# A perspective on inverse design of battery interphases using multi-scale modelling, experiments and generative deep learning

Bhowmik, Arghya; Castelli, Ivano Eligio; Garcia-Lastra, Juan Maria; Bjørn-Jørgensen, Peter; Winther, Ole; Vegge, Tejs

*Published in:*
Energy Storage Materials

*DOI:*
10.1016/j.ensm.2019.06.011

*Publication date:*
2019

*Document version*
Publisher's PDF, also known as Version of record

Energy Storage Materials 21 (2019) 446–456

Contents lists available at ScienceDirect

Energy Storage Materials

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

# A perspective on inverse design of battery interphases using multi-scale modelling, experiments and generative deep learning

Arghya Bhowmik [a], Ivano E. Castelli [a], Juan Maria Garcia-Lastra [a], Peter Bjørn Jørgensen [b], Ole Winther [b,c], Tejs Vegge [a,*]

[a] Department of Energy Conversion and Storage, Technical University of Denmark, DK-2800, Kgs. Lyngby, Denmark
[b] DTU Compute, Technical University of Denmark, DK-2800, Kgs. Lyngby, Denmark
[c] Department of Biology, University of Copenhagen, DK-2200, Copenhagen, Denmark

ARTICLE INFO

*Keywords:*
Battery interphases
Multi-scale modelling
Generative deep learning
Inverse materials design

ABSTRACT

Understanding and controlling the complex and dynamic processes at battery interfaces holds the key to developing more durable and ultra high performance secondary batteries. Interfacial processes like dendrite and Solid Electrolyte Interphase (SEI) formation span numerous time- and length scales, and despite decades of research, their formation, composition,structure and function still pose a conundrum. Consequently, "inverse design" of high-performance interfaces and interphases like the SEI, remains an elusive dream. Here, we present a perspective and possible blueprint for a future battery research strategy to reach this ambitious goal. Semi-supervised generative deep learning models trained on all sources of available data, i.e., extensive multi-fidelity datasets from multi-scale computer simulations and databases, operando characterization from large-scale research facilities, high-throughput synthesis and laboratory testing, need to work closely together to unlock this dream. We show how understanding and tracking different types of uncertainties in the experimental and simulation methods, as well as the machine learning framework for the generative model, is crucial for controlling and improving the fidelity in the predictive design of battery interfaces and interphases. We argue that simultaneous utilization of data from multiple domains, including data from failed experiments, will play a critical role in accelerating the development of reliable generative models to enable accelerated discovery and inverse design of durable ultra high performance batteries based on novel materials, structures and designs.

## 1. Introduction

Rechargeable (secondary) batteries play a critical role in the transition towards clean energy and e-mobility. The performance and durability of existing batteries are, however, limited; not only by the battery materials themselves but equally by our lack of understanding of the complex and dynamic processes taking place inside the materials and in particular at the interfaces within the battery cell. Despite decades of research, the formation of, e.g., dendrites and the composition and function of interfacial compounds like the Solid Electrolyte Interphase (SEI) and Cathode Electrolyte Interphase (CEI) remain a conundrum. The properties of these interfacial compounds depend in a highly complex and elusive manner on the specific characteristics of the composition and structures of the electrolyte and the electrode materials, as well as the external variables that decide the time evolution of the system [1–3].

The European Large Scale Research Initiative "Battery 2030+" has recently identified establishing the "Battery Interface Genome (BIG)" and a "Materials Acceleration Platform (MAP)" as essential milestones towards accelerated discovery of ultra high performance batteries (see http://www.battery2030.eu). In this perspective, we provide our vision of a methodological blueprint for enabling inverse design of battery interfaces in a "BIG-MAP" infrastructure based on a consolidated treatment of multiple and heterogeneous data sources, AI (Artificial Intelligence) orchestrated data acquisition from experiments and simulations, compressed representation of interfacial states at different time- and length scales, and semi-supervised generative deep learning. We present and discuss the complexity of this massive challenge, the state-of-the-art of selected critical methods and approaches - ranging from experimental techniques to physical simulation models and machine learning methods - and finally, how a closed-loop scheme for accelerated (inverse) interfacial design can be established, building upon the accumulated insights from the individual building blocks.

* Corresponding author.
*E-mail address:* teve@dtu.dk (T. Vegge).

https://doi.org/10.1016/j.ensm.2019.06.011
Received 15 February 2019; Received in revised form 7 June 2019; Accepted 7 June 2019
Available online 18 June 2019

While batteries consist of a multitude of different interfaces and interphases, few are as complex and pose as many challenges to the design of ultra high performance batteries as the SEI/CEI interphase at the electrode-electrolyte interface. In this perspective, special emphasis will therefore be placed on these interphases, but the described methodology will be equally applicable to other battery interfaces and interphases.

The complexity of the SEI and CEI arises from a host of different reactions and processes spanning a wide range of time- and length scales that define their formation, structure and, ultimately, their functionality in the battery. The initial steps in the formation of the SEI start at the atomic level, with its nucleation through chemical reactions between the electrode surface and molecules from the electrolyte. These reactions normally involve fast transfer of electrons at time-scales down to femtoseconds. During the growth of the interphase, different compounds are formed and the properties of the grains/phases and the boundaries between them, which extend from the nanometer to the micron scale, are crucial to allow ionic migration and to block electron conductivity. At the macroscopic scale, the mechanical stability of the interphase is responsible for the battery aging on the scale of years [4]. All these factors and their intricate entanglement lead to a combinatorial explosion that makes the rational (inverse) design of optimal interphases a daunting task.

Proactive and dynamic control of the SEI/CEI formation and complex composition and morphology remains a "Grand Challenge" in battery research and development - a challenge in striking need of more advanced and versatile prediction and design techniques. Radically new approaches are needed to accelerate the discovery and development of ultra high performing and durable materials and interphases in rechargeable batteries. The purpose of this perspective is to outline a possible path for bringing this ambitious task within reach through the use of AI-orchestrated closed-loop optimization and a common data infrastructure. Hybrid physical and machine learning models trained on large curated datasets from advanced multi-scale computational modelling, materials databases, operando characterization to data from synthesis and materials/cell level testing provide the basis for this infrastructure (Fig. 1).

The traditional paradigm of trial-and-error based sequential materials optimization is to start from a known materials composition and structure, and rely on human intuition to guide the optimization of the materials to improve the performance. Inverse materials design effectively inverts this process by allowing the desired performance goals to define the composition and structure which best fulfills these targets without a priori defining the starting material or structure [5,6].

State-of-the-art in battery interphase design is unidirectional - going from structure to the properties. Both the structure and the properties are measured or simulated for specific interphases to start with, owing to an inherent lack of one-to-one mapping between input structure/chemistry and system performance. On the contrary, inverse design, therefore, relies heavily on the use of generative models. Optimization challenges as complex as proactive engineering and design of a battery interphase, requires going from a many-to-one to a one-to-many type mapping (e.g., going from the desired SEI/CEI to specifying the optimal ratio of the numerous components in the electrolyte compositions [7]) over the specific dimensionality and termination of the electrode to the specific electrochemical and/or thermal preparation. To implement such models in battery interphase design, we need to incorporate the relevant physical understanding, and the model should be capable of performing an inverse mapping from the desired properties to the original composition of the materials and external parameters/conditions. Deep learning models represent an efficient way to optimize the data flow and build the required bridges between different domains, helping to solve the biggest challenges of battery interphases. In this perspective, we discuss the potential and main challenges facing such procedures for accelerated inverse design of battery interphases and outline our view towards future research direction.

## 2. Experimental and simulation methods for studying reactions and processes at battery interfaces

In this section, we discuss some of the key processes occurring at a battery interface and the existing modelling and experimental techniques, which can provide the necessary data and complement the generative design framework proposed in this perspective. A battery is a highly complex device where many (electro-)chemical reactions take place at different interfaces and stages of the battery lifetime. In typical Li-ion batteries (LIB), (electro-)chemical reactions are responsible for the formation of the SEI layer on the graphitic anode (negative electrode) and the CEI on the cathode (positive electrode), involving the electrolyte and its additives during the first cycles of the battery. Once the layers are formed, Li penetrates through the SEI/CEI films, which should be an ionic conductor and, at the same time, electronic insulating to prevent for further degradation of the electrolyte. The composition of the electrolyte, together with additives and impurities, directly influences the (electro-) chemical reactions and thus the performances, durability, and safety of the battery [8]. Electronic transport is one of the determining mechanisms in the formation of the SEI/CEI, critically influencing the battery performance. Especially critical is the electronic transport in electrodes where strongly insulating materials are involved. Some examples are metal-sulfur batteries and aprotic metal-air batteries. In the former, the poor conductivity of sulfur determines how much of this element can be loaded in the sulfur-carbon composites at the cathode [9]. In the latter,

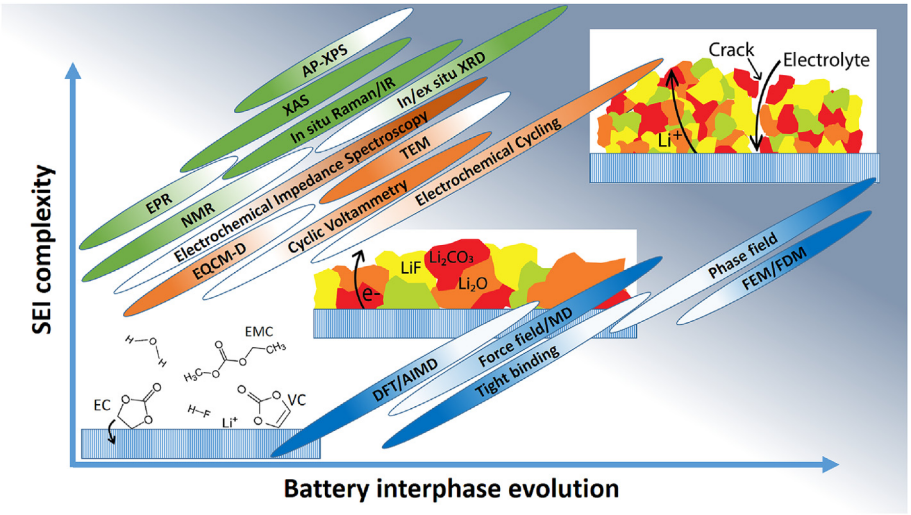


**Fig. 1.** Starting from the initial cell assembly/manufacturing, the battery interphase/SEI evolves continuously while adding complexity - from the molecular assembly, to SEI deposition and eventual macro-scale structure formation like cracks. This evolution is determined by the initial chemical, material and structural details, as well as the environmental variables during the lifetime of the battery, e.g. temperature and C-rates. The state of the interphase can be probed by various electrochemical (brown) and spectroscopic/structural (green) characterization techniques, as well as simulations (blue). The fidelity of a specific method in probing the interphase changes as the SEI matures. The variation from high (low) fidelity to low (high) fidelity of a given method due to the change in SEI complexity is denoted by the transition from solid color (higher fidelity) to a lighter tone (lower fidelity). (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

the peroxides formed as discharge product at the cathode passivates the carbon support and eventually result in the sudden death of the battery [10]. The insulating nature of both sulfur and peroxides force charges to move through mechanisms such as electron tunneling and polaron hopping. An accurate modelling of these two mechanisms is essential to determine the critical (limiting) thicknesses of these materials at the cathodes. The conventional method to study polaron hopping, is based on transition state theory combined with density functional theory (DFT), but this approach severely overestimates polaronic conductivities (e.g., by 5 and 10 orders of magnitude in sulfur [9] and peroxides, respectively), making it necessary to employ more expensive computational methods based on the Landau-Zenner (LZ) formalism combined with constrained DFT (cDFT) to get accurate results [11]. However, there is clearly room to accelerate the LZ-cDFT calculations by applying ML-based algorithms.

