Computational Materials Science 129 (2017) 311–322

Contents lists available at ScienceDirect

Computational Materials Science

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

Optimal experimental design for materials discovery

Roozbeh Dehghannasiri [c,b,*], Dezhen Xue [a,d], Prasanna V. Balachandran [a], Mohammadmahdi R. Yousefi [e], Lori A. Dalton [e,f], Turab Lookman [a,*], Edward R. Dougherty [c,b,*]

[a] Theoretical Division, Los Alamos National Laboratory, Los Alamos, NM 87545, USA
[b] TEES-AgriLife Center for Bioinformatics and Genomic Systems Engineering, Texas A&M University, College Station, TX 77843, USA
[c] Department of Electrical and Computer Engineering, Texas A&M University, College Station, TX 77843, USA
[d] State Key Laboratory for Mechanical Behavior of Materials, Xi'an Jiaotong University, Xi'an 710049, China
[e] Department of Electrical and Computer Engineering, The Ohio State University, Columbus, OH 43210, USA
[f] Department of Biomedical Informatics, The Ohio State University, Columbus, OH 43210, USA

ARTICLE INFO

Article history:
Received 6 July 2016
Received in revised form 7 November 2016
Accepted 23 November 2016
Available online 13 January 2017

Keywords:
Optimal experimental design
Materials design
Mean objective cost of uncertainty (MOCU)

ABSTRACT

In this paper, we propose a general experimental design framework for optimally guiding new experiments or simulations in search of new materials with desired properties. The method uses the knowledge of previously completed experiments or simulations to recommend the next experiment which can effectively reduce the pertinent model uncertainty affecting the materials properties. To illustrate the utility of the proposed framework, we focus on a computational problem that utilizes time-dependent Ginzburg-Landau (TDGL) theory for shape memory alloys to calculate the stress-strain profiles for a particular dopant at a given concentration. Our objective is to design materials with the lowest energy dissipation at a specific temperature. The aim of experimental design is to suggest the best dopant and its concentration for the next TDGL simulation. Our experimental design utilizes the mean objective cost of uncertainty (MOCU), which is an objective-based uncertainty quantification scheme that measures uncertainty based upon the increased operational cost it induces. We analyze the performance of the proposed method and compare it with other experimental design approaches, namely random selection and pure exploitation.

1. Introduction

The Materials Genome Initiative (MGI) in the U.S. [1] has catalyzed much recent interest in accelerating materials discovery. One of the outstanding challenges in materials science is to reduce the number of costly and time-consuming trial and error experiments required to find new materials with desired properties. This is not a trivial problem because the materials search space is vast due to the complex interplay of structural, chemical and microstructural degrees of freedom and only a small fraction has been experimentally investigated [2]. High-throughput efforts, including high-throughput calculations and combinatorial experiments, have largely been the approaches of choice to narrow the combinatorial search space [3–5]. Recently, there has been much interest in using data-driven machine learning tools for optimally guiding experiments or calculations towards materials with desired properties [2,6–16]. Such methods have met considerable success in fields such as game theory, pattern recognition, artificial intelligence, and event forecasting. However, the application of pure data-driven approaches to materials science can be biased and yield suboptimal results, as the available training data are quite limited compared to the number of features (or material descriptors) and size of the search space [17–19]. A distinct advantage of materials science is that knowledge in the form of constitutive or scaling relations and various constraint equations is often available from theory or known empirically. Such prior knowledge can be used in conjunction with data to quantify uncertainty and construct operators that are optimal relative to that uncertainty. Moreover, these operators on average outperform those designed solely using data. The efficacy of this approach has been successfully demonstrated in the biological sciences. For instance, in genomics there is a large body of knowledge regarding gene/protein signaling pathways. This knowledge can be transformed in such a way as to be useful for constructing biomarkers [20,21] and then

* Corresponding authors at: Department of Electrical and Computer Engineering, Texas A&M University, College Station, TX 77843, USA (R. Dehghannasiri, E.R. Dougherty), Theoretical Division, Los Alamos National Laboratory, Los Alamos, NM 87545, USA (T. Lookman).

E-mail addresses: roozbehdn@tamu.edu (R. Dehghannasiri), xuedezhen@mail.xjtu.edu.cn (D. Xue), pbalachandran@lanl.gov (P.V. Balachandran), yousefi@ece.osu.edu (M.R. Yousefi), dalton@ece.osu.edu (L.A. Dalton), txl@lanl.gov (T. Lookman), edward@ece.tamu.edu (E.R. Dougherty).

http://dx.doi.org/10.1016/j.commatsci.2016.11.041
0927-0256/

incorporated into a Bayesian framework to design optimal classifiers for decisions involving patient diagnosis, prognosis, and therapy [22–25]. How these methods can be utilized to predict properties and guide new experiments is therefore of significant importance and has not been demonstrated in materials science.

Our focus in this work is on experimental design. Experimental design has a long and varied history in science and engineering, the reason being that a properly designed experimental procedure provides much greater efficiency than simply making random probes. Indeed, Francis Bacon's call for experimental design in the *New Organon* in 1620 is often taken to be the beginning of modern science. In particular, in biomedicine, where the processes are very complex, one can gather a virtually limitless amount of information without getting to the crux of the matter. In view of the large number of measurements often needed in materials science, we deal with this problem as a multi-dimensional optimization problem, which typically requires training data in order to be solved [26,27]. Prior knowledge regarding parameters and features affecting the desired properties of materials is also crucial. However, it is often the case that prior knowledge is insufficient and the presence of large uncertainties degrades the experimental design. Therefore, one needs to improve the predictability of the model with respect to the ultimate objective by making additional measurements, which in turn requires synthesizing new materials. Therefore, it is necessary to target experimental efforts where the material with the desirable properties may be found by minimizing the number of experiments. This can be done via *experimental design* that distinguishes between different experiments based upon information they can provide.

