete feature descriptors can hardly achieve high-precision perdition[53]; however, they are easy to interpret and are suitable for searching physical rules, as discussed in 'Physical rules-guided materials design'. To fit the effect of subtle perturbations of the material structure on energy, many accurate local structure representation methods[68–73], represented by smooth overlap of atomic positions (SOAP)[70] and symmetry functions[69], have been developed, most of which can be described by a unified theoretical framework. Wavelet scattering transform is a unique local structure representation using a convolutional wavelet frame to represent the atomic density at different distances and orientations with symmetry invariance[72]. These methods can be combined with various ML methods, such as Gaussian process regression[80] and neural networks[69], to construct high-precision MLPs. Graph representations based on invariant features of material structure, such as atomic distances and bond angles, have been extensively used with message-passing GNNs to build predictive models for material properties[74–79]. Owing to its strong cross-element generalization capability, universal ML material property-prediction models are being established based on graph representations and message-passing models[77,79]. However, they are currently less spatially scalable than local representation-based models[73].

The introduction of tensor product-based equivariant neural networks (ENNs)[81] has substantially advanced the development of equivariant structure representations and property-prediction models, enhancing prediction performance for scalar properties[71,73,76,78]

and extending predictive capabilities to vectors and tensors. Equivariant local structure representations, such as λ-SOAP[71] and the newly developed Allegro[73], show high prediction precision and learning efficiency. Utilizing equivariant graph representations and ENNs has not only advanced energy prediction models, such as the development of neural equivariant interatomic potentials (NequIP)[76] and MACE[78], but also enabled the prediction of tensor properties, including electron Hamiltonians[82,83], and the development of effective material generation models[61]. As artificial intelligence for science gains more attention, equivariant material representations and property-prediction models are emerging as a forefront area within artificial intelligence research[84,85].

Beyond encoding material structure for predicting simple scalar and vector properties, effectively representing complex material properties, including one-dimensional spectral properties such as density of states, guided wave damage detection, and higher-dimensional outputs such as spectral images and electronic band structures, is essential for extracting pivotal information to enhance model performance. As an example, Rautela et al. transformed raw guided wave signals into wavelet-enhanced representations to train unsupervised ML models, achieving high-precision detection of delamination defects in composite materials[86]. Similarly, for studies that predict a high-order property from basic property inputs, such as predicting phonon scattering rates from phonon vectors[87] or predicting high-fidelity band structure from low-fidelity band structure[88], effective input property representations are crucial.

Optimizing thermal and electrical transport properties

The thermal management of electronic devices critically depends on the thermal conductivity of semiconductor materials. Data-driven methods have been extensively used to predict the thermal conductivity of inorganic materials through direct prediction[56,89–93] or indirectly by facilitating the calculations of force constants[94–96], scattering rates[87] and molecular dynamics simulations[97,98]. These methods can be used to enhance the thermal transport of semiconductor materials and further improve the thermal management of electronic devices. In the direct prediction of thermal conductivity, symbolic regression methods show high interpretability with strong extrapolation generalization performance[56,89,90]. Although these methods have substantially reduced computational costs compared with complete thermal conductivity calculations, they still rely on calculated properties, such as elastic modulus. In addition, some methods can predict thermal conductivity from crystal structures[91,92]. However, these methods have not been thoroughly tested for extrapolated performance, and they struggle with complex materials predictions. This limitation can be partially attributed to the scarcity of thermal conductivity data, which is both costly and challenging to obtain through computational and experimental methods. MLP-based indirect prediction methods can accelerate force-constant calculations, facilitate thermal conductivity evaluation based on the Boltzmann transport equation (BTE)[94–96] or accelerate thermal conductivity evaluation based on the Green–Kubo method by accelerating molecular dynamics simulations[97,98]. Guo et al. recently achieved the direct prediction of third- and fourth-order phonon scattering rates[87]. The researchers used phonon frequencies, wavevectors, eigenvectors and group velocities as descriptors of phonons, and a deep neural network was used to learn and predict the third- and fourth-order phonon scattering rates. However, the phonon eigenvector descriptor used in this work lacks scaling invariance with respect to the number of atoms in the unit cell, hindering the generalization of the model across different material systems. Given these factors, further development of dimension unified phonon eigenvector description method is an avenue worth exploring.

Data-driven approaches for electrical transport, especially through indirect approaches, are less common than for thermal transport due to the complex electron–phonon coupling (EPC) involved. Although direct predictions of carrier conductivity using basic descriptors has been attempted[99–101], they often and have limited accuracy and generalizability due to the limited dataset. Major intermediate properties required for carrier conductivity calculation include the electronic structure and EPC strength. The electrical transport of semiconductor heterojunctions has been accelerated by ML methods, by predicting their electronic structure[102]. However, the utilization of constant relaxation time approximation limits its accuracy. Recently, data-driven EPC prediction was successfully realized based on an ENN[103], paving the way for accurate data-driven prediction of semiconductor carrier mobility. Beyond the BTE method, the non-equilibrium Green's function (NEGF) method, combined with DFT or tight-binding models, can simulate ballistic electronic transport in nanostructures[104]. Bürkle et al. obtained electrical conductivity through NEGF based on local structural descriptors and a neural network[105], successfully predicting the electrical conductivity for large-scale nanowire systems with lengths up to 17.5 nm, and effectively extending the applicability range of the NEGF method.