Aprotic metal-air batteries are another example of the formation of a complex interphase at the electrode-electrolyte interface consisting of several compounds in which the (grain) boundaries can display improved electronic conductivity respect to their parent materials. Under particular circumstances, like carbon dioxide contamination or use of carbon-based electrolytes, a mixture of lithium peroxide and lithium carbonate is formed as a discharge product at the air-cathode. At the boundary between these two materials electrons can diffuse, which it is not possible in their respective bulk phases [10]. This particular case illustrates the level of complexity of an interphase in an emerging battery chemistry, where the performance is not only governed by the intrinsic properties of the materials but also by their boundaries.

Observing and simulating phenomenon like these requires dealing with a range of time- and length scales from both the experimental and modelling point of view, where different stages in the formation and growth of the interphase can be analyzed, e.g. through a combination of complementary experimental techniques targeting specific scales (Fig. 1). The formation of the SEI/CEI during the first cycles can be studied with electrochemical characterization methods like cyclic-voltammetry (CV), XPS, and in-situ spectroscopies (NMR, Raman, IR) as sketched in Fig. 2. The structural evolution can be investigated through in-situ XRD [12], the ionic transport by quasi-elastic neutron scattering (QENS) [13], or Galvanostatic/Potentiostatic Intermittent Titration Technique (GITT/PITT) [14]. The aging of the battery can be tracked through cycling tests, and eventually the battery can be analyzed post-mortem using Transmission electron microscopy (TEM) and by spectroscopic techniques, as mentioned above. Safety of the battery can also be tested using abuse testing [15]. With respect to the computational modelling, the length scales span from the atomistic level, where quantities as the open circuit voltage can be determined, to the device level, which is employed in battery management systems (BMS). Although in most cases, specific design tools have been developed to treat each of them independently [16–18], recently significant progress has been made in designing multi-scale models to cover multiple length- and time-scales [19–21].

Interfacial phenomena can be treated directly in short (time-/length) scale simulations, while long scale simulations usually neglect them or rely on averaged description. At the atomistic level (i.e., simulations involving no more than a few hundreds of atoms), first principle calculations, typically based on DFT, can be used to evaluate the electrolyte stability and predict if they will eventually decompose to form the SEI [22]. DFT calculations can also be used to assess the mechanical stability of interfaces and the charge transport (both ionic and electronic) across and along interfaces. A step further is to combine the DFT methods with molecular dynamic (MD) simulations (ab-initio molecular dynamics, AIMD) [23–26], to obtain a time-resolved analysis of the SEI formation. Another possibility is to use parameters derived from DFT calculations as an input for second-principle calculations, i.e., to build for force fields (FF) for conventional MD [27] or ReaxFF models [28], to obtain Slater-Koster integrals for Tight-Binding DFT (TBDFT) calculations [29] or probability rates in Kinetic Monte Carlo (KMC) simulations [30]. This allows studying model interphases using thousands of atoms in the simulations. Machine learning is now playing an important role to support and automatize the construction of parameters for these second principle simulations from first principle methods [31–36]. Examples of this span from Cluster Expansion methods [37], which are widely employed to study the disorder in electrode materials, to neural networks to systematically improve the reliability of MD simulations [38].

When moving from the microscopic modelling level to the macroscopic one, the atomic details about the interface/interphase are typically neglected, which is an important reason for the challenges in linking simulation scales. Macroscopic modelling is often based on finite element methods (FEM) to solve the partial differential equations used to describe the diffusion of species in the electrodes and the electrolyte (ignoring the interfaces). Examples of this are the pseudo two-dimensional model (P2D) [39], which can be used to capture Li migration and has been combined with KMC to investigate film growth mechanism [19], and the single particle model (SPM) [40], which can be employed to assess the battery aging.

Finally, empirical models, based on fitting of experimental data, are used to monitor the macroscopic behavior of a battery by the BMS.

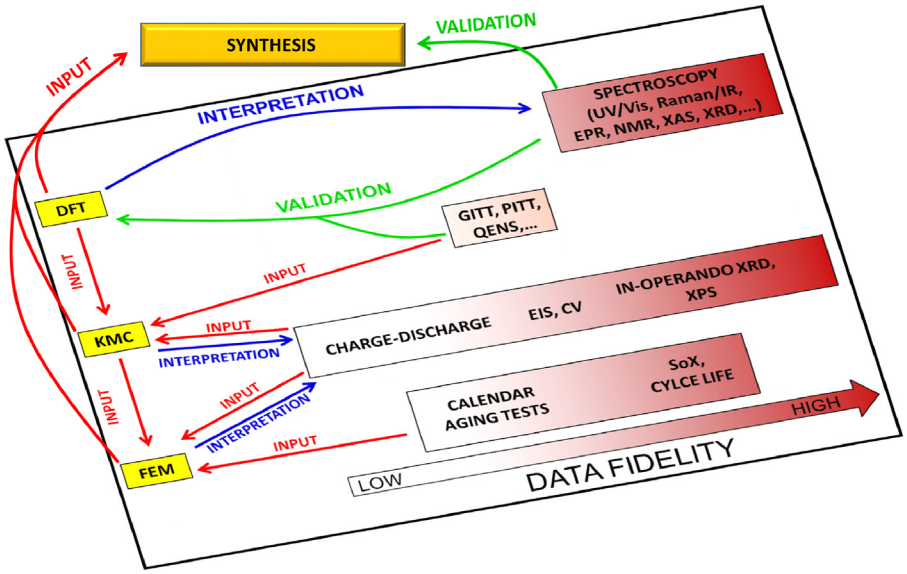


**Fig. 2.** The complex relation between the different state-of-the-art experimental and theoretical methods for studying battery interphases and their fidelity. Experimental techniques to characterize and test batteries are used to validate the theoretical modelling at different scales, while modelling activities help to interpret experiments. In addition, information from experiments and atomic-scale modelling is used as input in meso- and macroscopic modelling. Synthesis of materials is guided by input from theory and validated by experiments. The fidelity of each method is typically inversely proportional to its cost, although the fidelity/cost relationship can be optimized depending on the time frame and the experimental conditions (fidelity refers to the degree to which a model, simulation or experiment reproduces the state and behaviour of a real world object, feature or condition).

Traditional ML models generally fail in extrapolating beyond the parameter space, where adequate training data is available and thus have a limited predictive capability, when extrapolated (far) beyond the normal operating conditions. Such models are still used in practice, e.g. to predict state of charge, state of health, etc. (SoX) [41] and improve battery pack performance and lifetime [42–45]. Under the newly developed concept of physics informed [46] uncertainty-based deep learning, uncertainty propagation/estimation starting from training data as well as within the ML model itself can be used to obtain guidance on the reliability of the predictions [47–51].

Under the right conditions, i.e., appropriate models and datasets, ML techniques can accelerate the coupling between the different simulation scales, incorporating interfacial effects in mesoscopic and macroscopic simulations, which will enhance the prediction capabilities of the latter. For example, new ML methods like compressed sensing and autoencoder can help identify and automatically extract [52–54] necessary interface respresentations and "decriptors", e.g. thermodynamic, structural or kinetic parameters, from the expensive, short length/time scale physics simulations. Given the spatio-temporal complexity of the evolution of battery interfaces, the lack of simple intuitive physical descriptors can be overcome with a compressed sensing based ML framework. This can identify/select the physical descriptors for the limiting physical phenomena for battery interphases like charge and ionic transport, redox potentials, interfacial stress, etc. This approach has the inherent advantage of providing physical insights over black box ML modelling. ML driven physical intuition generation can help to develop an interface descriptor library for advancing battery interface design. It will enable long time/length scale coarse grained physics or empirical surrogate model based simulations [55,56] that helps to establish board trends in performance in a similar fashion as seen in, e.g., catalysis [57–59]. Additionally, ML can be used to make fast surrogate models for estimating properties of the SEI [60] (the most challenging region for atomistic simulations), to be used alongside macro scale models, thus improving reliability of multi-scale scale models.

A cardinal requirement for reliable ML or physics based models is the accuracy and comprehensiveness (quality) of the data it is mapping between, i.e., the structural/chemical details of the interphase and the observed properties. Cross validation of data from multiple sources generally improves the quality. Fig. 2 represents how data that can be gathered from experimental and simulation methods can be coupled in a complex and non-sequential manner, and that information often flow both ways between methods. In a reliable multi-scale approach to design battery materials, inclusion of interfacial details is necessary to feed the larger scale models with appropriate parameters. Ideally, the parameters for mesoscopic and macroscopic simulations should be derived from a set of relevant experiments combined with the output of microscopic scale calculations (for instance a KMC calculation can use DFT results and GITT data as input parameters). At the same time, experiments validate the results from calculations (for instance an XRD spectrum can validate a structure predicted by DFT), and calculations help to clarify experiments which could have several interpretations (for instance DFT can elucidate the origin of a peak in a Raman spectrum).

In this respect, special care should be taken as models typically assume a set of ideal conditions that are very difficult to achieve experimentally. In general, efforts should be made to bring models closer to the experimental conditions and also to perform experiments focusing and decoupling different effects (e.g., working with surfaces and electrolytes as free of impurities or defects as possible; in particular during the initial training of the model). An example of this is the formation of the SEI films on carbon- and silicon-anodes [61,62]. Solid bridges between different levels of experimental data and (multi-scale) computational models are thus required in order to make an intelligent choice of materials, additives, etc., which should be selected to form the desired interphase.

Deep learning models are particularly well suited to optimize the data flow and build such bridges, helping to solve the biggest challenges of battery interphases. Ideally, the ML algorithms should be fed with high-fidelity simulations and high-fidelity experiments (see definition of fidelity in the caption of Fig. 2). However, high-fidelity data is often very expensive and/or time-consuming to obtain and consequently scarce. A central element in the schemes for accelerated inverse interphase design discussed below, is for the generative model to know which experimental or computational technique can provide the most relevant/pertinent data at the lowest cost (in terms of time and/or money) for optimizing the model. It is important to stress that at certain times, a large, low-cost data set with moderate-to-low fidelity can be more useful than a limited, high-cost/high-fidelity dataset, e.g., in highly sparse regions of parameter space, and vice versa.