In the present study, we demonstrate experimental design by formulating a computational problem based on the time-dependent Ginzburg-Landau (TDGL) theory for shape memory alloys (SMAs) that has been shown to capture reasonably well the underlying physics of the shape memory effect (SME) and superelasticity (SE) [28]. SMAs are a subclass of martensitic structural phase transitions that display SME and SE. The SE effect arises when an alloy above a certain temperature is deformed such that on unloading it returns to the original strain state. They are important because of their high tensile strength and ability to recover [29–31]. When the high symmetry austenite structure is exposed to stress beyond a critical value, it transitions to the low symmetry martensite phase and when the stress is removed, the martensite reverts back to the parent austenite phase [32]. Typically, such first-order phase transitions are accompanied by large hysteresis in the stress-strain curve, as shown in Fig. 1(a). Therefore, the whole loading-unloading process results in a hysteresis stress-strain loop. The closed area inside the curve is a measure of the amount of energy dissipated during the stress-strain cycle. For practical applications, large energy dissipation or hysteresis is undesirable because it affects the fatigue properties of SMAs in devices (such as cardiovascular stents) that require high sensitivity, precision, and durability. Therefore, lowering the energy dissipation accompanying SE is critical for realizing SMAs in practical applications and serves as our computational *design target* in this work. More specifically, we integrate TDGL theory (which computes the stress-strain curve for prototypical SMAs such as FePd) with our experimental design to rapidly select the material-specific model parameters that minimize the energy dissipation associated with SE. In real experiments, chemical modification of SMAs, for example by doping alloying elements in the host alloy, can affect the shape of the stress-strain SE response and consequently, the energy dissipation. In our simulation, we "mimic" chemical doping by varying the model parameters (additional details discussed in the next section).

The outline of the paper is as follows. In Section 2, we discuss the computational model for SMAs based on mesoscale Landau theory. This serves as the "oracle" to calculate properties and we also use it as a source of data to fit an empirical model that serves as input to our design. In Section 3, we discuss and develop our experimental design strategy which acts on the results from the computational model of Section 2. Section 4 discusses the results from our design and evaluates its performance compared to using a random selection strategy for guiding the next experiment or using the best predicted model value. In Section 5, we distill the key ideas and state the general framework for materials design where model parameters or features are unknown, and illustrate the algorithm in Section 5.1 with a worked example using a polynomial cost function. In Section 5.2, we provide a second example but in the context of a network problem with discrete states, which again explains the mathematics underlying our approach in a simple form. In this example, the cost function is defined as the probability of undesirable states.

2. Ginzburg-Landau theory

We discuss here the TDGL theory for SMAs. Note that throughout this paper we denote vectors by bold letters. We also use uppercase and lowercase letters to represent random variables and their realizations, respectively. Our model for the alloy is a two-dimensional version of the cubic to tetragonal martensitic transformation appropriate for materials such as FePd or InTl [28]. The symmetry-adapted strains e_1, e_2, and e_3 represent hydrostatic, deviatoric, and shear modes, respectively, where $e_2 = (1/2)(\epsilon_{xx} - \epsilon_{yy})$ is the strain responsible for the transition (order parameter, OP), and $e_1 = (1/2)(\epsilon_{xx} + \epsilon_{yy})$ and $e_3 = \epsilon_{xy}$. These strains are defined in terms of the linear strain components of elasticity in 2D, $\epsilon_{ij} = 1/2[(\partial u_i/\partial x_j) + (\partial u_j/\partial x_i)]$, where u_i is the i-th

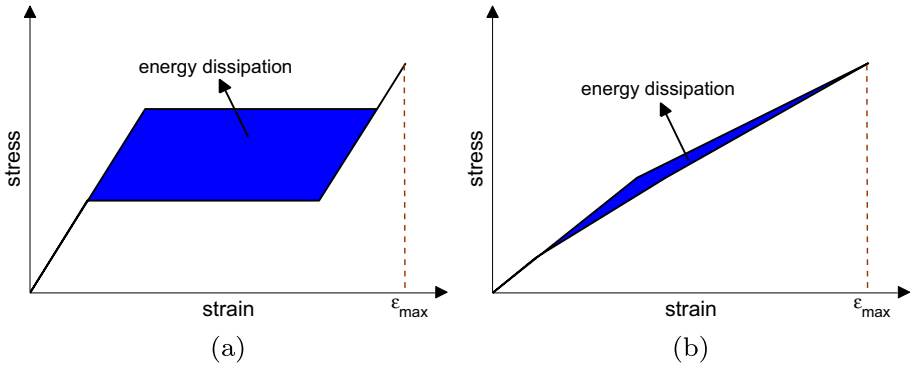

Fig. 1. Stress-strain curves before design and that showing targeted response. (a) Typical stress-strain curve for a shape memory alloy with typically large hysteresis giving rise to large dissipation. (b) The targeted stress-strain response with small hysteresis and dissipation. Our proposed experimental design guides the "next experiments" towards (b).

displacement component. Incorporating the influence of the internal stress field ρ associated with the dopants, the total free energy density is written as the summation of five contributions, namely,

$$F(e_1,e_2,e_3,\rho,\sigma_{11}) = f_h(e_2) + f_{grad}(\vec{\nabla} e_2) + f_{non-OP}(e_1,e_3) + f_{defect}(e_1,\rho) + f_{load}(e_1,e_2,\sigma_{11}), \quad (1)$$

where $f_h(e_2) = \frac{1}{2}A_2[T,\rho]e_2^2 + \frac{1}{4}\beta e_2^4 + \frac{1}{6}\gamma e_2^6$ is the homogeneous, bulk part accounting for the required nonlinearities in the order parameter, $f_{grad}(\vec{\nabla} e_2) = \frac{1}{2}g\left|\vec{\nabla} e_2\right|^2$ is the gradient (Ginzburg) term responsible for the interface energy in the order parameters, $f_{non-OP}(e_1,e_3) = \frac{1}{2}A_1e_1^2 + \frac{1}{2}A_3e_3^2$ is the contribution due to the non-order parameter components of the strain which we assume to be harmonic, $f_{defect}(e_1,\rho) = -e_1\rho$ captures the influence of the stress field ρ due to dopants, and $f_{load}(e_1,e_2,\sigma_{11}) = -\sqrt{2}(e_1+e_2)\sigma_{xx}$ is the energy associated with applying an external tensile load, σ_{xx} in the x direction.