ML-accelerated optical properties prediction

Data-driven methods, widely used in molecule excited-state property assessment[106,107], have not been frequently utilized in solid-state materials excited-state research owing to the high computational costs and complex electronic structures. Nevertheless, progress has been achieved in predicting accurate high-fidelity bandgaps[108–110] and electronic structures[88], as well as accelerating many-body quasiparticle and Bethe–Salpeter equation (BSE) calculations[111,112], which are all critical for obtaining accurate optical properties of semiconductors. For precise bandgap predictions, data from experiments or advanced computational techniques are required to prevent the typical underestimation of DFT. However, this requirement results in a scarcity of training data, limiting the performance of the resulting bandgap-prediction model[108]. Multi-fidelity models[113], such as Δ-learning[114], can bridge low-fidelity (DFT results) with high-fidelity data (obtained from hybrid functional, many-body quasiparticle calculations or experiments) and have shown promise in predicting the high-fidelity bandgaps of 2D semiconductors[110] and oxide perovskites[109]. Chen et al. developed a multi-fidelity GNN encompassing results from multiple computation methods with different fidelities and experimental data[113]. This approach can substantially enhance the predictive performance of high-fidelity bandgaps for both ordered and disordered materials[113]. High-fidelity electronic structures, such as G_0W_0 (ref. 88) quasiparticle band structures, have also been predicted using ML methods. However, these predictions cover only eigenvalues, and the lack of wavefunction information makes it impossible to obtain optical absorption. Recently, Dong et al. employed convolutional neural networks (CNNs) and established mapping between unscreened and screened Coulomb interactions, thus accelerating BSE computations[112]. The developed ML-BSE method was found to enhance the computational efficiency of finite-temperature optical absorption spectra by one to two orders of magnitude for a wide range of complex structures, including liquids, solids and nanostructures. Despite progress in the prediction of optical absorption properties, advancement in predicting luminescence properties is hindered by the complexities of quantifying excited-state renormalization in semiconductors. Further advancement relies on leveraging and enhancing existing excited-state computational techniques, such as occupancy constrained DFT and time-dependent DFT, to enable high-throughput calculations of semiconductor luminescence.

Recently, non-adiabatic molecular dynamics (NAMD) has been used to study non-radiative recombination in semiconductors[115,116]. Non-adiabatic couplings have a central role in NAMD calculations and require extensive calculations. The ML prediction of non-adiabatic couplings has been realized in multiple studies, which can considerably expedite research on non-radiative rates and carrier lifetimes in solid materials[116–118]. However, the classical path approximation used

in NAMD research neglects the impact of excited-state electrons on ion positions. This limitation makes it impractical to study localized excited states in solids where localized excited-state electrons can induce severe distortion of atomic positions.

Enhancing device performance through data-driven approaches

Although novel semiconductors offer promising intrinsic properties, material growth and device fabrication issues limit the manufacturing of high-performance electronics based on these materials. Gievn the high standards established by silicon-based technology, adapting these approaches to new semiconductors necessitates substantial revamping and investment. Therefore, computational and data-driven approaches are critical for accelerating such research.

Semiconductor growth through ML-assisted optimization

Semiconductor crystal growth involves complex multi-step processes, each with distinct thermodynamic and kinetic characteristics, making trial-and-error experimental explorations expensive and time intensive. Data-driven approaches offer assistance in both a theoretical understanding[119,120] of the growth processes and experimental optimization of growth parameters[121–127]. MLPs can accelerate and extend the scales of simulating material growth dynamical processes, from nucleation to growth, providing insights into the optimal conditions[119,120]. The main challenge involves constructing a high-quality training dataset that includes structures related to possible atomic configurations in growth processes. The development of high-throughput experiments and accumulation of experimental data have facilitated the data-driven experimental optimization of material growth. ML models trained on offline data collected from the literature or high-throughput experiments can explore uninvestigated experimental parameter spaces to identify parameters that may lead to high-performance materials[121,122,128–130], as shown in Fig. 4a. However, this open-loop approach requires extensive training data (at least hundreds of data points) and cannot guarantee a better result than the best sample in the training set. The iterative closed-loop optimization, as illustrated in Fig. 4b, can maximize the exploitation of each experiment, and therefore minimize the number of experiments required (typically less than 100) to achieve the optimizing target. The surrogate prediction model and experiment planning method are two key components of closed-loop optimization[131,132]. When real-time data and model uncertainty are available, an active learning strategy can function as an experiment planner to dynamically update the prediction model using an online training setting. This dynamic adjustment can also facilitate a balance between optimizing the model and exploiting the model to search for the optimal solution[131,133,134]. If no real-time data or model uncertainty are available, heuristic optimization algorithms can effectively serve as an experiment planning method, in conjunction with an offline-trained prediction model[135–137]. Reinforcement learning can break down decisions into isolated steps, and therefore aligns naturally with closed-loop optimization, particularly for multi-step material growth challenges. It needs an online training environment, and can simultaneously act as predictor and planner to find the optimal solutions while updating itself[124].

Promoting device fabrication via a data-driven approach

In addition to optimizing material growth quality, device fabrication optimization largely involves refining device interfaces. Although first-principles DFT calculations can be used to study interfacial properties, constructing low-energy stable interfacial structures remains a challenging task, due to the need to consider lattice mismatching and interfacial reconstruction. The limited simulation scale of DFT calculations also constrains their ability to simulate interfacial structures. Data-driven methods have shown potential in building realistic and stable interfacial structures. For instance, Li et al. introduced an ML-interface approach that enabled comprehensive solid–solid interface structure searches[138]. Using a combination of the phenomenological theory of martensite crystallography, graph theory and the stochastic surface walking method based on global neural network potentials, the researchers identified promising Si/SiO_2 interfacial structures suitable for 1 nm Si-based transistors. Experimentally, ML can optimize fabrication parameters for electronic device fabrication, leveraging data from various sources, including high-throughput experiments[139,140], literature mining[141] and technology computer-aided design simulations[142]. Using open-and closed-loop optimization paradigms, researchers have successfully used ML to predict materials and parameters for devices such as perovskite solar cells[140,141], field-effect transistors based on 2D materials[139] and integrated circuits[142]. As an example, Liu et al. realized the closed-loop optimization of open-air perovskite solar cell manufacturing using Bayesian optimization[140]. They achieved a success rate of 59%, which is much higher than the success rate (less than 20%) of traditional methods. In addition, the semiconductor manufacturing industry has made substantial progress in wafer defect identification and classification using ML techniques, resulting in enhanced production efficiency[143–145]. However, a comprehensive exploration of this subject is beyond the scope of this discussion.

Summary and outlook

This Perspective offers an in-depth exploration of the potential of data-driven frameworks within semiconductor research, shedding light on their pivotal role in accelerating material discovery, predicting semiconductor properties and optimizing device fabrication methodologies. In Supplementary Table 1, we summarize the representative studies from this Perspective, highlighting their main discovery, data-driven method, public dataset, and access link. Looking ahead, there are vast opportunities to further refine data-driven approaches for addressing computational challenges in semiconductor design, device optimization and manufacturing in the post-Moore era, paving the way for groundbreaking innovations and major advancements in semiconductor technology.