## 3. Representative modelling challenges at battery interphases

To illustrate the complexity of interfacial phenomena and processes in batteries and the concomitant difficulties to model and rationalize them, we present few representative challenges, i.e., the growth of dendrites in metallic-anodes, the SEI formation in model electrode surfaces (carbon- and silicon-based anodes), and electronic transport through the interphase. The later is schematized in Fig. 1, in which different phenomena are indicated as a function of the battery interphase evolution and SEI complexity with various modelling, structural and chemical characterization methods that can be used to elucidate and rationalize them.

The formation of dendrites upon cycling of a battery is one of the most vivid examples of an interfacial process, where several time- and length scales are involved. Preventing the formation of dendrites at the negative electrode would mean tremendous progress in several battery technologies, by enabling the use of high energy density metal electrodes. Thus, there is a enormous interest in understanding their formation and pursuing strategies to inhibit their realization. At the atomic level, it is possible to find dopants which block the nucleation sites for the dendrites. This strategy has allowed elucidation of the beneficial impact of additives as Bi and In in Zn anodes for secondary zinc-air batteries [63].

At larger length scales, MD calculations are used to look at the charge distribution in the cracks present at the interface between metallic lithium and solid electrolytes and how that influences the growth rate and the shape of dendrites [64]. With phase-field models, we can model grain-level localized anisotropic misfit strain during battery cycling [65]. FEM is employed to determine the impact of the specific geometry of the electrode and the operating conditions (i.e., current density and temperature) on the dendritic growth [66]. It has been used to design a method to reduce the impact of dendrites in copper electrodes [67].

The initial SEI is formed during the first charge/discharge cycles and its passivating nature prevents further decomposition of the electrolyte. The SEI should, however, still allow transport of Li-ions from the electrolyte to the anode electrode and vice versa, relying on different mechanisms [68]. Various compounds are present in the SEI coming from different decomposition mechanisms: inorganic compounds like LiF is obtained from decomposition of HF and water impurities [61] and organic compounds like $Li_2CO_3$ from the decomposition of the organic electrolyte [69,70], etc. Multiple atomic- and multi-scale models addressing different aspects of the SEI formation have been proposed [71,72] in which experimental data are often used as input.

Electrochemical response experiments on single crystal model surfaces together with state-of-the-art AIMD DFT simulations [61] and in-/ex-situ X-ray experiments coupled with first principle calculations [62] - are two recent examples of how a combination of simulations and experiments can provide insight into the formation of the SEI on carbon-based and Si-anodes. Such studies cover the formation of the SEI layer during the first few cycles, but also provide possible design criteria for improved anode/SEI interphases. To go beyond that, continuum models have been proposed to investigate the long-term growth mechanism of the SEI [2,73,74].

On the cathode side, descriptors like adsorption energy, the acid dissociation constant or the oxygen band center can be used to rationalize the reactivity of the cathode for the electrolyte oxidation reactions by

correlating simulations with mass spectrometry experiments [22,75,76].

While the state-of-the-art has already progressed in multi-modal data gathering, we envision that a future "BIG-MAP" will integrate all experimental and simulation based data sources and databases under a shared infrastructure with seamless data sharing between all methods as well as collective AI-orchestrated control and utilization of them. This constitutes a paradigm shift concerning the current status for generation and collection of data. Nowadays, data is typically generated using single experiments or simulations, which encompasses a well defined length- and time-scale. Databases, on the other hand, are mostly defined to contain data coming from similar experimental techniques or computational models. To reach the point of inverse design of battery interphases, this has to change.

Workflows need to be created to cover higher (preferably complete) levels of connectivity among databases gathered from different experiments and simulations, and the underlying interfacial details they cover. It will allow the AI to recognize which type of data is most critical/valuable at any given time, and launch the appropriate experiment or simulation. The database structure should also allow the interconnections between the various sources, and tools that bridge among various databases are thus a requirement. Progress is currently underway to create shared data infrastructure within computational materials area focusing on reproducibility, automation and transparency [77]. Our aspiration is a shared data infrastructure that is centered around the structural and dynamic properties of the interphase and encompasses all the relevant experimental/computational techniques.

## 4. Deep generative models for inverse design

The principles of battery interphase design are currently being explored independently through experimental characterization, life cycle and degradation tests, as well as through simulations at different length scales. Such results allow for the creation of either physical models or observation based rules. Given sufficiently detailed information about an interphase (or interface) and an accurate physics-based model, all relevant properties can, in theory, be predicted. Nonetheless, the lack of detailed and accurate information about the composition, structure and dynamic processes at the interphase, as well as the inability to perform quantum mechanical simulations at the required time- and length scale makes this approach impractical. Recent developments and deployment of ML-based tools that improve with training have created a paradigm shift in many scientific fields [78–87]. The availability of large datasets from high throughput simulations and large experimental facilities have made battery research ripe for benefiting from ML orchestrated modelling and design principle development [88,89].

The comprehensive term of AI can be broadly defined as "the study of agents that receive precepts from the environment and perform actions", following Russel and Norvig [90]. Machine learning is a subfield of AI, where the functional mapping from input to output learns from examples and improves upon such training. For example, a ML model can learn from long-term charge/discharge cycling experiments of a battery system, e.g., at varying C-rates and temperatures, and become very good at predicting the online state of charge (SoC) and estimating the state of health (SoH) for a specific type of battery (e.g. in a BMS), or predict and classify cells by cycle life [41,91–93]. Such training data can be either provided by human guided experiments or can be computer generated. If the ML model can also orchestrate the experiments, simulations or tests to be performed to further improve the model, it is defined as *active learning*. We define a *generative model* as a probabilistic model for the observed data.

In supervised learning, the objective is to learn a mapping from a feature vector $x$ (an n-dimensional vector of numerical features that represent some object) to a target variable $y$. In unsupervised learning, we only have access to a data set of feature vectors $x$ and the goal is to model the data distribution $p(x)$. A supervised generative model represents the joint probability distribution $p(x,y)$, which can be factorized as $p(x, y) = p(y|x)p(x)$. Many real world applications operate in a semi-supervised setting, where only some of the feature vectors have a corresponding target variable, but the unlabeled data points are still useful for learning the distribution of all features of interest. Recent developments in neural network based ML models (popularized as *deep learning* [94]), employs many layers of interconnected neurons. The structure of the model provides immense flexibility in the functional space and is thus well suited for approximating complex physical interactions at battery interphases. Here, a plethora of physical and chemical processes occur at multiple time and length scales, and the mapping from feature vector $x$ to target variable $p(y|x)$ and the distribution of features $p(x)$ are highly complicated functions and powerful models such as deep neural networks [95,96] are needed to learn these functions from the available data.

Generative deep learning is the family of models that combine the generative modelling approach with deep neural networks. One of the main reasons for employing deep generative models is that it enables the simulation (prediction) of new data points. The new data points can be used for qualitative model evaluation by assessing whether the generated data looks realistic. Furthermore, the generated data is of practical use for inverse design, where the goal is to generate data with the desired property, e.g. a high performance interphase. In this section, we discuss how emergent ML tools can be used to connect and utilize heterogeneous data from various experimental techniques and physics-based simulation methods to create a holistic closed-loop design framework as shown in Fig. 3 for predictive and generative battery interphase design. The proposed framework would enable inverse design of battery interphases such that the specific performance metrics are achieved, while retaining a (reasonable) degree of control over how the interphase evolves over battery lifetime. *Inverse design* [97] of battery materials and interphases, i.e., optimizing and engineering materials and interphases from the starting point of a particular desired functionality without a priori defining a specific composition, structures, etc. [98], has until now been computationally unfeasible for most battery challenges due to the massive complexity. However, the complexity of the design space makes inverse design based on generative models particularly enticing.

The apparent lack of a one-to-one mapping between input variables such as the chemical/structural starting point, the lifetime environment and the observables, makes the functional relationship between the design details and the performance and evolution "non-invertible", necessitating a generative inverse design approach. Furthermore, the inverse design process must be able to handle incomplete specifications, where not all properties and observables are specified at any given time, and needs to be configured such that the overall performance is optimized. For inverse molecular design, where a one-to-one mapping is not possible and the design space is immensely large, traditional screening strategies do not work and evolutionary and generative models are often used [98,99]. The complexity of the inverse design framework extends onwards for interphases, since the time evolution of the system must also be taken into account.

Computational materials design has made enormous progress as outlined in the recent 2019 roadmap [100], but successful examples of inverse computational design are still rare. A central element in inverse materials design is the ability to identify and suggest material compositions and structures, which can, in fact, be synthesized. Recent work by Aykol et al. uses a network representation to identify synthesizable materials [101] from the OQMD database [102,103]. The future challenge lies in extending current inverse design strategies to be able to handle battery interphases and systems of similar complexity.

The full integration of physical, experimental and data-derived phenomenological models towards closed-loop battery interphase design is the long term goal. It would allow greater flexibility towards optimal utilization of all available experimental/simulation techniques based on the reliability and cost, and accelerate the materials discovery process dramatically. For example, given a specific performance-limiting physical process, one can choose between time consuming but

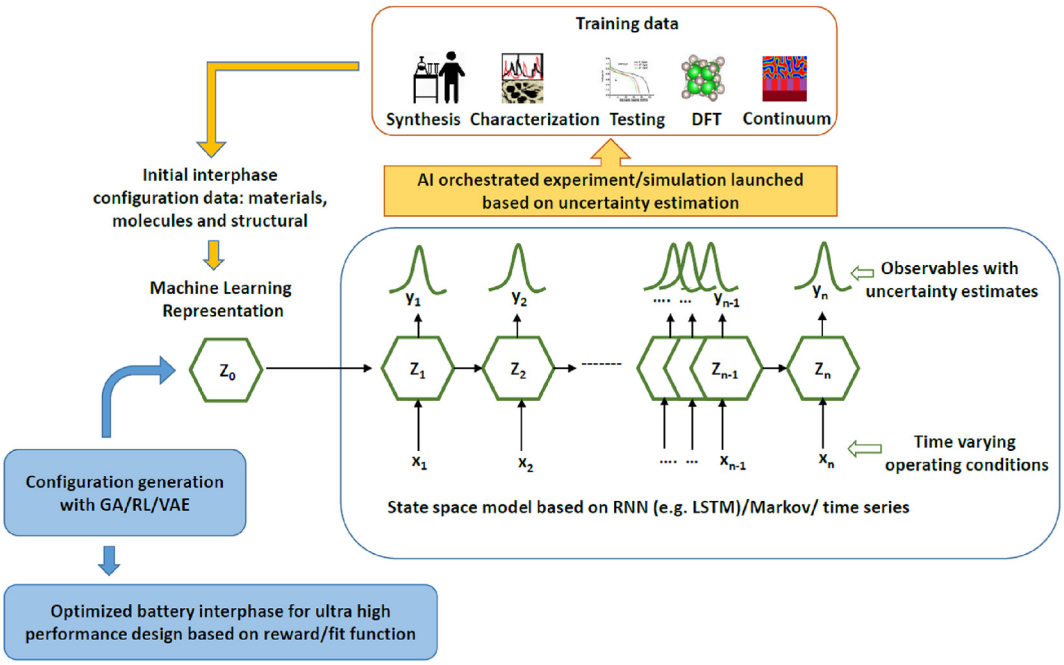


**Fig. 3.** A framework for semi-supervised deep generative model with ML-optimized active learning from experiments and simulations. The battery interphase is modeled using a time series model and based on the predictions of the model, an agent decides whether to launch a new experiment/simulation, which then becomes part of the training data. Given sufficient data, different techniques can be used to explore the input representation space to find new battery interphases that outperforms the ones in the training data.

comparatively cheap cycling experiments, synchrotron based XRD investigation, expensive atomic scale characterization, a handful of high accuracy DFT simulations, many low accuracy force field simulations or utilize ML model fitted to existing data. Based on the specific system and physical process(es), each of these methods has an inherent level of uncertainty. Intelligent selection is thus needed to obtain the optimal direction to follow based on the estimated fidelity.