In general, the dopants have different atomic sizes compared to the host alloy and the size mismatch gives rise to a local dilatational strain or stress. Thus, the effect of dopants can be modeled as a randomly distributed dilatational stress in the system and the concentrations can be varied by changing the number of dopants. The local dilatation stress will not only affect the dopant crystallographic site but also its nearest neighbors and beyond because of long-range elastic forces. Therefore, the dilatation stress will acquire a certain distribution rather than a Delta function. We assume a Gaussian distribution of the form

$$\rho(\vec{r}) = h\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{\vec{r}^2}{2\sigma^2}}. \quad (2)$$

Different chemical dopants will thus have varied strengths in the host lattices and the affected region will vary. Therefore, the combination of parameters h (strength) and σ (range of stress disturbance) in Eq. (2) characterize the dopants. By solving the following time dependent TDGL evolution equation,

$$\frac{\partial e_2(\vec{r},t)}{\partial t} = -\Gamma\frac{\delta F}{\delta e_2(\vec{r},t)}, \quad (3)$$

which relaxes the free energy $F(e_1,e_2,e_3,\rho,\sigma_{11})$ associated with the alloy, we are able to calculate the stress-strain profiles for a particular dopant at a given concentration (number fraction).

3. Experimental design using mean objective cost of uncertainty

Given that the aim of new measurements is to reduce model uncertainties associated with finding new materials with desirable properties, we quantify uncertainty by taking into account its effect on the dissipation, the objective. In other words, we need an objective-based uncertainty quantification scheme. The mean objective cost of uncertainty (MOCU) measures the deterioration in the performance of the designed operator, herein simulating materials, due to presence of model uncertainty – that is, it measures the degradation in performance between an operator optimally designed based on partial prior knowledge and data as compared to an operator optimally designed when there is full knowledge of the underlying system [33]. Recently, MOCU has also been utilized for experimental design in gene regulatory networks (GRNs) [34–36]. GRNs are used to study interactions among genes in order to develop drugs to mitigate aberrant phenotypes such as cancers [37]. These networks usually suffer from significant uncertainty as a result of the sophisticated gene regulatory mechanism in living organisms.

It should be recognized that MOCU is different from conventional uncertainty quantification schemes such as variance and entropy, which only consider the statistical characteristics of uncertainty irrespective of the objective. Classical experimental design with the aim of reducing variance or entropy has long been studied in machine learning [38–40] where the information gain of each experiment is measured in terms of the reduction in the entropy or variance.

In this section, we show how MOCU applies to the problem of designing materials with the lowest energy dissipation. We assume that there is a set of different dopants that can be used and the range (σ) and the strength or potency (h) of the dopants are unknown. As shown in Fig. 2, we first fit a model to the energy dissipation as a function of dopant concentration, dopant potency, and dopant range, and then use the trained model in the experimental design step to find the best dopant and concentration for

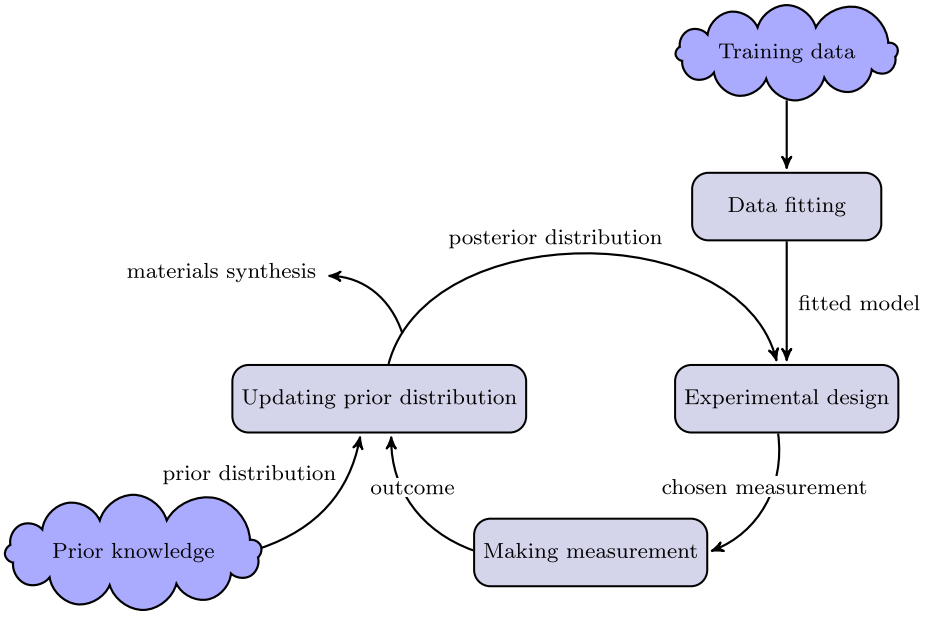

Fig. 2. An illustrative view of the experimental design process when used recursively.

the next measurement. Based on the observed outcome of the chosen measurement, the prior distribution that encodes the prior knowledge regarding unknown dopant parameters is updated to the posterior distribution. If one is interested in making more measurements, this distribution can be used as the new prior distribution for the next experimental design loop. In the case that no more measurement is needed, the posterior distribution can be used to design the low energy dissipation material. In Section 5, we will state the general framework for experimental design in materials discovery for designing materials with desirable properties given that some parameters or features in the problem are unknown.

3.1. Designing materials with low energy dissipation

For experimental design, we define the goal as finding the lowest energy dissipation possible at a specific temperature. The temperature that we consider is 285 K. The native (or host) material is characterized by three parameters: elastic strain gradient coefficient κ, fourth-order elastic constant β of x_2 which is the deviatoric shear strain, and sixth-order elastic constant γ of x_2 [28]. Parameters β and γ can be determined from thermodynamic, lattice parameter data.