Discovering new semiconductor materials with performance surpassing existing ones in the extrapolative materials design space remains an exciting yet practically daunting task, given the challenging aspects involved. First, among functional solids, semiconductors face substantial challenges in materials design owing to the curse of dimensionality, primarily stemming from their diverse elemental composition and often complex mix of ionic and covalent bonds. In this regard, continuously expanding the training set through active learning or transfer learning to cover the complex material design space is a promising strategy[20,37]. In addition, extrapolative sampling and generating candidate materials in the latent space based on generative models offer another promising avenue to be explored[23,24]. Effectively integrating domain knowledge and physical principles is essential but can be challenging owing to complexity of the design space[59,140]. Second, designing high-performance semiconductors necessitates consideration of multiple properties simultaneously, such as a suitable bandgap, superior carrier mobility, controllable dopability and stability. This presents a challenging multiple-objective optimization problem[41], which calls for developing applicable data-driven methods to effectively balance trade-offs between conflicting properties and explore the high-dimensional design space of semiconductor materials, as mentioned earlier[146]. The methods' capability to incorporate physical knowledge and constraints into the optimization process, as well as to adaptively adjust the optimization strategy based on feedback[131,133], is crucial. Finally, transferability and uncertainty quantification are two critical factors for developing any effective ML-based predictive models for realistic semiconductor design. To address the transferability issue, it is crucial to use strategies that enhance the model's ability to generalize across different domains or datasets such as domain adaptation, meta-learning, data augmentation and ensemble techniques[147].

https://doi.org/10.1038/s43588-024-00632-5

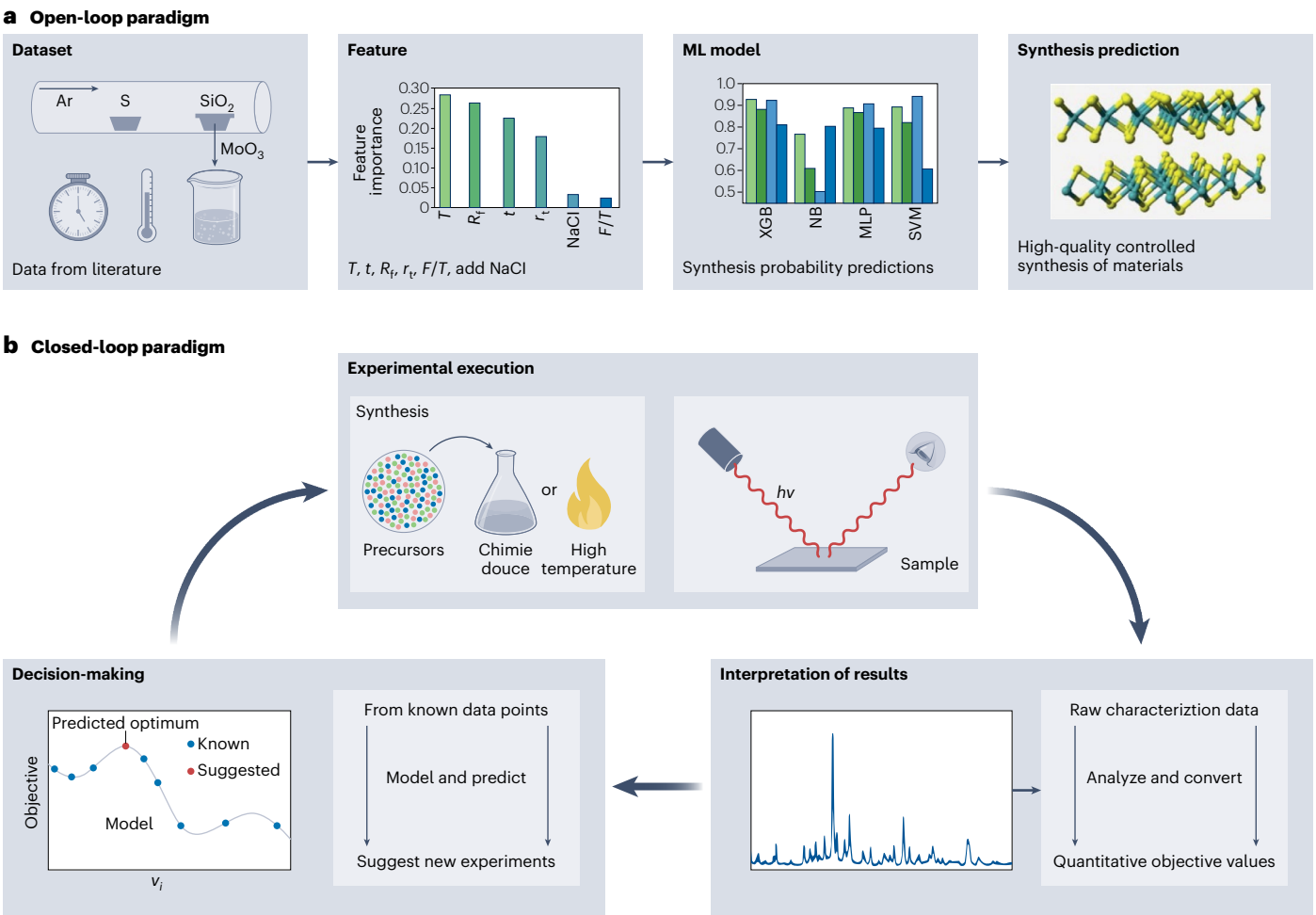

Fig. 4 | Two fundamental research paradigms in data-driven material growth experimental studies. a, Open-loop optimization characterized by directly correlating experimental parameters with the resulting material properties, to explore optimal material growth conditions. The four panels represent, from left to right: collecting data from the literature; finding important features from reaction temperature (T), reaction time (t), Ar gas flow rate (R_f), ramp time (r_t), boat configuration (F/T), where F and T stand for flat and tilted, and addition of NaCl; constructing ML synthesis probability prediction models, including XGBboost, support vector machine (SVM), Naïve Bayes (NB) and multilayer perceptron (MLP); and finding the synthesis conditions for high-quality materials. **b**, Closed-loop optimization characterized by utilizing the outcomes of each experiment to guide and optimize the planning of subsequent experiments. After experiment execution and results interpretation, the known data points (v_i) are passed to an intelligent decision-making algorithm that predicts data points with a high objective value and suggests new experiments. Panels reproduced with permission from: **a**, ref. 122, American Chemical Society; **b**, ref. 26, RSC.

Strategies such as ensemble techniques, dropout regularization and calibration methods can enable robust uncertainty quantification and enhance the reliability of ML-based predictive models[148].