ML based surrogate models can reduce the complexity pertaining to many simultaneous physical/chemical phenomenon by coalescing effects from multiple physical phenomena into learned functions. However, the vast number of chemical, environmental and structural details that affect the battery interphase performance makes it unfeasible to provide even semi-accurate prediction for all possible configurations. Existing experimental data and that accessible through future experimental efforts cover only a fraction of the combinations possible. This obstructs using state-of-the-art supervised ML models as the majority of widely used modelling frameworks require the availability of reliable data for different input variable combinations. The reliability of such models is highly dependent on the quality and quantity of training data [104,105]. As experiments/simulations are merely an end to data gathering in this framework, it is important recognize that data from failed experiments can also be utilized to train ML models [106].

The fidelity of the predictions from such models is limited if the underlying physical mechanism is different for the training and the test data. For similar reason, predictive ML models trained with state-of-the-art simulation and experimental results can not directly support evaluation of exploratory new battery interphases. For instance, an ML model trained on data from a battery interphase where the charge transport [107] is polaronic, will have limited success in predicting properties of interphases where electron tunneling [71] is the dominant mechanism for charge transport.

A battery interphase changes throughout its lifetime depending on its configuration and the operating conditions it is exposed to. The initial battery state is represented through configuration parameters like the materials and molecules present, as well as structural details like microstructure of the electrode composite, etc. How to represent this data in an ML model is discussed in more detail in section 4.1, and here it is discussed how to model the interphase. One possibility is to model the state of the interphase as a time series with a latent vector representation state $z_t$ at time $t$. The latent space representation evolves as a Markov chain [108], i.e., the state at $z_{t+1}$ depends only on the current state $z_t$ and the operating conditions $x_t$. The actual (electro-)chemical reactions taking place between species present are a function of time dependent variables like operating current rate and temperature etc. Chemical species are thus formed as a function of those present at the previous state, as well as the environmental variables. The state transitions can be modeled with a recurrent neural network (RNN) (e.g., Long Short Term Memory (LSTM) cells [109]) and include stochastic layers e.g. Ref. [110], to model the inherent uncertainty of the process. At any time step, the model can be queried to map the latent state to a set of observable outputs $y_t$. By defining a cost function that evaluates the cost of performing different simulation/experiments together with an acquisition function that decides the added value of performing a given experiment, an active learning scheme can be used to decide to optimal simulation/experiment to explore next (Fig. 3).

Given the performance metrics provided by the model, it is possible to envisage inverse design, i.e., determine the required input configuration and external parameters to yield the desired output, given the output. One way to achieve this is to do optimization in the input representation space. Utilizing back-propagation of gradients through the state space model, it is possible to search for a configuration and set of conditions that yield the desired performance metrics using gradient descent or other guided search methods such as genetic algorithms (GA) or reinforcement learning (RL) (Fig. 4) [111–114]. Alternatively, a direct mapping from response variable to configurations can be learned using a neural network [99] or invertible neural networks [115,116].

### 4.1. Representation of interphases for machine learning

ML models are limited by training data in two different ways. The sheer complexity of systems like battery interphases leads to large variability in the methods and parameters for which experimental and

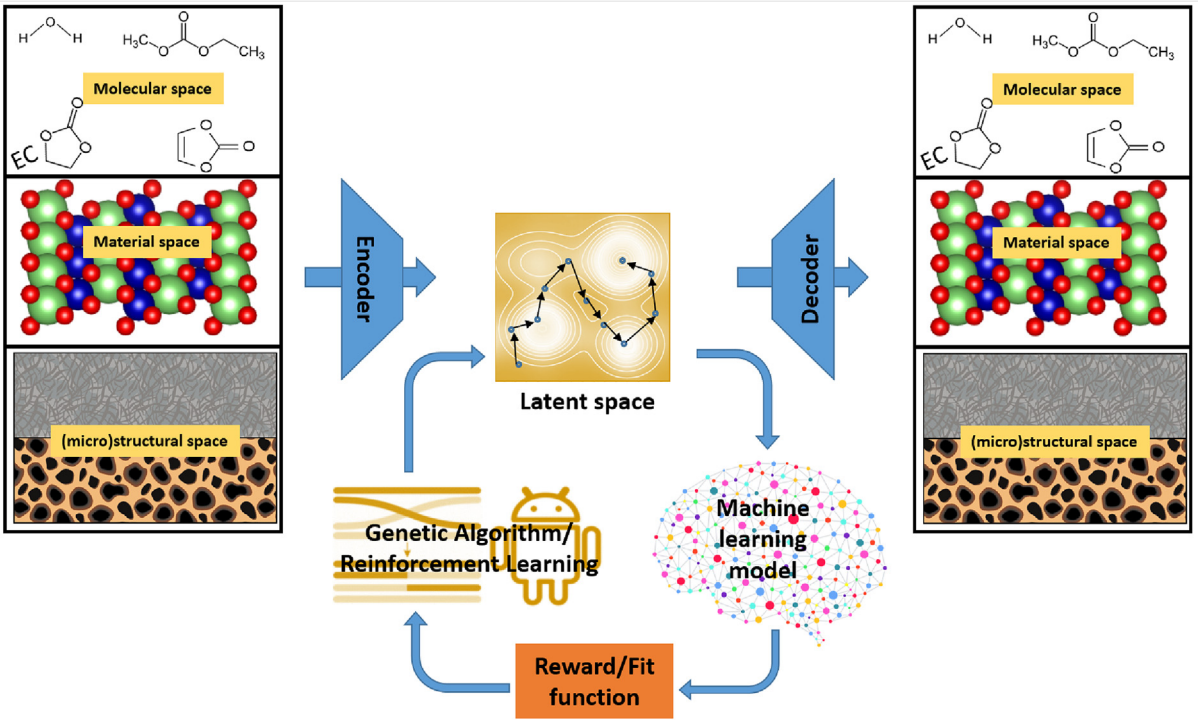


**Fig. 4.** Variational auto encoder (VAE) based encoding and decoding of chemical and structural information on a battery interphase into latent space, to enable generative battery interphase design through the use of genetic algorithms or reinforcement learning based exploration.

theoretical data (for training) is produced. This is especially true for databases aggregated from multiple research groups across the world. Furthermore, reported data often is incomplete in terms of the details of the battery interphase. The lack of standardization and completeness of the information can be addressed through the establishment of guidelines and data ontology [77,117]. Raw data from experiments and simulations need to be curated and represented in a interoperable and specific manner (fingerprint) to improve the training of ML models and their predictive power. This is especially true for supervised learning [118], where the physical model is already well established, e.g. predicting total energy from a molecular or crystal structure [119–122]. The choice of representation depends both on the physics that is emulated and the specific ML algorithm. Improved ML methods can learn pattern like symmetry operations from unrefined data and thus do away with complex fingerprinting processes. Molecular energies learned through e.g., kernel ridge regression [123], needs symmetry to be built into the molecular structure fingerprints, whereas convolutional neural networks can automatically learn underlying symmetries [124]. The complexity of the physical models and the diversity in the data landscape coming from a variety of experimental and simulation based sources for a given battery interphase makes effective fingerprint creation highly challenging. By utilizing the flexibility of the learning framework, it is possible to escape creating a large cohort of fingerprinting that is needed to encompass the known and unknown physics that occur at the battery interphase.

Deep generative models (DGL) will become progressively more efficient in an autopilot mode by allowing the creation and assimilation of more data through ML orchestrated simulations and experiments. DGL models allow exploration of uncharted territories in phase space and are able to detect outlier systems, where the current physical modelling approaches might be inadequate. Variational auto encoder (VAE) [125] based deep generative models are viable for this approach. VAE enables the creation of unique representations that are both lightweight and appropriate for the specific physical process that needs to be learned (Fig. 4). VAE is a powerful unsupervised learning approach to model an unlabeled distribution of the data. By utilizing data from particularly well characterized battery interphases, the VAE can be trained in a semi-supervised setting which is expected to improve the representation and classification.

VAEs are good at filtering out the scatter [126] in experimental and simulation data originating from differences in simulation parameters and experimental conditions. Another critical benefit of VAE in semi-supervised learning for battery interphases is the possibility of working with incomplete datasets (as is often the case with experimental data) and make the best utilization of a small set of available experimental data through the completion of the dataset with the technique of neural inpainting [127]. Using disentangled VAE for semi-supervised learning can achieve further improvement by making reliable assumptions on a subset of interpretable variables and rely on the flexibility of neural networks to learn representations for the remaining variables [128]. Utilization of compressed representation from disentangled VAE will likely enable much faster environment exploration through optimized explorer agents [129], based on, e.g., GA or RL in the final step of the proposed learning framework.

An alternative approach is to utilize another recently developed generative model of generative adversarial network (GAN) [130,131] for conjuring new interphase configurations, which have similar characteristics as existing interphases for which the synthesis conditions are well established. GANs can be trained to generate interphase systems such that generated systems have a high probability of being stable (stability as refining factor). This is in start contrast to VAEs, which are known to generate unphysical systems from random latent representations [132]. Yet another direction can be explored, where unsupervised learning is used to cluster battery interphase into groups with similar performance defining physical processes [133]. This is followed by identification of a descriptive physical model, using existing data point within the group for which experimental and(or) theoretical information exist.

For all of the approaches described above, access to sufficient training data is a bottleneck. This can, however, be addressed in part by the latest developments in meta-learning methods like one shot learning, which can work with few reliable data points [134,135].

### 4.2. Uncertainty estimation in atomic-scale simulations and deep learning

While uncertainty estimation is frequently used in experimental characterization of batteries, e.g., in estimations of state of charge, state of health, etc. (SoX) [136], the concept is less established in supervised deep learning techniques [137]. It can be particularly challenging when the network encounters conditions it has not been exposed to during training [47]. Methods like the use of probabilistic weights in deep neural network is gaining popularity in deep learning methods; an approach which is also compatible with reinforcement learning [138]. Although novel approaches like noise contrastive priors can be used to obtain uncertainty estimate [50], quantifiable predictive uncertainty estimates remain a central challenge for DGL models in general [48]. This is especially true in molecular/atomic science [139], where predictions are often overconfident. Reliable uncertainty quantification is particularly important using deep generative models to explore materials optimization in numerous parameter search space using highly sparse datasets, as these can be used to determine when particular experiments or simulations are needed in a specific part of parameter space [140].