We assume that there are n different available dopants each with a specific strength h_i and range σ_i. We represent the unknown potency and spread for each dopant by random variables H_i and Σ_i, respectively. The random vectors $\mathbf{H} = [H_1, H_2, \ldots, H_n]$ and $\mathbf{\Sigma} = [\Sigma_1, \Sigma_2, \ldots, \Sigma_n]$ denote the unknown potency and spread for all dopants. Let c be the concentration of dopant. We want to find which dopant i and concentration $c \in C$, C being the range of possible concentrations, should be used for the next measurement.

We split the whole experimental design problem into two parts. First we need a function to model the energy dissipation at a specific temperature as a function of dopant potency h, dopant spread σ, and dopant concentration c. We use a reciprocal function of the following form for energy dissipation:

$$g(h,\sigma,c;\mathbf{b}) = \max\left(\frac{1}{(b_1h+b_2\sigma+b_3h\sigma+b_4)c+b_5h+b_6\sigma+b_7h\sigma+b_8}, 0\right), \tag{4}$$

where $\mathbf{b} = [b_1, \ldots, b_8]$ is the vector of coefficients for the fitted model. Energy dissipation takes non-negative values. In Eq. (4), the operator $\max(p, q)$ chooses p if $p \geqslant q$ and q otherwise. The coefficient vector $\mathbf{b}$ can be estimated using training data generated from the TDGL code that numerically solves the partial differential equation (PDE) required for finding strain-stress values. The numerical algorithm defines and keeps track of a matrix for storing the stress at each location in a material. We randomly distribute c dopant atoms in the native material and our TDGL code incrementally applies stress to the material and then incrementally lowers stress to produce the stress-strain curve at the given temperature. We fit $g(h, \sigma, c; \mathbf{b})$ for energy dissipation to the training data obtained from the code using the iterative least squares estimation method.

After the data fitting step, we use the fitted model to find the next best measurement in the experimental design. Table 1 summarizes the notation we use for the minimum energy dissipation problem and the relationship to the general experimental design framework we will outline in Section 5. The uncertainty class is the class of unknown parameters in the problem and is $\Theta = \mathbf{H} \times \mathbf{\Sigma}$, the Cartesian product of random vectors $\mathbf{H}$ and $\mathbf{\Sigma}$. Therefore, each element $\theta = [\mathbf{h}, \boldsymbol{\sigma}]$ in the uncertainty class corresponds to a specific realization $\mathbf{h} = [h_1, \ldots, h_n]$ of $\mathbf{H}$ and $\boldsymbol{\sigma} = [\sigma_1, \ldots, \sigma_n]$ of $\mathbf{\Sigma}$. The prior probability distribution governing the uncertainty class is $f_{\mathbf{H},\mathbf{\Sigma}}(\mathbf{h}, \boldsymbol{\sigma}) = f_{\mathbf{H}}(\mathbf{h}) f_{\mathbf{\Sigma}}(\boldsymbol{\sigma})$, $f_{\mathbf{H}}(\mathbf{h})$ and $f_{\mathbf{\Sigma}}(\boldsymbol{\sigma})$ being the probability distributions for the random vectors $\mathbf{H}$ and $\mathbf{\Sigma}$, respectively. The cost function is defined as the energy dissipation $g(h_i, \sigma_i, c; \mathbf{b})$ obtained after adding dopant i with concentration c. Corresponding to each $\theta = [\mathbf{h}, \boldsymbol{\sigma}]$ value, the *model-specific optimal material* can be found as

$$i(\theta), c(\theta) = \underset{1 \leqslant i \leqslant n,\, c \in C}{\arg\min}\, g(h_i, \sigma_i, c; \mathbf{b}). \tag{5}$$

This defines the material with the lowest cost relative to the specific model parameter θ and there will be an optimal material for every value of θ. We define the *robust material* for the uncertainty class Θ as that with the lowest energy dissipation *on average* across the whole class. When dealing with uncertainties, it is this robust material we seek which performs optimally on average. This is given by

$$i(\Theta), c(\Theta) = \underset{1 \leqslant i \leqslant n,\, c \in C}{\arg\min}\, \mathrm{E}_{h_i,\sigma_i}[g(h_i, \sigma_i, c; \mathbf{b})], \tag{6}$$

where the expectation is taken over the distribution function $f_{H_i,\Sigma_i}(h_i, \sigma_i) = f_{H_i}(h_i) f_{\Sigma_i}(\sigma_i)$ of the parameters for the i-th dopant. It will have the lowest value on average, but not necessarily for any specific model θ. Having defined the robust and optimal materials, we define the *objective cost of uncertainty* relative to θ as the difference between the energy dissipations $g\big(h_{i(\Theta)}, \sigma_{i(\Theta)}, c(\Theta); \mathbf{b}\big)$ and $g\big(h_{i(\theta)}, \sigma_{i(\theta)}, c(\theta); \mathbf{b}\big)$. The mean objective cost of uncertainty (MOCU) is then defined by

$$\mathrm{M}(\Theta) = \mathrm{E}_{\theta}\big[g\big(h_{i(\Theta)}, \sigma_{i(\Theta)}, c(\Theta); \mathbf{b}\big) - g\big(h_{i(\theta)}, \sigma_{i(\theta)}, c(\theta); \mathbf{b}\big)\big], \tag{7}$$

where the expectation is taken over the distribution function $f_{\mathbf{H},\mathbf{\Sigma}}(\mathbf{h}, \boldsymbol{\sigma})$.