Continuously optimizing semiconductor device performance through exploiting generative and optimization capabilities of data-driven strategies can be anticipated, but still faces challenges on the computational front. Developing effective data-driven approaches to guide semiconductor growth and synthesis, which usually involves complex parameters such as high-temperature, pressure, precursor gas composition and substrate surface preparation, is imperative yet challenging. This is critical not only for fabricating electronic devices, which demands large-size, low-defect-density single-crystal samples, but also for designing new semiconductors, which often involves predicting high-performance yet metastable materials with uncertain synthesizability[149]. This task necessitates integrating diverse sources of data, encompassing experimental success and failure data, computational data, theoretical models and domain expertise, to train predictive models capable of capturing the intricate relationships between synthesis parameters and the thermodynamic and kinetic stability of materials[26,121,122]. Along this direction, further exploration is warranted in the development of accurate and high-transferability MLPs[77,79], complemented by multiscale simulation methods (such as molecular dynamics, Monte Carlo, coarse-grained simulations)[126] and feedbacks from experimental data[134]. Designing the interfaces between solid–solid or solid–molecule with desired properties within the device architecture is a challenging task owing to factors such as the high computational cost and the curse of dimensionality stemming from the complexity of the system[138]. To achieve such goals, substantial efforts are required to develop effective energetics and properties calculation protocols, as well as robust data-driven design strategies.

Utilizing data-driven strategies to address the complex challenges in semiconductor manufacturing and testing, intensified by the reduced device size in the post-Moore era, represents a vital and promising area of exploration. In the many processes of semiconductor manufacturing, lithography stands out as one of the few critical non-chemical steps, noted for its precision and complexity. Incorporating ML techniques into lithography processes, specifically for generating optimal mask patterns during mask synthesis and optimization, and for detecting lithography hotspots (defects) in mask verification, can substantially enhance production yield and efficiency[144]. However, as lithography scales down, existing models and data risk obsolescence necessitating the use of methods such as transfer learning

to continually update current models and develop new approaches to push the boundaries of scaling down further. Testing is pivotal in semiconductor manufacturing, consuming approximately 70% of resources and time in very large-scale integration production[144]. Although design-for-test strategies are employed to reduce costs and enhance test coverage, the increasing integration of very large-scale integration introduces more complicated test parameters, thereby reducing test efficiency. ML-guided design-for-test (MLDFT) is addressing semiconductor testing challenges, by optimizing parameters of scan chain tests, refining test pattern generation for better test efficiency, and predicting test power to minimize device damage during testing[143,144]. As in lithography, MLDFT must evolve continually to support technological advances. The increase in circuit integration demands both the extraction of more features from intricate circuits for a thorough analysis and the application of feature engineering to remove redundancy. Utilizing deep learning models such as CNNs and GNNs is important for enhancing MLDFT capabilities. Adaptive learning schemes in MLDFT are currently underutilized but have the potential to closely integrate ML with real-time manufacturing, thereby substantially enhancing productivity.

References

1. Waldrop, M. M. The chips are down for Moore's law. *Nature* **530**, 144–147 (2016).
2. Powell, J. R. The quantum limit to Moore's law. *Proc. IEEE* **96**, 1247–1248 (2008).
3. Vorwerk, C., Sheng, N., Govoni, M., Huang, B. & Galli, G. Quantum embedding theories to simulate condensed systems on quantum computers. *Nat. Comput. Sci.* **2**, 424–432 (2022).
4. Schuman, C. D. et al. Opportunities for neuromorphic computing algorithms and applications. *Nat. Comput. Sci.* **2**, 10–19 (2022).
5. Akinwande, D. et al. Graphene and two-dimensional materials for silicon technology. *Nature* **573**, 507–518 (2019).
6. Li, W. et al. Approaching the quantum limit in two-dimensional semiconductor contacts. *Nature* **613**, 274–279 (2023).
7. Zhang, S. et al. Minimizing buried interfacial defects for efficient inverted perovskite solar cells. *Science* **380**, 404–409 (2023).
8. Luo, J. et al. Efficient and stable emission of warm-white light from lead-free halide double perovskites. *Nature* **563**, 541–545 (2018).