The specific ML-based framework for generative battery interphase design outlined above is well suited for treating and predicting uncertainties. It is, however, essential to prioritize the requirements of further experimental work and physics-based simulations specifically for generative exploration in sparsely sampled parameter space. The proposed time series model output includes uncertainty estimates of each parameter. In essence, it is modelling two types of uncertainties, the inherent uncertainty of the process, also called *aleatoric uncertainty*, as well as the systematic uncertainty or *epistemic uncertainty*, which is the uncertainty that arises due to insufficient data to provide reliable estimates of the response variables. The aleatoric uncertainty can, for example, be modeled by outputting a mean and variance for each variable to model, e.g. the probability distribution $p(y_t|z_t)$ as a Gaussian distribution with mean $\mu(z_t)$ and variance $\sigma^2(z_t)$. Another option is quantile regression [141], where the model is trained to output quantiles rather than point estimates. The epistemic uncertainty is more challenging for deep learning models, but initial steps in this direction have been taken via approximate Bayesian inference [51,142,143] and ensemble models [49,144,145]. The type of uncertainty and at which time step it occurs aids the decision about which simulation, synthesis, characterization or test to run next. Note that while some of the observable variables will come from the training data, the rest will be generated by the model.

In addition to the model uncertainty, the uncertainty of the training data must also be taken into consideration. In atomic scale simulations and (high-fidelity) materials prediction, uncertainty estimation is an emerging field with a limited number of applications in the design of clean energy materials. In DFT-level calculations, the specific choice of exchange-correlation functional can have large implications on whether predictive accuracy can be achieved for a given system. The development of the Bayesian error estimation functional with van der Waals correlation (BEEF–vdW) [146] has set a new standard for uncertainty estimation in atomic-scale simulations and, e.g., used to predict high-fidelity lithium-graphite phase diagrams [147]. The generated ensemble of possible exchange correlation functionals can also be used to identify intrinsic DFT-errors associated with the formation and breaking of specific chemical bonds [148,149]. By correcting the DFT-level training data, improved accuracy of the DGL models can thus be obtained.

## 5. Outlook

Designing high-performance battery interfaces in general and interphases like the SEI and CEI specifically remains a Grand Challenge in the discovery and development of emerging and future battery chemistries. Here, we have outlined a possible path to achieve the highly ambitious goal of enabling the accelerated inverse design of battery interphases through utilization of semi-supervised generative deep learning (DGL) models with uncertainty estimation. The outlined multimodal approach would combine data-driven models with physical insights and models when available, e.g. by enforcing parameter bounds based on the laws of physics. The described approach utilizes training data spanning all relevant domains of the battery discovery and development cycle, i.e., multi-scale computer simulations and materials databases, structural and electrochemical characterization, synthesis and manufacturing, as well as cell and pack level testing and utilization.

The DGL approach stand on automated learning of compressed representation of chemical/structural details of battery interphases and identification of descriptors that define the limiting physical phenomena. This data assimilation and generation method not only facilitate explorative DGL in whole latent space but also provide physical insights, beyond that is feasible with only human intuition, from models with identified battery interface descriptors. A central aspect in the approach is that the DGL models should be capable of utilizing uncertainty estimation to identify when new training data is needed. More importantly, the model needs to query/request data from the specific domain or technique, where new data can bring maximum value for this specific training task. For example, whether high-fidelity atomic-scale computer simulations or operando characterization data is needed, or (low-fidelity) high-throughput electrochemical testing is more valuable.

The success of the outlined AI-orchestrated approach will depend extensively on the availability of data. A guiding example of the required dataset size can be found in deep learning for geosciences. The corresponding heterogeneity of data sources, parameters bound by the laws of physics, and a spatio-temporal dependence of external parameter is comparable to the SEI and CEI. Dozens of petabyte of data is currently available, but generative models for predicting long-time scale events, e.g., droughts, are still limited by the model uncertainty [150]. Comparable or larger data sets is expected to be required for inverse design of dynamic battery interphases, which stresses the need for standardization and curation of data across research disciplines and domains.

A first validation of the generative approach can be found within drug-like molecule and synthesis reaction design [151], where the datasets employed are now in excess of a million molecules [152], e.g. for the training of long short-term memory (LSTM) recurrent neural network (RNN) [153]. As long as only static properties are considered, similar dataset sizes should be applicable to battery interphases. Prediction of properties over time can be achieved with a semi-supervised approach, where cycle data is available for only a subset of materials.

The promise of AI and deep learning in materials science shines brightly [154], but it is important also to keep potential failure risk in mind. In the presence of adequate data and a fully operational BIG-MAP infrastructure, we believe that a 10x acceleration in the total discovery time would be within reach, leading to a 1–2 year discovery time within the next 5–10 years [6]. That being said, it is important to stress that the acceleration pace will depend on the specific battery chemistry and interface/phase in question and the time-scale of the associated targets/goals for the inverse design procedure.

Ultimately, the outlined approaches could even be used to identify possible self-healing additives that could be added to the electrolyte and activated via external stimuli, if needed to heal or re-active specific battery interphase possesses. This provided another distinctly interesting opportunity in the development of ultra-high performance battery interphases in future battery technologies.

## Acknowledgements

The authors acknowledges support from VILLUM FONDEN by DeepDFT research grant (0023105) and the European Union's Horizon 2020 research and innovation programs FET Open (766581 and 711792) and FET Proactive (824066 and 854472).

## Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ensm.2019.06.011.

## References


[1] M. Winter, The solid electrolyte interphase – the most important and the least understood solid electrolyte in rechargeable li batteries, Zeitschrift Für Phys Chemie 223 (2009) 1395–1406, https://doi.org/10.1524/zpch.2009.6086.

[2] F. Single, A. Latz, B. Horstmann, Identifying the mechanism of continued growth of the solid-electrolyte interphase, ChemSusChem 11 (2018), https://doi.org/10.1002/cssc.201800077, 1950–5.

[3] J. Alvarado, M.A. Schroeder, M. Zhang, O. Borodin, E. Gobrogge, M. Olguin, et al., A carbonate-free, sulfone-based electrolyte for high-voltage li-ion batteries, Mater Today 21 (2018) 341–353, https://doi.org/10.1016/j.mattod.2018.02.005.

[4] M.M. Kabir, D.E. Demirocak, Degradation mechanisms in li-ion batteries : a state-of- the-art review, Int J Energy Res 41 (2017) 1963–1986, https://doi.org/10.1002/er.3762.

[5] Tejs Vegge, Kristian Sommer Thygesen, A. Aspuru-Guzik, clean energy materials innovation challenge, in: Accelerating the clean energy revolution - perspectives on innovation challenges: DTU International Energy Report, 2018.

[6] A.A. Aspuru-Guzik, K. Persson, Materials Acceleration Platform: Accelerating Advanced Energy Materials Discovery by Integrating High-Throughput Methods and Artificial Intelligence, Mission Innovation IC6, 2018.

[7] R. Younesi, G.M. Veith, P. Johansson, K. Edström, T. Vegge, Lithium salts for advanced lithium batteries: Li–metal, Li–$O_2$ , and Li–S, Energy Environ Sci 8 (2015) 1905–1922, https://doi.org/10.1039/C5EE01215E.

[8] M. Gauthier, T.J. Carney, A. Grimaud, L. Giordano, N. Pour, H.-H. Chang, et al., Electrode–electrolyte interface in li-ion batteries: current understanding and new insights, J Phys Chem Lett 6 (2015) 4653–4672, https://doi.org/10.1021/acs.jpclett.5b01727.

[9] H. Park, N. Kumar, M. Melander, T. Vegge, J.M. Garcia Lastra, D.J. Siegel, Adiabatic and nonadiabatic charge transport in li–s batteries, Chem Mater 30 (2018) 915–928, https://doi.org/10.1021/acs.chemmater.7b04618.

[10] Y.S. Mekonnen, J.M. Garcia-Lastra, J.S. Hummelshøj, C. Jin, T. Vegge, Role of $Li_2O_2$@$Li_2CO_3$ interfaces on charge transport in nonaqueous li–air batteries, J Phys Chem C 119 (2015) 18066–18073, https://doi.org/10.1021/acs.jpcc.5b04432.

[11] M. Melander, E.Ö. Jónsson, J.J. Mortensen, T. Vegge, J.M. García Lastra, Implementation of constrained dft for computing charge transfer rates within the projector augmented wave method, J Chem Theory Comput 12 (2016) 5367–5378, https://doi.org/10.1021/acs.jctc.6b00815.

[12] M. Morcrette, Y. Chabre, G. Vaughan, G. Amatucci, J.-B. Leriche, S. Patoux, et al., In situ X-ray diffraction techniques as a powerful tool to study battery electrode materials, Electrochim Acta 47 (2002) 3137–3149, https://doi.org/10.1016/S0013-4686(02)00233-5.

[13] J.S.G. Myrdal, D. Blanchard, D. Sveinbjörnsson, T. Vegge, Li-ion conduction in the $LiBH_4$:LiI system from density functional theory calculations and quasi-elastic neutron scattering, J Phys Chem C 117 (2013) 9084–9091, https://doi.org/10.1021/jp311980h.

[14] Y. Zhu, C. Wang, Galvanostatic intermittent titration technique for phase-transformation electrodes, J Phys Chem C 114 (2010) 2830–2841, https://doi.org/10.1021/jp9113333.

[15] V. Ruiz, A. Pfrang, A. Kriston, N. Omar, P. Van den Bossche, L. Boon-Brett, A review of international abuse testing standards and regulations for lithium ion batteries in electric and hybrid electric vehicles, Renew Sustain Energy Rev 81 (2018) 1427–1452, https://doi.org/10.1016/j.rser.2017.05.195.

[16] S. Abada, G. Marlair, A. Lecocq, M. Petit, V. Sauvant-Moynot, F. Huet, Safety focused modeling of lithium-ion batteries: A review, J Power Sources 306 (2016) 178–192, https://doi.org/10.1016/j.jpowsour.2015.11.100.

[17] A.A. Franco, Multiscale modelling and numerical simulation of rechargeable lithium ion batteries: concepts, methods and challenges, RSC Adv 3 (2013) 13027, https://doi.org/10.1039/c3ra23502e.

[18] V. Ramadesigan, P.W.C. Northrop, S. De, S. Santhanagopalan, R.D. Braatz, V.R. Subramanian, Modeling and simulation of Lithium-Ion batteries from a systems engineering perspective, J Electrochem Soc 159 (2012) R31–45, https://doi.org/10.1149/2.018203jes.

[19] F. Röder, R.D. Braatz, U. Krewer, Multi-scale simulation of heterogeneous surface film growth mechanisms in lithium-ion batteries, J Electrochem Soc 164 (2017) E3335–E3344, https://doi.org/10.1149/2.0241711jes.

[20] Y. Zhao, P. Stein, Y. Bai, M. Al-Siraj, Y. Yang, B.-X. Xu, A review on modeling of electro-chemo-mechanics in lithium-ion batteries, J Power Sources 413 (2019) 259–283, https://doi.org/10.1016/j.jpowsour.2018.12.011.