The motivation for these definitions and concepts is as follows. If we knew the exact values of h and σ parameters for each dopant, we would be able to obtain the optimal material according to Eq. (5). However, when dopant parameters are unknown, we have to find the robust material according to Eq. (6). Of course, the energy dissipation of the robust material is not as low as that of the optimal one for each specific value of dopant parameters, but it has the lowest energy dissipation on average when we consider all possible models (parameter values) and incorporate them via the expectation. That is, our best choice when we do not have any unknown parameter is Eq. (5) and the best choice when we have uncertainty is the robust solution Eq. (6). The difference between the energy dissipation of these two choices in the absence and the presence of uncertainty shows how much additional cost we have incurred due to the uncertainty. Therefore, we compute the energy dissipation difference for each model θ and as θ is unknown, we take the expectation and refer to it as MOCU, a measure of the expected additional cost in the presence of uncertainty.

In order to bring in the aspects guiding the next experiments, let $X_{i,c}$ be the random variable representing the measurement outcome using dopant i with concentration c. For $X_{i,c}$, we assume the conditional distribution $f_{X_{i,c}}(x_{i,c}|H_i = h_i, \Sigma_i = \sigma_i)$ given the parameters for dopant i are $H_i = h_i$ and $\Sigma_i = \sigma_i$. This distribution reflects the measurement error. It should be emphasized that although

Table 1
Correspondence between the specific experimental design with the aim of minimizing energy dissipation discussed in Section 3 and the general experimental design framework outlined in Section 5.

Experimental design problem for minimum dissipation	General framework
$[\mathbf{h}, \boldsymbol{\sigma}] \in \mathbf{H} \times \mathbf{\Sigma}$ (dopants parameters)	Uncertainty class $\theta \in \Theta$
(i, c), $1 \leqslant i \leqslant n$, $c \in C$ (doping process)	Operator ψ
$f_{\mathbf{H},\mathbf{\Sigma}}(\mathbf{h}, \boldsymbol{\sigma})$ (distribution of the parameters)	Prior distribution $f(\theta)$
$X_{i,c}$, $1 \leqslant i \leqslant n$, $c \in C$(measurement)	Experiment X
$g(h_i, \sigma_i, c; \mathbf{b})$ (energy dissipation)	Cost function $\zeta_\theta(\psi)$

we do not know the actual outcome of the measurements, we do know the distribution governing the outcomes. We use this distribution for the experimental design to determine the measurement to be made first.

Using Eq. (26), which we derive in Section 5, the dopant i^* and concentration c^* suggested by the proposed experimental design are those that result in the minimum expected remaining MOCU:

$$i^*, c^* = \arg\min_{1\leqslant i\leqslant n,\, c\in C} \mathrm{E}_{x_{i,c}}\left[\mathrm{E}_{\theta|x_{i,c}}\left[g\left(h_{i(\Theta|x_{i,c})}, \sigma_{i(\Theta|x_{i,c})}, c(\Theta|x_{i,c}); \mathbf{b}\right)\right]\right], \tag{8}$$

where the outer expectation is taken with respect to the distribution function $f_{X_{i,c}}(x_{i,c})$ of the measurement outcomes, obtained using the law of total probability:

$$\begin{aligned} f_{X_{i,c}}(x_{i,c}) &= \mathrm{E}_{h_i,\sigma_i}\left[f_{X_{i,c}}(x_{i,c}|H_i = h_i, \Sigma_i = \sigma_i)\right] \\ &= \int f_{X_{i,c}}(x_{i,c}|H_i = h_i, \Sigma_i = \sigma_i) f_{H_i,\Sigma_i}(h_i, \sigma_i)\, dh_i d\sigma_i, \end{aligned} \tag{9}$$

and the inner expectation in Eq. (8) is taken over $f_{\mathbf{H},\mathbf{\Sigma}}(\mathbf{h}, \boldsymbol{\sigma}|X_{i,c} = x_{i,c})$, which can be calculated using Bayes' rule,

$$f_{\mathbf{H},\mathbf{\Sigma}}(\mathbf{h}, \boldsymbol{\sigma}|X_{i,c} = x_{i,c}) = \frac{f_{X_{i,c}}(x_{i,c}|\mathbf{H} = \mathbf{h}, \mathbf{\Sigma} = \boldsymbol{\sigma}) f_{\mathbf{H},\mathbf{\Sigma}}(\mathbf{h}, \boldsymbol{\sigma})}{f_{X_{i,c}}(x_{i,c})}. \tag{10}$$

In the preceding equation, $f_{X_{i,c}}(x_{i,c}|\mathbf{H} = \mathbf{h}, \mathbf{\Sigma} = \boldsymbol{\sigma}) = f_{X_{i,c}}(x_{i,c}|H_i = h_i, \Sigma_i = \sigma_i)$ because the measurement outcome using dopant i depends only on the h and σ parameters of that dopant and not those of the other dopants. In Eq. (8), we also need to compute the posterior robust dopant $i(\Theta|x_{i,c})$ and robust concentration $c(\Theta|x_{i,c})$, which are robust relative to the posterior distribution $f_{\mathbf{H},\mathbf{\Sigma}}(\mathbf{h}, \boldsymbol{\sigma}|X_{i,c} = x_{i,c})$ after observing $x_{i,c}$ for the measurement $X_{i,c}$. They can be obtained similarly to Eq. (6):

$$i(\Theta|x_{i,c}), c(\Theta|x_{i,c}) = \arg\min_{1\leqslant i\leqslant n,\, c\in C} \mathrm{E}_{h_i\sigma_i|x_{i,c}}[g(h_i, \sigma_i, c; \mathbf{b})]. \tag{11}$$