9. Yang, H. et al. Two-dimensional materials prospects for non-volatile spintronic memories. *Nature* **606**, 663–673 (2022).
10. Schram, T., Sutar, S., Radu, I. & Asselberghs, I. Challenges of wafer-scale integration of 2D semiconductors for high-performance transistor circuits. *Adv. Mater.* **34**, 2109796 (2022).
11. Wang, R. et al. A review of perovskites solar cell stability. *Adv. Funct. Mater.* **29**, 1808843 (2019).
12. Sabatier, P. C. Past and future of inverse problems. *J. Math. Phys.* **41**, 4082–4124 (2000).
13. Zunger, A. Inverse design in search of materials with target functionalities. *Nat. Rev. Chem.* **2**, 0121 (2018).
14. Peng, J. et al. Human- and machine-centred designs of molecules and materials for sustainability and decarbonization. *Nat. Rev. Mater.* **7**, 991–1009 (2022).
15. Yang, D. et al. Functionality-directed screening of Pb-free hybrid organic–inorganic perovskites with desired intrinsic photovoltaic functionalities. *Chem. Mater.* **29**, 524–538 (2017).
16. Zhao, X.-G. et al. Design of lead-free inorganic halide perovskites for solar cells via cation-transmutation. *J. Am. Chem. Soc.* **139**, 2630–2638 (2017).
17. Liu, Z. et al. Computational functionality-driven design of semiconductors for optoelectronic applications. *InfoMat* **2**, 879–904 (2020).
18. Liao, T. W. & Li, G. Metaheuristic-based inverse design of materials—a survey. *J. Materiomics* **6**, 414–430 (2020).
19. Peng, H. et al. Li-doped Cr_2MnO_4: a new p-type transparent conducting oxide by computational materials design. *Adv. Funct. Mater.* **23**, 5267–5276 (2013).
20. Schmidt, J. et al. Machine-learning-assisted determination of the global zero-temperature phase diagram of materials. *Adv. Mater.* **35**, 2210788 (2023).
21. Wang, H. et al. Scientific discovery in the age of artificial intelligence. *Nature* **620**, 47–60 (2023).
22. Kulik, H. J. et al. Roadmap on machine learning in electronic structure. *Electron. Struct.* **4**, 023004 (2022).
23. Yan, D., Smith, A. D. & Chen, C.-C. Structure prediction and materials design with generative neural networks. *Nat. Comput. Sci.* **3**, 572–574 (2023).
24. Anstine, D. M. & Isayev, O. Generative models as an emerging paradigm in the chemical sciences. *J. Am. Chem. Soc.* **145**, 8736–8750 (2023).
25. Szymanski, N. J. et al. An autonomous laboratory for the accelerated synthesis of novel materials. *Nature* **624**, 86–91 (2023).
26. Szymanski, N. J. et al. Toward autonomous design and synthesis of novel inorganic materials. *Mater. Horiz.* **8**, 2169–2198 (2021).
27. Luo, S., Li, T., Wang, X., Faizan, M. & Zhang, L. High-throughput computational materials screening and discovery of optoelectronic semiconductors. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **11**, e1489 (2021).
28. Curtarolo, S. et al. The high-throughput highway to computational materials design. *Nat. Mater.* **12**, 191–201 (2013).
29. Han, D. et al. Design of high-performance lead-free quaternary antiperovskites for photovoltaics via ion type inversion and anion ordering. *J. Am. Chem. Soc.* **143**, 12369–12379 (2021).
30. Li, Y. et al. Design of organic–inorganic hybrid heterostructured semiconductors via high-throughput materials screening for optoelectronic applications. *J. Am. Chem. Soc.* **144**, 16656–16666 (2022).
31. He, X., Singh, D. J., Boon-on, P., Lee, M.-W. & Zhang, L. Dielectric behavior as a screen in rational searches for electronic materials: metal pnictide sulfosalts. *J. Am. Chem. Soc.* **140**, 18058–18065 (2018).
32. Zhao, X.-G. et al. Cu–In halide perovskite solar absorbers. *J. Am. Chem. Soc.* **139**, 6718–6725 (2017).
33. Zhao, X.-G. et al. JAMIP: an artificial-intelligence aided data-driven infrastructure for computational materials informatics. *Sci. Bull.* **66**, 1973–1985 (2021).
34. Mounet, N. et al. Two-dimensional materials from high-throughput computational exfoliation of experimentally known compounds. *Nat. Nanotechnol.* **13**, 246–252 (2018).
35. Pizzi, G., Cepellotti, A., Sabatini, R., Marzari, N. & Kozinsky, B. AiiDA: automated interactive infrastructure and database for computational science. *Comput. Mater. Sci.* **111**, 218–230 (2016).
36. Gómez-Bombarelli, R. et al. Design of efficient molecular organic light-emitting diodes by a high-throughput virtual screening and experimental approach. *Nat. Mater.* **15**, 1120–1127 (2016).
37. Merchant, A. et al. Scaling deep learning for materials discovery. *Nature* **624**, 80–85 (2023).
38. Oganov, A. R., Pickard, C. J., Zhu, Q. & Needs, R. J. Structure prediction drives materials discovery. *Nat. Rev. Mater.* **4**, 331–348 (2019).
39. Zhang, L., Luo, J.-W., Saraiva, A., Koiller, B. & Zunger, A. Genetic design of enhanced valley splitting towards a spin qubit in silicon. *Nat. Commun.* **4**, 2396 (2013).
40. Zhang, Y.-Y., Gao, W., Chen, S., Xiang, H. & Gong, X.-G. Inverse design of materials by multi-objective differential evolution. *Comput. Mater. Sci.* **98**, 51–55 (2015).
41. Allahyari, Z. & Oganov, A. R. in *Handbook of Materials Modeling* (eds Andreoni, W. & Yip, S.) 2777–2790 (Springer, 2020).