[21] A.A. Franco, A. Rucci, D. Brandell, C. Frayret, M. Gaberscek, P. Jankowski, et al., Boosting rechargeable batteries R&D by multiscale modeling: myth or reality? Chem Rev 119 (2019) 4569–4627, https://doi.org/10.1021/acs.chemrev.8b00239.

[22] T.M. Østergaard, L. Giordano, I.E. Castelli, F. Maglia, B.K. Antonopoulos, Y. Shao-Horn, et al., Oxidation of ethylene carbonate on li metal oxide surfaces, J Phys Chem C 122 (2018) 10442–10449, https://doi.org/10.1021/acs.jpcc.8b01713.

[23] H.H. Kristoffersen, T. Vegge, H.A. Hansen, OH formation and $H_2$ adsorption at the liquid water–Pt(111) interface, Chem Sci 9 (2018) 6912–6921, https://doi.org/10.1039/C8SC02495B.

[24] L. Zhang, J. Han, H. Wang, R. Car, E W, Deep potential molecular dynamics: a scalable model with the accuracy of quantum mechanics, Phys Rev Lett 120 (2018) 143001, https://doi.org/10.1103/PhysRevLett.120.143001.

[25] M.M. Melander, M.J. Kuisma, T.E.K. Christensen, K. Honkala, Grand-canonical approach to density functional theory of electrocatalytic systems: thermodynamics of solid-liquid interfaces at constant ion and electrode potentials, J Chem Phys 150 (2019), 041706, https://doi.org/10.1063/1.5047829.

[26] N.G. Hörmann, O. Andreussi, N. Marzari, Grand canonical simulations of electrochemical interfaces in implicit solvation models, J Chem Phys 150 (2019), 041730, https://doi.org/10.1063/1.5054580.

[27] J.R. Maple, U. Dinur, A.T. Hagler, Derivation of force fields for molecular mechanics and dynamics from ab initio energy surfaces, Proc Natl Acad Sci 85 (1988) 5350–5354.

[28] S.S. Han, A.C.T. van Duin, W.A. Goddard, H.M. Lee, Optimization and application of lithium parameters for the reactive force field, ReaxFF. J Phys Chem A 109 (2005) 4575–4582, https://doi.org/10.1021/jp051450m.

[29] D. Porezag, T. Frauenheim, T. Köhler, G. Seifert, R. Kaschner, Construction of tight-binding-like potentials on the basis of density-functional theory: application to carbon, Phys Rev B 51 (1995) 12947–12957, https://doi.org/10.1103/PhysRevB.51.12947.

[30] S. Hong, A. Karim, T.S. Rahman, K. Jacobi, G. Ertl, Selective oxidation of ammonia on RuO2(110): A combined DFT and KMC study, J Catal 276 (2010) 371–381, https://doi.org/10.1016/j.jcat.2010.09.029.

[31] G. Klimeck, R.C. Bowen, T.B. Boykin, C. Salazar-Lazaro, T.A. Cwik, A. Stoica, Si tight-binding parameters from genetic algorithm fitting, Superlattices Microstruct 27 (2000) 77–88, https://doi.org/10.1006/spmi.1999.0797.

[32] Z. Li, J.R. Kermode, A. De Vita, Molecular dynamics with on-the-fly machine learning of quantum-mechanical forces, Phys Rev Lett 114 (2015), 096405, https://doi.org/10.1103/PhysRevLett.114.096405.

[33] V. Botu, R. Batra, J. Chapman, R. Ramprasad, Machine learning force fields: construction, validation, and outlook, J Phys Chem C 121 (2017) 511–522, https://doi.org/10.1021/acs.jpcc.6b10908.

[34] A. Khorshidi, A.A. Peterson, Amp : a modular approach to machine learning in atomistic simulations, Comput Phys Commun 207 (2016) 310–324, https://doi.org/10.1016/j.cpc.2016.05.010.

[35] R. Ramakrishnan, P.O. Dral, M. Rupp, O.A. von Lilienfeld, Big data meets quantum chemistry approximations: the δ-machine learning approach, J Chem Theory Comput 11 (2015) 2087–2096, https://doi.org/10.1021/acs.jctc.5b00099.

[36] B. Huang, O.A. von Lilienfeld, The "DNA" of chemistry: Scalable quantum machine learning with "amons, 2017. 170704146.

[37] J.H. Chang, D. Kleiven, M. Melander, J. Akola, J.M. Garcia-Lastra, T. Vegge, CLEASE: a versatile and user-friendly implementation of cluster expansion method, J Phys Condens Matter 31 (2019) 325901, https://doi.org/10.1088/1361-648X/ab1bbc.

[38] V.L. Deringer, G. Csányi, Machine learning based interatomic potential for amorphous carbon, Phys Rev B 95 (2017), 094203, https://doi.org/10.1103/PhysRevB.95.094203.

[39] E. Prada, D. Di Domenico, Y. Creff, J. Bernard, V. Sauvant-Moynot, F. Huet, Simplified electrochemical and thermal model of $LiFePO_4$ -graphite li-ion batteries for fast charge applications, J Electrochem Soc 159 (2012) A1508–A1519, https://doi.org/10.1149/2.064209jes.

[40] B. Rajabloo, A. Jokar, M. Desilets, M. Lacroix, An inverse method for estimating the electrochemical parameters of lithium-ion batteries: ii: implementation, J Electrochem Soc 164 (2017), https://doi.org/10.1149/2.0221702jes. A99–105.

[41] K.A. Severson, P.M. Attia, N. Jin, N. Perkins, B. Jiang, Z. Yang, et al., Data-driven prediction of battery cycle life before capacity degradation, Nat Energy 4 (2019) 383–391, https://doi.org/10.1038/s41560-019-0356-8.

[42] K. Goebel, B. Saha, A. Saxena, J. Celaya, J. Christophersen, Prognostics in battery health management, IEEE Instrum Meas Mag 11 (2008) 33–40, https://doi.org/10.1109/MIM.2008.4579269.

[43] N.A. Chaturvedi, R. Klein, J. Christensen, J. Ahmed, A. Kojic, Algorithms for advanced battery-management systems, IEEE Control Syst 30 (2010) 49–68, https://doi.org/10.1109/MCS.2010.936293.

[44] D.N. Rakhmatov, S.B.K. Vrudhula, An analytical high-level battery model for use in energy management of portable electronic systems, in: IEEE/ACM Int. Conf. Comput. Aided Des. ICCAD 2001. IEEE/ACM Dig. Tech. Pap. (Cat. No.01CH37281), IEEE, 2001, pp. 488–493, https://doi.org/10.1109/ICCAD.2001.968687.

[45] Rong Peng, M. Pedram, An analytical model for predicting the remaining battery capacity of lithium-ion batteries, IEEE Trans Very Large Scale Integr Syst 14 (2006) 441–451, https://doi.org/10.1109/TVLSI.2006.876094.

[46] Y. Yang, P. Perdikaris, Adversarial uncertainty quantification in physics-informed neural networks, J Comput Phys 394 (2019) 136–152, https://doi.org/10.1016/j.jcp.2019.05.027.

[47] M. Teye, H. Azizpour, K. Smith, Bayesian Uncertainty Estimation for Batch Normalized Deep Networks, 2018. 180206455.

[48] B. Lakshminarayanan, A. Pritzel, C. Blundell, Simple and scalable predictive uncertainty estimation using deep ensembles, NIPS'17 Proc. 31st Int. Conf. Neural Inf. Process. Syst. (2017) 6402–6413.

[49] T. Pearce, M. Zaki, A. Brintrup, N. Anastassacos, A. Neely, Uncertainty in Neural Networks: Bayesian Ensembling, 2018. ArXiv Prepr.

[50] D. Hafner, D. Tran, T. Lillicrap, A. Irpan, J. Davidson, Reliable Uncertainty Estimates in Deep Neural Networks using Noise Contrastive Priors, 2018. 180709289.

[51] C. Blundell, J. Cornebise, K. Kavukcuoglu, D. Wierstra, Weight Uncertainty in Neural Networks, 2015. 150505424.

[52] L.J. Nelson, G.L.W. Hart, F. Zhou, V. Ozoliņš, Compressive sensing as a paradigm for building physics models, Phys Rev B 87 (2013), 035125, https://doi.org/10.1103/PhysRevB.87.035125.

[53] D. Duvenaud, D. Maclaurin, J. Aguilera-Iparraguirre, R. Gómez-Bombarelli, T. Hirzel, A. Aspuru-Guzik, et al., Convolutional Networks on Graphs for Learning Molecular Fingerprints, 2015. 150909292.

[54] L.M. Ghiringhelli, J. Vybiral, E. Ahmetcik, R. Ouyang, S.V. Levchenko, C. Draxl, et al., Learning physical descriptors for materials science by compressed sensing, New J Phys 19 (2017), 023017, https://doi.org/10.1088/1367-2630/aa57bf.

[55] F. Bianchini, A. Glielmo, J.R. Kermode, A. De Vita, Enabling QM-accurate simulation of dislocation motion in γ-Ni and α-Fe using a hybrid multiscale approach, Phys Rev Mater 3 (2019), https://doi.org/10.1103/PhysRevMaterials.3.043605.

[56] M. Caccin, Z. Li, J.R. Kermode, A. De Vita, A framework for machine-learning-augmented multiscale atomistic simulations on parallel supercomputers, Int J Quantum Chem 115 (2015) 1129–1139, https://doi.org/10.1002/qua.24952.

[57] A.F. Zahrt, J.J. Henle, B.T. Rose, Y. Wang, W.T. Darrow, S.E. Denmark, Prediction of higher-selectivity catalysts by computer-driven workflow and machine learning, Science (80- ) (2019) 363, https://doi.org/10.1126/science.aau5631, eaau5631.

[58] J.K. Nørskov, T. Bligaard, J. Rossmeisl, C.H. Christensen, Towards the computational design of solid catalysts, Nat Chem 1 (2009) 37–46, https://doi.org/10.1038/nchem.121.

[59] Z. Li, S. Wang, H. Xin, Toward artificial intelligence in catalysis, Nat Catal 1 (2018) 641–642, https://doi.org/10.1038/s41929-018-0150-1.

[60] Z. Ahmad, T. Xie, C. Maheshwari, J.C. Grossman, V. Viswanathan, Machine learning enabled computational screening of inorganic solid electrolytes for suppression of dendrite formation in lithium metal anodes, ACS Cent Sci 4 (2018) 996–1006, https://doi.org/10.1021/acscentsci.8b00229.

[61] D. Strmcnik, I.E. Castelli, J.G. Connell, D. Haering, M. Zorko, P. Martins, et al., Electrocatalytic transformation of HF impurity to $H_2$ and LiF in lithium-ion batteries, Nat Catal 1 (2018) 255–262, https://doi.org/10.1038/s41929-018-0047-z.

[62] C. Cao, Abate II, E. Sivonxay, B. Shyam, C. Jia, B. Moritz, et al., Solid Electrolyte Interphase on Native Oxide-Terminated Silicon Anodes for Li-Ion Batteries, Joule 3 (2019) 762–781, https://doi.org/10.1016/j.joule.2018.12.013.