Algorithm 1. Experimental design for materials with low energy dissipation

1: **input**:
$f_{\mathbf{H},\mathbf{\Sigma}}(\mathbf{h}, \boldsymbol{\sigma}), \{X_{i,c}, 1 \leqslant i \leqslant n, c \in C\}, f_{X_{i,c}}(x_{i,c}|H_i = h_i, \Sigma_i = \sigma_i)$
2: **output**: Optimal measurement (i^*, c^*)
3: **for all** $X_{i,c}, 1 \leqslant i \leqslant n, c \in C$ **do**
4: **for all** possible outcomes $x_{i,c}$ of $X_{i,c}$ **do**
5: Compute $f_{\mathbf{H},\mathbf{\Sigma}}(\mathbf{h}, \boldsymbol{\sigma}|X_{i,c} = x_{i,c})$ using (10)
6: Find the robust material $i(\Theta|x_{i,c}), c(\Theta|x_{i,c})$ using Eq. (11)
7: $\omega(x_{i,c}) \leftarrow \mathrm{E}_{\theta|x_{i,c}}\left[g\left(h_{i(\Theta|x_{i,c})}, \sigma_{i(\Theta|x_{i,c})}, c(\Theta|x_{i,c}); \mathbf{b}\right)\right]$
8: Compute $f_{X_{i,c}}(x_{i,c})$ using Eq. (9)
9: Compute $\mathrm{E}_{x_{i,c}}[\omega(x_{i,c})]$
10: $i^*, c^* \leftarrow \arg\min_{i=1,\ldots,n, c\in C} \mathrm{E}_{x_{i,c}}[\omega(x_{i,c})]$ (optimization in Eq. (8))
11: **return** i^*, c^* (optimal measurement)

Note that in Eq. (8), in order to find the best measurement, we need to search over all potential measurements (i, c) and enumerate their corresponding outcomes $x_{i,c}$. For each possible outcome $x_{i,c}$, we update the prior distribution $f_{\mathbf{H},\mathbf{\Sigma}}(\mathbf{h}, \boldsymbol{\sigma})$ to the posterior distribution $f_{\mathbf{H},\mathbf{\Sigma}}(\mathbf{h}, \boldsymbol{\sigma}|X_{i,c} = x_{i,c})$ and then find the posterior robust material $i(\Theta|x_{i,c}), c(\Theta|x_{i,c})$. The algorithm of the proposed method for this specific experimental design problem is given in Algorithm 1. A summary of the different probability distributions used in this specific experimental design problem is provided in Table 2.

Table 2
Summary of the different probability distributions required for the specific experimental design problem being studied.

Probability distributions	Source
Distribution of the potency parameters $f_{\mathbf{H}}(\mathbf{h})$	Prior knowledge
Distribution of the spread parameters $f_{\mathbf{\Sigma}}(\boldsymbol{\sigma})$	Prior knowledge
Distribution of the dopants parameters $f_{\mathbf{H},\mathbf{\Sigma}}(\mathbf{h}, \boldsymbol{\sigma})$	$f_{\mathbf{H}}(\mathbf{h}) f_{\mathbf{\Sigma}}(\boldsymbol{\sigma})$
Conditional distribution of the measurement outcome $f_{X_{i,c}}(x_{i,c}\|h_i, \sigma_i)$	Prior knowledge
Conditional distribution of the parameters $f_{\mathbf{H},\mathbf{\Sigma}}(\mathbf{h}, \boldsymbol{\sigma}\|X_{i,c} = x_{i,c})$	Bayes' rule Eq. (10)
Distribution of the measurement outcome $f_{X_{i,c}}(x_{i,c})$	Law of total probability Eq. (9)

Prior probabilities should reflect our prior knowledge regarding the uncertain parameters and constraints of the problem [41]. We assume that, based on the prior knowledge, although the parameters h and σ of the dopants are unknown, their relative rank is known. Suppose there are n dopants for which $h_{min} \leqslant h_i \leqslant h_{max}$ and $\sigma_{min} \leqslant \sigma_i \leqslant \sigma_{max}$ for $i = 1, \ldots, n$. The following arbitrary rankings for the h and σ parameters of the dopants are assumed:

$$h_{min} \leqslant h_{i_1} < h_{i_2} < \cdots < h_{i_n} \leqslant h_{max} \tag{12}$$

$$\sigma_{min} \leqslant \sigma_{j_1} < \sigma_{j_2} < \cdots < \sigma_{j_n} \leqslant \sigma_{max}. \tag{13}$$

In order to build appropriate priors for the uncertainty class, we assume that any $\mathbf{h}$ and $\boldsymbol{\sigma}$ vectors whose components satisfy Eqs. (12) and (13), respectively, are in the uncertainty class and all elements in the uncertainty class are equally likely. Any combination that does not follow the proper ranking has prior probability zero. Let random vector $\mathbf{P} = [P_1, P_2, \ldots, P_{n+1}]$ denote a set of new random variables defined by

$$\begin{aligned} P_1 &= \frac{H_{i_1} - h_{min}}{h_{max} - h_{min}} \\ P_2 &= \frac{H_{i_2} - H_{i_1}}{h_{max} - h_{min}} \\ &\vdots \\ P_{n+1} &= \frac{h_{max} - H_{i_n}}{h_{max} - h_{min}}. \end{aligned} \tag{14}$$

Similarly, random vector $\mathbf{Q} = [Q_1, Q_2, \ldots, Q_{n+1}]$ is defined by

$$\begin{aligned} Q_1 &= \frac{\Sigma_{j_1} - \sigma_{min}}{\sigma_{max} - \sigma_{min}} \\ Q_2 &= \frac{\Sigma_{j_2} - \Sigma_{j_1}}{\sigma_{max} - \sigma_{min}} \\ &\vdots \\ Q_{n+1} &= \frac{\sigma_{max} - \Sigma_{j_n}}{\sigma_{max} - \sigma_{min}}. \end{aligned} \tag{15}$$

In this way, random vectors $\mathbf{H}$ and $\mathbf{\Sigma}$ are mapped to new random vectors $\mathbf{P}$ and $\mathbf{Q}$, respectively. It can be seen that $\sum_{i=1}^{n+1} P_i = 1$ and $\sum_{i=1}^{n+1} Q_i = 1$.