42. Li, C.-N., Liang, H.-P., Zhang, X., Lin, Z. & Wei, S.-H. Graph deep learning accelerated efficient crystal structure search and feature extraction. *npj Comput. Mater.* **9**, 176 (2023).
43. Cheng, G., Gong, X.-G. & Yin, W.-J. Crystal structure prediction by combining graph network and optimization algorithm. *Nat. Commun.* **13**, 1492 (2022).
44. Kruglov, I. A. et al. Crystal structure prediction at finite temperatures. *npj Comput. Mater.* **9**, 197 (2023).
45. Kang, S. et al. Accelerated identification of equilibrium structures of multicomponent inorganic crystals using machine learning potentials. *npj Comput. Mater.* **8**, 108 (2022).
46. Wang, X. et al. Data-driven prediction of complex crystal structures of dense lithium. *Nat. Commun.* **14**, 2924 (2023).
47. Hong, C. et al. Training machine-learning potentials for crystal structure prediction using disordered structures. *Phys. Rev. B* **102**, 224104 (2020).
48. Hwang, S. et al. Stability and equilibrium structures of unknown ternary metal oxides explored by machine-learned potentials. *J. Am. Chem. Soc.* **145**, 19378–19386 (2023).
49. Wei, L. et al. TCSP: a template-based crystal structure prediction algorithm for materials discovery. *Inorg. Chem.* **61**, 8431–8439 (2022).
50. Kusaba, M., Liu, C. & Yoshida, R. Crystal structure prediction with machine learning-based element substitution. *Comput. Mater. Sci.* **211**, 111496 (2022).
51. Wang, H.-C., Botti, S. & Marques, M. A. L. Predicting stable crystalline compounds using chemical similarity. *npj Comput. Mater.* **7**, 12 (2021).
52. Choubisa, H. et al. Interpretable discovery of semiconductors with machine learning. *npj Comput. Mater.* **9**, 117 (2023).
53. Zhao, R., Xing, B., Mu, H., Fu, Y. & Zhang, L. Evaluation of performance of machine learning methods in mining structure–property data of halide perovskite materials. *Chin. Phys. B* **31**, 056302 (2022).
54. Ouyang, R., Curtarolo, S., Ahmetcik, E., Scheffler, M. & Ghiringhelli, L. M. SISSO: a compressed-sensing method for identifying the best low-dimensional descriptor in an immensity of offered candidates. *Phys. Rev. Mater.* **2**, 083802 (2018).
55. Bartel, C. J. et al. New tolerance factor to predict the stability of perovskite oxides and halides. *Sci. Adv.* **5**, eaav0693 (2019).
56. Purcell, T. A. R., Scheffler, M., Ghiringhelli, L. M. & Carbogno, C. Accelerating materials-space exploration for thermal insulators by mapping materials properties via artificial intelligence. *npj Comput. Mater.* **9**, 112 (2023).
57. Kim, S., Noh, J., Gu, G. H., Aspuru-Guzik, A. & Jung, Y. Generative adversarial networks for crystal structure prediction. *ACS Cent. Sci.* **6**, 1412–1420 (2020).
58. Zhao, Y. et al. High-throughput discovery of novel cubic crystal materials using deep generative neural networks. *Adv. Sci.* **8**, 2100566 (2021).
59. Zhao, Y. et al. Physics guided deep learning for generative design of crystal materials with symmetry constraints. *npj Comput. Mater.* **9**, 38 (2023).
60. Ren, Z. et al. An invertible crystallographic representation for general inverse design of inorganic crystals with targeted properties. *Matter* **5**, 314–335 (2022).
61. Xie, T., Fu, X., Ganea, O.-E., Barzilay, R. & Jaakkola, T. S. Crystal diffusion variational autoencoder for periodic material generation. In *International Conference on Learning Representations* (2022).
62. Fu, N. et al. Material transformers: deep learning language models for generative materials design. *Mach. Learn. Sci. Technol.* **4**, 015001 (2023).
63. Zeni, C. et al. MatterGen: a generative model for inorganic materials design. Preprint at https://arxiv.org/abs/2312.03687 (2023).
64. Lyngby, P. & Thygesen, K. S. Data-driven discovery of 2D materials by deep generative models. *npj Comput. Mater.* **8**, 232 (2022).
65. Moustafa, H., Lyngby, P. M., Mortensen, J. J., Thygesen, K. S. & Jacobsen, K. W. Hundreds of new, stable, one-dimensional materials from a generative machine learning model. *Phys. Rev. Mater.* **7**, 014007 (2023).
66. Wines, D., Xie, T. & Choudhary, K. Inverse design of next-generation superconductors using data-driven deep generative models. *J. Phys. Chem. Lett.* **14**, 6630–6638 (2023).
67. Siriwardane, E. M. D., Zhao, Y. & Hu, J. Data-driven deep generative design of stable spintronic materials. *CrystEngComm* **25**, 6017–6029 (2023).
68. Langer, M. F., Goeßmann, A. & Rupp, M. Representations of molecules and materials for interpolation of quantum-mechanical simulations via machine learning. *npj Comput. Mater.* **8**, 41 (2022).
69. Behler, J. Atom-centered symmetry functions for constructing high-dimensional neural network potentials. *J. Chem. Phys.* **134**, 074106 (2011).
70. Bartók, A. P., Kondor, R. & Csányi, G. On representing chemical environments. *Phys. Rev. B* **87**, 184115 (2013).
71. Grisafi, A., Wilkins, D. M., Csányi, G. & Ceriotti, M. Symmetry-adapted machine learning for tensorial properties of atomistic systems. *Phys. Rev. Lett.* **120**, 036002 (2018).
72. Eickenberg, M., Exarchakis, G., Hirn, M. & Mallat, S. Solid harmonic wavelet scattering: predicting quantum molecular energy from invariant descriptors of 3D electronic densities. In *Advances in Neural Information Processing Systems* Vol. 30, 6543–6552 (Curran Associates, 2017).
73. Musaelian, A. et al. Learning local equivariant representations for large-scale atomistic dynamics. *Nat. Commun.* **14**, 579 (2023).
74. Xie, T. & Grossman, J. C. Crystal graph convolutional neural networks for an accurate and interpretable prediction of material properties. *Phys. Rev. Lett.* **120**, 145301 (2018).
75. Chen, C., Ye, W., Zuo, Y., Zheng, C. & Ong, S. P. Graph networks as a universal machine learning framework for molecules and crystals. *Chem. Mater.* **31**, 3564–3572 (2019).
76. Batzner, S. et al. E(3)-equivariant graph neural networks for data-efficient and accurate interatomic potentials. *Nat. Commun.* **13**, 2453 (2022).
77. Chen, C. & Ong, S. P. A universal graph deep learning interatomic potential for the periodic table. *Nat. Comput. Sci.* **2**, 718–728 (2022).
78. Batatia, I., Kovacs, D. P., Simm, G., Ortner, C. & Csanyi, G. MACE: higher order equivariant message passing neural networks for fast and accurate force fields. *Adv. Neural Inf. Process. Syst.* **35**, 11423–11436 (2022).
79. Choudhary, K. et al. Unified graph neural network force-field for the periodic table: solid state applications. *Digit. Discov.* **2**, 346–355 (2023).
80. Deringer, V. L. et al. Gaussian process regression for materials and molecules. *Chem. Rev.* **121**, 10073–10141 (2021).
81. Thomas, N. et al. Tensor field networks: rotation- and translation-equivariant neural networks for 3D point clouds. Preprint at https://arxiv.org/abs/1802.08219 (2018).
82. Gong, X. et al. General framework for E(3)-equivariant neural network representation of density functional theory Hamiltonian. *Nat. Commun.* **14**, 2848 (2023).
83. Zhong, Y., Yu, H., Su, M., Gong, X. & Xiang, H. Transferable equivariant graph neural networks for the Hamiltonians of molecules and solids. *npj Comput. Mater.* **9**, 182 (2023).
84. Zhang, X. et al. Artificial intelligence for science in quantum, atomistic, and continuum systems. Preprint at https://arxiv.org/abs/2307.08423 (2023).