[63] S. Lysgaard, M.K. Christensen, H.A. Hansen, J.M. García Lastra, P. Norby, T. Vegge, Combined DFT and differential electrochemical mass spectrometry investigation of the effect of dopants in secondary zinc-air batteries, ChemSusChem 11 (2018) 1933–1941, https://doi.org/10.1002/cssc.201800225.

[64] L.A. Selis, J.M. Seminario, Dendrite formation in silicon anodes of lithium-ion batteries, RSC Adv 8 (2018) 5255–5267, https://doi.org/10.1039/C7RA12690E.

[65] Y.-M. Chiang, Building a better battery, Science (80- ) 330 (2010) 1485–1486, https://doi.org/10.1126/science.1198591.

[66] C. Lupo, D. Schlettwein, Modeling of dendrite formation as a consequence of diffusion-limited electrodeposition, J Electrochem Soc 166 (2019) D3182–D3189, https://doi.org/10.1149/2.0231901jes.

[67] D. Wheeler, T.P. Moffat, D. Josell, Spatial-temporal modeling of extreme bottom-up filling of through-silicon-vias, J Electrochem Soc 160 (2013) D3260–D3265, https://doi.org/10.1149/2.040312jes.

[68] F.A. Soto, A. Marzouk, F. El-Mellouhi, P.B. Balbuena, Understanding ionic diffusion through sei components for lithium-ion and sodium-ion batteries: insights from first-principles calculations, Chem Mater 30 (2018) 3315–3322, https://doi.org/10.1021/acs.chemmater.8b00635.

[69] E. Peled, S. Menkin, Review—sei: past, present and future, J Electrochem Soc 164 (2017) A1703–A1719, https://doi.org/10.1149/2.1441707jes.

[70] S.J. An, J. Li, C. Daniel, D. Mohanty, S. Nagpure, D.L. Wood, The state of understanding of the lithium-ion-battery graphite solid electrolyte interphase (SEI) and its relationship to formation cycling, Carbon N Y 105 (2016) 52–76, https://doi.org/10.1016/j.carbon.2016.04.008.

[71] A. Wang, S. Kadam, H. Li, S. Shi, Y. Qi, Review on modeling of the anode solid electrolyte interphase (SEI) for lithium-ion batteries, Npj Comput Mater 4 (2018) 15, https://doi.org/10.1038/s41524-018-0064-0.

[72] B. Horstmann, F. Single, A. Latz, Review on multi-scale models of solid-electrolyte interphase formation, Curr Opin Electrochem 13 (2019) 61–69, https://doi.org/10.1016/j.coelec.2018.10.013.

[73] F. Hao, Z. Liu, P.B. Balbuena, P.P. Mukherjee, Mesoscale elucidation of solid electrolyte interphase layer formation in li-ion battery anode, J Phys Chem C 121 (2017) 26233–26240, https://doi.org/10.1021/acs.jpcc.7b09465.

[74] M. Farkhondeh, M. Pritzker, M. Fowler, C. Delacourt, Mesoscopic modeling of a $LiFePO_4$ electrode: experimental validation under continuous and intermittent operating conditions, J Electrochem Soc 164 (2017) E3040–E3053, https://doi.org/10.1149/2.0211706jes.

[75] R. Imhof, Oxidative electrolyte solvent degradation in lithium-ion batteries: an in situ differential electrochemical mass spectrometry investigation, J Electrochem Soc 146 (1999) 1702, https://doi.org/10.1149/1.1391829.

[76] L. Giordano, P. Karayaylali, Y. Yu, Y. Katayama, F. Maglia, S. Lux, et al., Chemical reactivity descriptor for the oxide-electrolyte interface in li-ion batteries, J Phys Chem Lett 8 (2017) 3881–3887, https://doi.org/10.1021/acs.jpclett.7b01655.

[77] G. Pizzi, A. Cepellotti, R. Sabatini, N. Marzari, B. Kozinsky, AiiDA: automated interactive infrastructure and database for computational science, Comput Mater Sci 111 (2016) 218–230, https://doi.org/10.1016/j.commatsci.2015.09.013.

[78] X. Shi, Z. Chen, H. Wang, D.-Y. Yeung, W. Wong, W. Woo, Convolutional LSTM Network: A Machine Learning Approach for Precipitation Nowcasting, 2015, pp. 1–12. 13126114.

[79] J. Kremer, K. Stensbo-Smidt, F. Gieseke, K.S. Pedersen, C. Igel, Big universe, big data: machine learning and image analysis for astronomy, IEEE Intell Syst 32 (2017) 16–22, https://doi.org/10.1109/MIS.2017.40.

[80] D.C. Parkes, M.P. Wellman, Economic reasoning and artificial intelligence, Science (80- ) 349 (2015) 267–272, https://doi.org/10.1126/science.aaa8403.

[81] T. Ching, D.S. Himmelstein, B.K. Beaulieu-Jones, A.A. Kalinin, B.T. Do, G.P. Way, et al., Opportunities and obstacles for deep learning in biology and medicine, J R Soc Interface 15 (2018) 20170387, https://doi.org/10.1098/rsif.2017.0387.

[82] D.J. Lary, A.H. Alavi, A.H. Gandomi, A.L. Walker, Machine learning in geosciences and remote sensing, Geosci Front 7 (2016) 3–10, https://doi.org/10.1016/j.gsf.2015.07.003.

[83] A. Radovic, M. Williams, D. Rousseau, M. Kagan, D. Bonacorsi, A. Himmel, et al., Machine learning at the energy and intensity frontiers of particle physics, Nature 560 (2018) 41–48, https://doi.org/10.1038/s41586-018-0361-2.

[84] M.W. Libbrecht, W.S. Noble, Machine learning applications in genetics and genomics, Nat Rev Genet 16 (2015) 321–332, https://doi.org/10.1038/nrg3920.

[85] E. Gawehn, J.A. Hiss, G. Schneider, Deep learning in drug discovery, Mol Inform 35 (2016) 3–14, https://doi.org/10.1002/minf.201501008.

[86] Z. Obermeyer, E.J. Emanuel, Predicting the future — big data, machine learning, and clinical medicine, N Engl J Med 375 (2016) 1216–1219, https://doi.org/10.1056/NEJMp1606181.

[87] M. Chen, Y. Hao, K. Hwang, L. Wang, L. Wang, Disease prediction by machine learning over big data from healthcare communities, IEEE Access 5 (2017) 8869–8879, https://doi.org/10.1109/ACCESS.2017.2694446.

[88] C. Zheng, K. Mathew, C. Chen, Y. Chen, H. Tang, A. Dozier, et al., Automated generation and ensemble-learned matching of X-ray absorption spectra, Npj Comput Mater 4 (2018) 12, https://doi.org/10.1038/s41524-018-0067-x.

[89] S.K. Suram, Y. Xue, J. Bai, R. Le Bras, B. Rappazzo, R. Bernstein, et al., Automated phase mapping with AgileFD and its application to light absorber discovery in the V–Mn–Nb oxide system, ACS Comb Sci 19 (2017) 37–46, https://doi.org/10.1021/acscombsci.6b00153.

[90] S. Russell, P. Norvig, Artificial Intelligence: A Modern Approach, 3rd ed., Prentice Hall Press, Upper Saddle River, NJ, USA, 2009.

[91] X. Hu, S.E. Li, Y. Yang, Advanced machine learning approach for lithium-ion battery state estimation in electric vehicles, IEEE Trans Transp Electrif 2 (2016) 140–149, https://doi.org/10.1109/TTE.2015.2512237.

[92] T. Zahid, K. Xu, W. Li, C. Li, H. Li, State of charge estimation for electric vehicle power battery using advanced machine learning algorithm under diversified drive cycles, Energy 162 (2018) 871–882, https://doi.org/10.1016/j.energy.2018.08.071.

[93] E. Chemali, P.J. Kollmeyer, M. Preindl, A. Emadi, State-of-charge estimation of Li-ion batteries using deep neural networks: A machine learning approach, J Power Sources 400 (2018) 242–255, https://doi.org/10.1016/j.jpowsour.2018.06.104.

[94] I. Goodfellow, Y. Bengio, A. Courville, Deep Learning, MIT Press, 2016. www.deeplearningbook.org.

[95] Y. LeCun, Y. Bengio, G. Hinton, Deep learning, Nature 521 (2015) 436–444, https://doi.org/10.1038/nature14539.

[96] J. Schmidhuber, Deep learning in neural networks: An overview, Neural Networks 61 (2015) 85–117, https://doi.org/10.1016/j.neunet.2014.09.003.

[97] A. Jain, J.A. Bollinger, T.M. Truskett, Inverse methods for material design, AIChE J 60 (2014) 2732–2740, https://doi.org/10.1002/aic.14491.

[98] B. Sanchez-Lengeling, A. Aspuru-Guzik, Inverse molecular design using machine learning: Generative models for matter engineering, Science (80- ) 361 (2018) 360–365, https://doi.org/10.1126/science.aat2663.

[99] P. Nguyen, T. Tran, S. Gupta, S. Rana, S. Venkatesh, Hybrid generative-discriminative models for inverse materials design, ArXiv Prepr ArXiv181106060 1–18 (2018).

[100] K. Alberi, M.B. Nardelli, A. Zakutayev, L. Mitas, S. Curtarolo, A. Jain, et al., The 2019 materials by design roadmap, J Phys D Appl Phys 52 (2019), 013001, https://doi.org/10.1088/1361-6463/aad926.

[101] M. Aykol, V.I. Hegde, L. Hung, S. Suram, P. Herring, C. Wolverton, et al., Network analysis of synthesizable materials discovery, Nat Commun 10 (2019) 2018, https://doi.org/10.1038/s41467-019-10030-5.

[102] S. Kirklin, J.E. Saal, B. Meredig, A. Thompson, J.W. Doak, M. Aykol, et al., The Open Quantum Materials Database (OQMD): assessing the accuracy of DFT formation energies, Npj Comput Mater 1 (2015) 15010, https://doi.org/10.1038/npjcompumats.2015.10.

[103] J.E. Saal, S. Kirklin, M. Aykol, B. Meredig, C. Wolverton, Materials design and discovery with high-throughput density functional theory: the open quantum materials database (oqmd), JOM 65 (2013) 1501–1509, https://doi.org/10.1007/s11837-013-0755-4.

[104] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, et al., Imagenet large scale visual recognition challenge, Int J Comput Vis 115 (2015) 211–252, https://doi.org/10.1007/s11263-015-0816-y.

[105] M.I. Jordan, T.M. Mitchell, Machine learning: trends, perspectives, and prospects, Science (80- ) 349 (2015) 255–260, https://doi.org/10.1126/science.aaa8415.

[106] P. Raccuglia, K.C. Elbert, P.D.F. Adler, C. Falk, M.B. Wenny, A. Mollo, et al., Machine-learning-assisted materials discovery using failed experiments, Nature 533 (2016) 73–76, https://doi.org/10.1038/nature17439.