The probability distribution that can be used for a random vector whose components are non-negative and add up to 1 is the Dirichlet distribution [42]. For a random vector $\mathbf{Y}$ of k components in $\mathbb{R}^k$, a Dirichlet distribution with parameters $\alpha_i > 0, 1 \leqslant i \leqslant k$, is defined as:

$$\mathbf{Y} = [Y_1, Y_2, \ldots, Y_k] \sim Dir(\alpha_1, \ldots, \alpha_k) \;\Rightarrow\; f_{\mathbf{Y}}(\mathbf{y}) \propto \prod_{i=1}^{k} y_i^{\alpha_i - 1}. \tag{16}$$

In Bayesian statistics, the Dirichlet distribution is often used as the prior distribution for the parameters of the multinomial distribution, an example being the use of Dirichlet priors in discrete

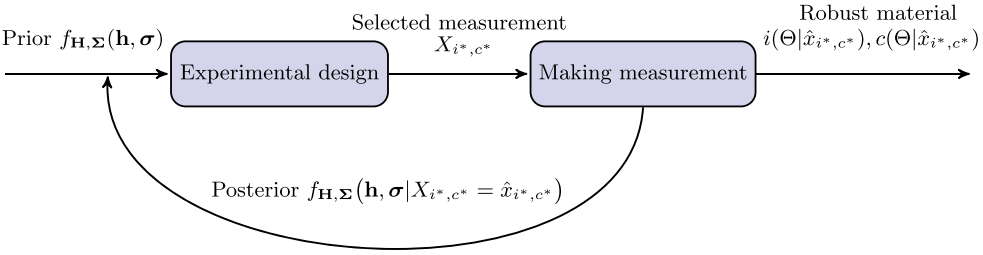

Fig. 3. The experimental design applied recursively to a sequence of measurements.

classification when the feature-label distribution contains unknown parameters [22,24]. In the case that $\alpha_i = 1$ for all i, the distribution is equivalent to a uniform distribution over the support region.

In general, two Dirichlet priors with different parameters α_i and μ_i can be assumed for $\mathbf{P}$ and $\mathbf{Q}$, respectively: $\mathbf{P} \sim Dir(\alpha_1, \ldots, \alpha_{n+1})$ and $\mathbf{Q} \sim Dir(\mu_1, \ldots, \mu_{n+1})$. Choosing appropriate values for the Dirichlet parameters is not straightforward, particularly due to the effect they can have on the statistical performance of the experimental design. When there is no reliable information regarding the distribution, it is reasonable to set the Dirichlet parameters to 1, making the Dirichlet distribution be equivalent to a uniform distribution over the support region. Consequently, all elements in the uncertainty have the same weight and we are not biased towards any of them *a priori*. Therefore, in this paper, we assume that $\mu_i = \alpha_i = 1$ for all i.

As illustrated in Fig. 3, the proposed experimental design method can be applied recursively for a sequence of measurements. Assuming that the prior $f_{\mathbf{H},\mathbf{\Sigma}}(\mathbf{h},\boldsymbol{\sigma})$ is used to find the best measurement i^*, c^* and the outcome of the measurement is $\hat{x}_{i^*,c^*}$, after observing the outcome, the prior distribution is updated to the posterior distribution $f_{\mathbf{H},\mathbf{\Sigma}}(\mathbf{h},\boldsymbol{\sigma}|X_{i^*,c^*} = \hat{x}_{i^*,c^*})$. Now for the next iteration, we consider the updated posterior distribution as the new prior distribution and use it to find the next measurement. As this process is repeated, the posterior distribution will eventually converge to a point mass on the underlying true parameters. It should be emphasized that, as can be seen in the figure, the best choices for measurement and designing materials are different. While the best choice for the measurement is i^*, c^*, which is selected by the proposed experimental design method in Eq. (8), the best choice for the materials design after observing $\hat{x}_{i^*,c^*}$ is the robust material $i(\Theta|\hat{x}_{i^*,c^*}), c(\Theta|\hat{x}_{i^*,c^*})$ obtained similar to Eq. (11).

4. Simulation results and performance analysis

To analyze the performance of the proposed experimental design method, we desire a metric to evaluate the effectiveness of a chosen measurement. Let $\mathbf{h}^{true}$ and $\boldsymbol{\sigma}^{true}$ be the underlying true parameter vectors, which are assumed to be unknown during experimental design. Suppose that we have made the measurement using dopant i with concentration c and the outcome of the measurement is x, i.e., $X_{i,c} = x$. The prior is updated from $f_{\mathbf{H},\mathbf{\Sigma}}(\mathbf{h},\boldsymbol{\sigma})$ to $f_{\mathbf{H},\mathbf{\Sigma}}(\mathbf{h},\boldsymbol{\sigma}|X_{i,c} = x)$ and the robust material relative to the updated prior is $i(\Theta|x), c(\Theta|x)$. The energy dissipation of the robust material relative to the underlying true parameters is $g\left(h^{true}_{i(\Theta|x)}, \sigma^{true}_{i(\Theta|x)}, c(\Theta|x); \mathbf{b}\right)$, which we use as a metric to evaluate the effectiveness of a chosen measurement.

What we "ideally" desire from an effective experimental design is that the energy dissipation of the robust material that would be obtained after running the simulation suggested by the experimental design method should result in lower energy dissipation compared to some other selection protocols. It is possible that for some cases the energy dissipation of the robust material obtained after the measurement chosen by the experimental design method results in larger energy dissipation, but our desire is that it be as low as possible on average for a large number of simulations when following the experimental design.

For simulations, we set the parameters of the native material to $\beta = -17000, \kappa = 25$, and $\gamma = 3 \times 10^7$. We run the code that numerically solves the PDE to find the stress-strain curve for $h = 20$, 25, and 30, $\sigma = 0.5$, 1, 2, and 3, and dopant concentration c taking values between 50 and 1500 with the step size of 50 and between 1500 and 2500 with the step size of 100. We repeat running the code four times and then use the average of the simulation results for the data fitting step. The coefficients of the energy dissipation model in Eq. (4) are found using the iterative least squares estimation method. Having performed data fitting, the coefficient vector is $\mathbf{b} = [1.4, -11.27, 0.75, -17.16, -39.07, 301.18, -11.23, 2.63 \times 10^3]$. Fig. 4 shows the data obtained from running the code and the result of the data fitting step for a number of values for dopant parameters h and σ, and concentration c.