85. Liu, S. et al. Symmetry-informed geometric representation for molecules, proteins, and crystalline materials. *Adv. Neural Inf. Process. Syst.* **36**, 66084–66101 (2023).
86. Rautela, M., Senthilnath, J., Monaco, E. & Gopalakrishnan, S. Delamination prediction in composite panels using unsupervised-feature learning methods with wavelet-enhanced guided wave representations. *Compos. Struct.* **291**, 115579 (2022).
87. Guo, Z. et al. Fast and accurate machine learning prediction of phonon scattering rates and lattice thermal conductivity. *npj Comput. Mater.* **9**, 95 (2023).
88. Knøsgaard, N. R. & Thygesen, K. S. Representing individual electronic states for machine learning GW band structures of 2D materials. *Nat. Commun.* **13**, 468 (2022).
89. Loftis, C., Yuan, K., Zhao, Y., Hu, M. & Hu, J. Lattice thermal conductivity prediction using symbolic regression and machine learning. *J. Phys. Chem. A* **125**, 435–450 (2021).
90. Juneja, R., Yumnam, G., Satsangi, S. & Singh, A. K. Coupling the high-throughput property map to machine learning for predicting lattice thermal conductivity. *Chem. Mater.* **31**, 5145–5151 (2019).
91. Jaafreh, R., Kang, Y. S. & Hamad, K. Lattice thermal conductivity: an accelerated discovery guided by machine learning. *ACS Appl. Mater. Interfaces* **13**, 57204–57213 (2021).
92. Zhu, T. et al. Charting lattice thermal conductivity for inorganic crystals and discovering rare earth chalcogenides for thermoelectrics. *Energy Environ. Sci.* **14**, 3559–3566 (2021).
93. Luo, Y., Li, M., Yuan, H., Liu, H. & Fang, Y. Predicting lattice thermal conductivity via machine learning: a mini review. *npj Comput. Mater.* **9**, 4 (2023).
94. Ladygin, V. V., Korotaev, P. Y., Yanilkin, A. V. & Shapeev, A. V. Lattice dynamics simulation using machine learning interatomic potentials. *Comput. Mater. Sci.* **172**, 109333 (2020).
95. Li, R. et al. A deep neural network interatomic potential for studying thermal conductivity of β-Ga_2O_3. *Appl. Phys. Lett.* **117**, 152102 (2020).
96. Babaei, H., Guo, R., Hashemi, A. & Lee, S. Machine-learning-based interatomic potential for phonon transport in perfect crystalline Si and crystalline Si with vacancies. *Phys. Rev. Mater.* **3**, 074603 (2019).
97. Qian, X., Peng, S., Li, X., Wei, Y. & Yang, R. Thermal conductivity modeling using machine learning potentials: application to crystalline and amorphous silicon. *Mater. Today Phys.* **10**, 100140 (2019).
98. Verdi, C., Karsai, F., Liu, P., Jinnouchi, R. & Kresse, G. Thermal transport and phase transitions of zirconia by on-the-fly machine-learned interatomic potentials. *npj Comput. Mater.* **7**, 156 (2021).
99. Mukherjee, M., Satsangi, S. & Singh, A. K. A statistical approach for the rapid prediction of electron relaxation time using elemental representatives. *Chem. Mater.* **32**, 6507–6514 (2020).
100. Na, G. S., Jang, S. & Chang, H. Predicting thermoelectric properties from chemical formula with explicitly identifying dopant effects. *npj Comput. Mater.* **7**, 106 (2021).
101. Antunes, L. M., Butler, K. T. & Grau-Crespo, R. Predicting thermoelectric transport properties from composition with attention-based deep learning. *Mach. Learn. Sci. Technol.* **4**, 015037 (2023).
102. Pimachev, A. K. & Neogi, S. First-principles prediction of electronic transport in fabricated semiconductor heterostructures via physics-aware machine learning. *npj Comput. Mater.* **7**, 93 (2021).
103. Li, H. et al. Deep-learning density functional perturbation theory. *Phys. Rev. Lett.* **132**, 096401 (2024).
104. Ozaki, T., Nishio, K. & Kino, H. Efficient implementation of the nonequilibrium Green function method for electronic transport calculations. *Phys. Rev. B* **81**, 035116 (2010).
105. Bürkle, M. et al. Deep-learning approach to first-principles transport simulations. *Phys. Rev. Lett.* **126**, 177701 (2021).
106. Dral, P. O. & Barbatti, M. Molecular excited states through a machine learning lens. *Nat. Rev. Chem.* **5**, 388–405 (2021).
107. Westermayr, J. & Marquetand, P. Machine learning for electronically excited states of molecules. *Chem. Rev.* **121**, 9873–9926 (2021).
108. Zhuo, Y., Mansouri Tehrani, A. & Brgoch, J. Predicting the band gaps of inorganic solids by machine learning. *J. Phys. Chem. Lett.* **9**, 1668–1673 (2018).
109. Li, W. et al. Predicting band gaps and band-edge positions of oxide perovskites using density functional theory and machine learning. *Phys. Rev. B* **106**, 155156 (2022).
110. Zhang, L. et al. Accurate band gap prediction based on an interpretable Δ-machine learning. *Mater. Today Commun.* **33**, 104630 (2022).
111. Zauchner, M. G., Horsfield, A. & Lischner, J. Accelerating GW calculations through machine-learned dielectric matrices. *npj Comput. Mater.* **9**, 184 (2023).
112. Dong, S. S., Govoni, M. & Galli, G. Machine learning dielectric screening for the simulation of excited state properties of molecules and materials. *Chem. Sci.* **12**, 4970–4980 (2021).
113. Chen, C., Zuo, Y., Ye, W., Li, X. & Ong, S. P. Learning properties of ordered and disordered materials from multi-fidelity data. *Nat. Comput. Sci.* **1**, 46–53 (2021).
114. Ramakrishnan, R., Dral, P. O., Rupp, M. & von Lilienfeld, O. A. Big data meets quantum chemistry approximations: the Δ-machine learning approach. *J. Chem. Theory Comput.* **11**, 2087–2096 (2015).
115. Wang, S. et al. Effective lifetime of non-equilibrium carriers in semiconductors from non-adiabatic molecular dynamics simulations. *Nat. Comput. Sci.* **2**, 486–493 (2022).
116. Zhang, Z., Wang, J., Zhang, Y., Xu, J. & Long, R. Charge recombination dynamics in a metal halide perovskite simulated by nonadiabatic molecular dynamics combined with machine learning. *J. Phys. Chem. Lett.* **13**, 10734–10740 (2022).
117. Akimov, A. V. Extending the time scales of nonadiabatic molecular dynamics via machine learning in the time domain. *J. Phys. Chem. Lett.* **12**, 12119–12128 (2021).
118. Wang, B., Chu, W., Tkatchenko, A. & Prezhdo, O. V. Interpolating nonadiabatic molecular dynamics hamiltonian with artificial neural networks. *J. Phys. Chem. Lett.* **12**, 6070–6077 (2021).
119. Niu, H., Bonati, L., Piaggi, P. M. & Parrinello, M. Ab initio phase diagram and nucleation of gallium. *Nat. Commun.* **11**, 2654 (2020).
120. Zhang, D., Yi, P., Lai, X., Peng, L. & Li, H. Active machine learning model for the dynamic simulation and growth mechanisms of carbon on metal surface. *Nat. Commun.* **15**, 344 (2024).
121. Xu, M. et al. Machine learning driven synthesis of few-layered WTe_2 with geometrical control. *J. Am. Chem. Soc.* **143**, 18103–18113 (2021).
122. Lu, M. et al. Machine learning-assisted synthesis of two-dimensional materials. *ACS Appl. Mater. Interfaces* **15**, 1871–1878 (2023).
123. Zhang, J., Wang, F., Shenoy, V. B., Tang, M. & Lou, J. Towards controlled synthesis of 2D crystals by chemical vapor deposition (CVD). *Mater. Today* **40**, 132–139 (2020).
124. Volk, A. A. et al. AlphaFlow: autonomous discovery and optimization of multi-step chemistry using a self-driven fluidic lab guided by reinforcement learning. *Nat. Commun.* **14**, 1403 (2023).
125. Beckham, J. L. et al. Machine learning guided synthesis of flash graphene. *Adv. Mater.* **34**, 2106506 (2022).
126. Bhowmik, S. & Govind Rajan, A. Chemical vapor deposition of 2D materials: a review of modeling, simulation, and machine learning studies. *iScience* **25**, 103832 (2022).