[107] J.M. Garcia-Lastra, J.S.G. Myrdal, R. Christensen, K.S. Thygesen, T. Vegge, DFT+U Study of Polaronic Conduction in Li 2 O 2 and Li 2 CO 3 : Implications for Li–Air Batteries, J Phys Chem C 117 (2013) 5568–5577, https://doi.org/10.1021/jp3107809.

[108] J.R. Norris, Markov Chains (Cambridge Series in Statistical and Probabilistic Mathematics), Cambridge University Press, 1998.

[109] S. Hochreiter, J. Schmidhuber, Long short-term memory, Neural Comput 9 (1997) 1735–1780, https://doi.org/10.1162/neco.1997.9.8.1735.
[110] M. Fraccaro, S.K. Sønderby, U. Paquet, O. Winther, Sequential Neural Models with Stochastic Layers, 2016. 160507571.
[111] V. Mnih, K. Kavukcuoglu, D. Silver, A.A. Rusu, J. Veness, M.G. Bellemare, et al., Human-level control through deep reinforcement learning, Nature 518 (2015) 529, https://doi.org/10.1038/nature14236.
[112] R.S. Sutton, A.G. Barto, Reinforcement learning: An introduction, MIT press, 2018.
[113] K. Deb, A. Pratap, S. Agarwal, T. Meyarivan, A fast and elitist multiobjective genetic algorithm: NSGA-II, IEEE Trans Evol Comput 6 (2002) 182–197, https://doi.org/10.1109/4235.996017.
[114] S.N. Sivanandam, S.N. Deepa, Genetic Algorithm Optimization Problems. Introd. to Genet. Algorithms, Springer Berlin Heidelberg, Berlin, Heidelberg, 2008, pp. 165–209, https://doi.org/10.1007/978-3-540-73190-0_7.
[115] L. Ardizzone, J. Kruse, S. Wirkert, D. Rahner, E.W. Pellegrini, R.S. Klessen, et al., Analyzing Inverse Problems with Invertible Neural Networks, 2018. 180804730.
[116] R.T.Q. Chen, Y. Rubanova, J. Bettencourt, D. Duvenaud, Neural Ordinary Differential Equations, 2018, pp. 6572–6583. 180607366.
[117] R. Jose, S. Ramakrishna, Materials 4.0: Materials big data enabled materials discovery, Appl Mater Today 10 (2018) 127–132, https://doi.org/10.1016/j.apmt.2017.12.015.
[118] S.B. Kotsiantis, Supervised machine learning: A review of classification techniques, in: Proc. 2007 Conf. Emerg. Artif. Intell. Appl. Comput. Eng. Real Word AI Syst. with Appl. eHealth, HCI, Inf. Retr. Pervasive Technol, 160, 2007, pp. 3–24.
[119] G. Imbalzano, A. Anelli, D. Giofré, S. Klees, J. Behler, M. Ceriotti, Automatic selection of atomic fingerprints and reference configurations for machine-learning potentials, J Chem Phys 148 (2018) 241730, https://doi.org/10.1063/1.5024611.
[120] O. Isayev, D. Fourches, E.N. Muratov, C. Oses, K. Rasch, A. Tropsha, et al., Materials cartography: representing and mining materials space using structural and electronic fingerprints, Chem Mater 27 (2015) 735–743, https://doi.org/10.1021/cm503507h.
[121] R. Ramprasad, R. Batra, G. Pilania, A. Mannodi-Kanakkithodi, C. Kim, Machine learning in materials informatics: recent applications and prospects, Npj Comput Mater 3 (2017) 54, https://doi.org/10.1038/s41524-017-0056-5.
[122] O.A. von Lilienfeld, R. Ramakrishnan, M. Rupp, A. Knoll, Fourier series of atomic radial distribution functions: A molecular fingerprint for machine learning models of quantum chemical properties, Int J Quantum Chem 115 (2015) 1084–1093, https://doi.org/10.1002/qua.24912.
[123] P.C. Jennings, S. Lysgaard, J.S. Hummelshøj, T. Vegge, T. Bligaard, Genetic algorithms for computational materials discovery accelerated by machine learning, Npj Comput Mater 5 (2019) 46, https://doi.org/10.1038/s41524-019-0181-4.
[124] K.T. Schütt, F. Arbabzadah, S. Chmiela, K.R. Müller, A. Tkatchenko, Quantum-chemical insights from deep tensor neural networks, Nat Commun 8 (2017) 13890, https://doi.org/10.1038/ncomms13890.
[125] D.P. Kingma, M. Welling, Auto-Encoding Variational Bayes, 2013. 13126114.
[126] D.J. Im, S. Ahn, R. Memisevic, Y. Bengio, Denoising Criterion for Variational Auto-Encoding Framework, 2015. 151106406.
[127] R.A. Yeh, C. Chen, T.Y. Lim, A.G. Schwing, M. Hasegawa-Johnson, M.N. Do, Semantic image inpainting with deep generative models, in: 2017 IEEE Conf. Comput. Vis. Pattern Recognit, IEEE, 2017, pp. 6882–6890, https://doi.org/10.1109/CVPR.2017.728.
[128] N. Siddharth, B. Paige, J.-W. van de Meent, A. Desmaison, N.D. Goodman, P. Kohli, et al., Learning Disentangled Representations with Semi-Supervised Deep Generative Models, 2017. 170600400.
[129] L.P. Kaelbling, M.L. Littman, A.W. Moore, Reinforcement learning: a survey, J Artif Intell Res 4 (1996) 237–285, https://doi.org/10.1613/jair.301.
[130] I.J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, et al., Generative Adversarial Networks. 14062661 2014:2672–80.
[131] A. Radford, L. Metz, S. Chintala, Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks, 2015, https://doi.org/10.1051/0004-6361/201527329. 151106434.
[132] R. Gómez-Bombarelli, J.N. Wei, D. Duvenaud, J.M. Hernández-Lobato, B. Sánchez-Lengeling, D. Sheberla, et al., Automatic chemical design using a data-driven continuous representation of molecules, ACS Cent Sci 4 (2018) 268–276, https://doi.org/10.1021/acscentsci.7b00572.
[133] L. Maaløe, M. Fraccaro, O. Winther, Semi-Supervised Generation with Cluster-aware Generative Models, 2017. 170400637.
[134] T. Yu, C. Finn, A. Xie, S. Dasari, T. Zhang, P. Abbeel, et al., One-Shot Imitation from Observing Humans via Domain-Adaptive Meta-Learning, 2018. 180201557.
[135] O. Vinyals, C. Blundell, T. Lillicrap, K. Kavukcuoglu, D. Wierstra, Matching Networks for One Shot Learning, 2016. 160604080.
[136] Z. Li, J. Huang, B.Y. Liaw, J. Zhang, On state-of-charge determination for lithium-ion batteries, J Power Sources 348 (2017) 281–301, https://doi.org/10.1016/j.jpowsour.2017.03.001.
[137] D. Krishnamurthy, H. Weiland, A. Barati Farimani, E. Antono, J. Green, V. Viswanathan, Machine learning based approaches to accelerate energy materials discovery and optimization, ACS Energy Lett 4 (2018) 187–191.
[138] C. Blundell, J. Cornebise, K. Kavukcuoglu, D. Wierstra, Weight uncertainty in neural networks, ICML'15 Proc. 32nd Int. Conf. Int. Conf. Mach. Learn. 37 (2015) 1613–1622.
[139] P.B. Jørgensen, M.N. Schmidt, O. Winther, Deep generative models for molecular science, Mol Inform 37 (2018) 1700133, https://doi.org/10.1002/minf.201700133.
[140] J. Ling, M. Hutchinson, E. Antono, S. Paradiso, B. Meredig, High-dimensional materials and process optimization using data-driven experimental design with well-calibrated uncertainty estimates, Integr Mater Manuf Innov 6 (2017) 207–217, https://doi.org/10.1007/s40192-017-0098-z.
[141] R. Koenker, K.F. Hallock, Quantile regression, J Econ Perspect 15 (2001) 143–156, https://doi.org/10.1257/jep.15.4.143.
[142] D.J. Rezende, S. Mohamed, D. Wierstra, Stochastic Backpropagation and Approximate Inference in Deep Generative Models, 2014. 14014082.
[143] Y. Gal, Z. Ghahramani, Dropout as a Bayesian Approximation: Representing Model Uncertainty in Deep Learning, 2015. 150602142.
[144] B. Lakshminarayanan, A. Pritzel, C. Blundell, Simple and Scalable Predictive Uncertainty Estimation using Deep Ensembles, 2016. 161201474.
[145] I. Osband, J. Aslanides, A. Cassirer, Randomized Prior Functions for Deep Reinforcement Learning, 2018. 180603335.
[146] J. Wellendorff, K.T. Lundgaard, A. Møgelhøj, V. Petzold, D.D. Landis, J.K. Nørskov, et al., Density functionals for surface science: Exchange-correlation model development with Bayesian error estimation, Phys Rev B 85 (2012) 235149, https://doi.org/10.1103/PhysRevB.85.235149.
[147] V. Pande, V. Viswanathan, Robust high-fidelity DFT study of the lithium-graphite phase diagram, Phys Rev Mater 2 (2018) 125401, https://doi.org/10.1103/PhysRevMaterials.2.125401.
[148] R. Christensen, H.A. Hansen, C.F. Dickens, J.K. Nørskov, T. Vegge, Functional Independent Scaling Relation for ORR/OER Catalysts, J Phys Chem C 120 (2016) 24910–24916, https://doi.org/10.1021/acs.jpcc.6b09141.
[149] R. Christensen, H.A. Hansen, T. Vegge, Catalysis science & technology, Catal Sci Technol 5 (2015) 4946–4949, https://doi.org/10.1039/C5CY01332A.
[150] M. Reichstein, G. Camps-Valls, B. Stevens, M. Jung, J. Denzler, N. Carvalhais, et al., Deep learning and process understanding for data-driven earth system science, Nature 566 (2019) 195–204, https://doi.org/10.1038/s41586-019-0912-1.
[151] T. Dimitrov, C. Kreisbeck, J.S. Becker, A. Aspuru-Guzik, S.K. Saikin, Autonomous molecular design: then and now, ACS Appl Mater Interfaces (2019), https://doi.org/10.1021/acsami.9b01226 acsami.9b01226.
[152] M. Nakata, T. Shimazaki, M. Hashimoto, T. Maeda, PubChemQC PM6: A dataset of 221 million molecules with optimized molecular geometries and electronic properties, 2019. 190406046.
[153] M.H.S. Segler, T. Kogej, C. Tyrchan, M.P. Waller, Generating focused molecule libraries for drug discovery with recurrent neural networks, ACS Cent Sci 4 (2018) 120–131, https://doi.org/10.1021/acscentsci.7b00512.
[154] M. Umehara, H.S. Stein, D. Guevarra, P.F. Newhouse, D.A. Boyd, J.M. Gregoire, Analyzing machine learning models to accelerate generation of fundamental materials insights, Npj Comput Mater 5 (2019) 34, https://doi.org/10.1038/s41524-019-0172-5.