127. Tao, H. et al. Nanoparticle synthesis assisted by machine learning. *Nat. Rev. Mater.* **6**, 701–716 (2021).
128. Costine, A., Delsa, P., Li, T., Reinke, P. & Balachandran, P. V. Data-driven assessment of chemical vapor deposition grown MoS_2 monolayer thin films. *J. Appl. Phys.* **128**, 235303 (2020).
129. Dahl, J. C., Wang, X., Huang, X., Chan, E. M. & Alivisatos, A. P. Elucidating the weakly reversible Cs–Pb–Br perovskite nanocrystal reaction network with high-throughput maps and transformations. *J. Am. Chem. Soc.* **142**, 11915–11926 (2020).
130. Han, Y. et al. Machine-learning-driven synthesis of carbon dots with enhanced quantum yields. *ACS Nano* **14**, 14761–14768 (2020).
131. Balachandran, P. V., Xue, D., Theiler, J., Hogden, J. & Lookman, T. Adaptive strategies for materials design using uncertainties. *Sci. Rep.* **6**, 19660 (2016).
132. Schmidt, J., Marques, M. R. G., Botti, S. & Marques, M. A. L. Recent advances and applications of machine learning in solid-state materials science. *npj Comput. Mater.* **5**, 83 (2019).
133. Lookman, T., Balachandran, P. V., Xue, D. & Yuan, R. Active learning in materials science with emphasis on adaptive sampling using uncertainties for targeted design. *npj Comput. Mater.* **5**, 21 (2019).
134. Ohkubo, I. et al. Realization of closed-loop optimization of epitaxial titanium nitride thin-film growth via machine learning. *Mater. Today Phys.* **16**, 100296 (2021).
135. Salley, D. et al. A nanomaterials discovery robot for the Darwinian evolution of shape programmable gold nanoparticles. *Nat. Commun.* **11**, 2771 (2020).
136. Liow, C. H. et al. Machine learning assisted synthesis of lithium-ion batteries cathode materials. *Nano Energy* **98**, 107214 (2022).
137. Li, J. et al. Autonomous discovery of optically active chiral inorganic perovskite nanocrystals through an intelligent cloud lab. *Nat. Commun.* **11**, 2046 (2020).
138. Li, Y.-F. & Liu, Z.-P. Smallest stable Si/SiO_2 interface that suppresses quantum tunneling from machine-learning-based global search. *Phys. Rev. Lett.* **128**, 226102 (2022).
139. Chen, X. et al. Wafer-scale functional circuits based on two dimensional semiconductors with fabrication optimized by machine learning. *Nat. Commun.* **12**, 5953 (2021).
140. Liu, Z. et al. Machine learning with knowledge constraints for process optimization of open-air perovskite solar cell manufacturing. *Joule* **6**, 834–849 (2022).
141. Liu, W. et al. Machine learning enables intelligent screening of interface materials towards minimizing voltage losses for p–i–n type perovskite solar cells. *J. Energy Chem.* **83**, 128–137 (2023).
142. Yoo, J. et al. Machine-learning based TCAD optimization method for next generation BCD process development. In *2021 33rd International Symposium on Power Semiconductor Devices and ICs (ISPSD)* 279–282 (IEEE, 2021).
143. Zhao, Z. & Cai, Z. A review of intelligent design for test based on machine learning. In *2023 International Symposium of Electronics Design Automation (ISEDA)* 116–120 (IEEE, 2023).
144. Amuru, D. et al. AI/ML algorithms and applications in VLSI design and technology. *Integration* **93**, 102048 (2023).
145. Kim, T. & Behdinan, K. Advances in machine learning and deep learning applications towards wafer map defect recognition and classification: a review. *J. Intell. Manuf.* **34**, 3215–3247 (2023).
146. Gopakumar, A. M., Balachandran, P. V., Xue, D., Gubernatis, J. E. & Lookman, T. Multi-objective optimization for materials discovery via adaptive design. *Sci. Rep.* **8**, 3738 (2018).
147. Tan, C. et al. A survey on deep transfer learning. In *Artificial Neural Networks and Machine Learning—ICANN 2018* (eds Kůrková, V. et al.) 270–279 (Springer, 2018).
148. Gawlikowski, J. et al. A survey of uncertainty in deep neural networks. *Artif. Intell. Rev.* **56**, 1513–1589 (2023).
149. Tolborg, K., Klarbring, J., M. Ganose, A. & Walsh, A. Free energy predictions for crystal stability and synthesisability. *Digit. Discov.* **1**, 586–595 (2022).
150. Dotson, J. J. et al. Data-driven multi-objective optimization tactics for catalytic asymmetric reactions using bisphosphine ligands. *J. Am. Chem. Soc.* **145**, 110–121 (2023).

Acknowledgements

L.Z. acknowledges funding support from the National Key Research and Development Program of China (grant number 2022YFA1402500) and the National Natural Science Foundation of China (grant numbers 62125402 and 62321166653).

Author contributions

L.Z. led the preparation, writing and editing of this Perspective. J.X. contributed most of the text and figures. Y.Z. and Z.L. assisted in writing and figure preparation. X.W. and M.F. reviewed and refined the paper. All authors contributed to discussions and feedback.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s43588-024-00632-5.

Correspondence should be addressed to Xinjiang Wang or Lijun Zhang.

Peer review information *Nature Computational Science* thanks Senthilnath Jayavelu and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Primary Handling Editor: Kaitlin McCardle, in collaboration with the *Nature Computational Science* team.