We assumed that there are $n = 7$ different dopants whose σ parameter is between 0 and 5 and h parameter is between 12 and 25, and they have the following relative ranking:

$$12 \leqslant h_1 < h_2 < h_3 < h_4 < h_5 < h_6 < h_7 \leqslant 25$$
$$0 \leqslant \sigma_4 < \sigma_5 < \sigma_7 < \sigma_1 < \sigma_2 < \sigma_6 < \sigma_3 \leqslant 5.$$

We generated 150 different samples for each random vector $\mathbf{P}$ and $\mathbf{Q}$ using Dirichlet distributions whose parameters are 1. Then we used Eqs. (14) and (15) to find the corresponding samples for random vectors $\mathbf{H}$ and $\mathbf{\Sigma}$, respectively. Hence, in total the uncertainty class contains 150×150 different elements, each having a probability of $\frac{1}{150 \times 150}$. These generated samples are used to approximate the prior distribution $f_{\mathbf{H},\mathbf{\Sigma}}(\mathbf{h},\boldsymbol{\sigma})$. We assume that dopants can have concentration levels between 500 and 1000 with the step size of 50. Therefore, there are $7 \times 11 = 77$ different choices for simulations.

For the experimental design problem, we also need to define $f_{X_{i,c}}(x_{i,c}|h_i, \sigma_i)$. In order to do this, for a simulation (i, c) relative to the parameter values h_i and σ_i, we first compute $m = round(g(h_i, \sigma_i, c; \mathbf{b}) \times 10^6)$, where the $round(.)$ operator rounds to the nearest integer. Having done this, we discretize the possible outcomes. Based on the value of m, we use one of the distribution functions shown in Fig. 5 as the conditional distribution function $f_{X_{i,c}}(x_{i,c}|h_i, \sigma_i)$.

We demonstrate the effectiveness of the proposed experimental design method by comparing it with two other selection policies: pure exploitation and random selection. In the random selection policy, all dopants and concentrations can be chosen randomly for the next simulation with equal likelihood. In the pure exploitation approach, the next simulation is run using $i^{exploit}, c^{exploit}$, which

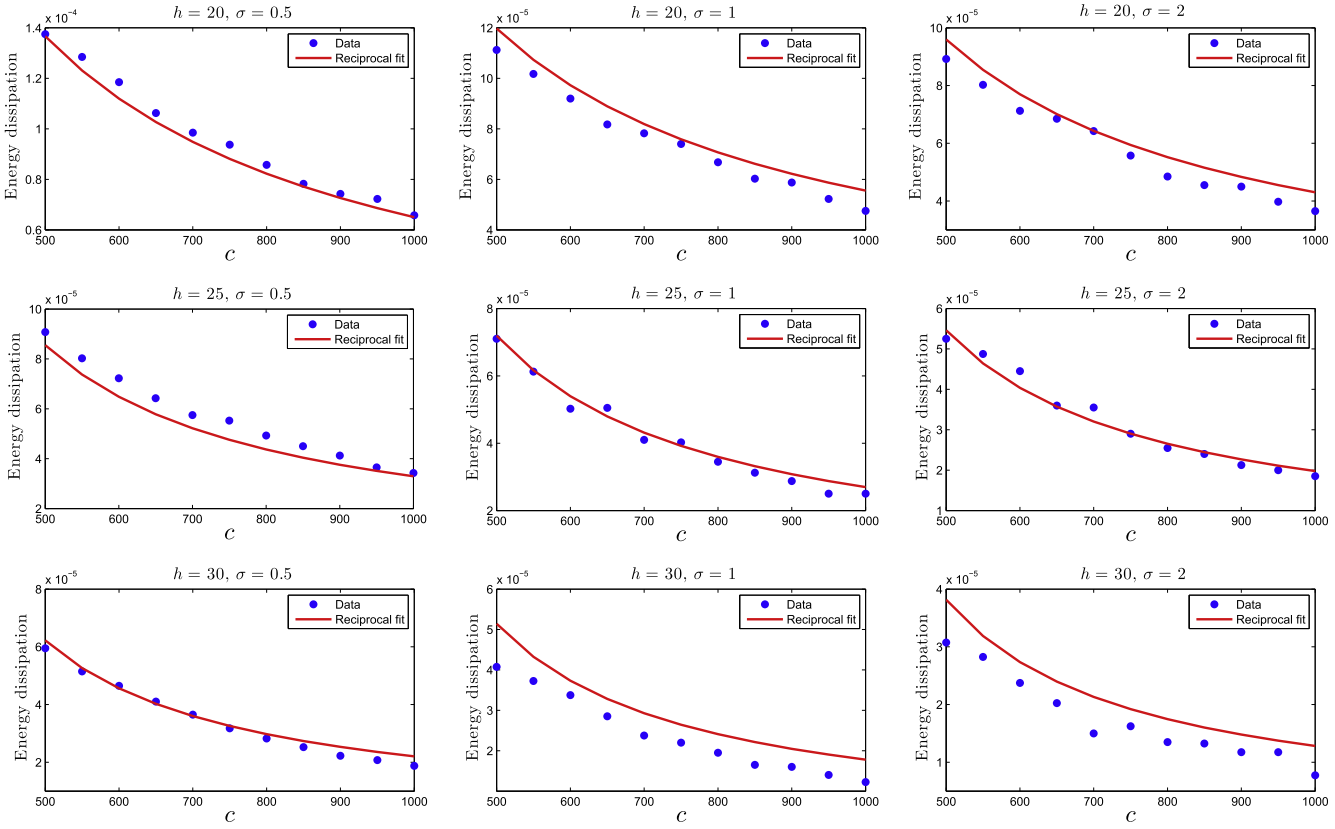

Fig. 4. The training data and fitted results for different values of h, σ and c values.

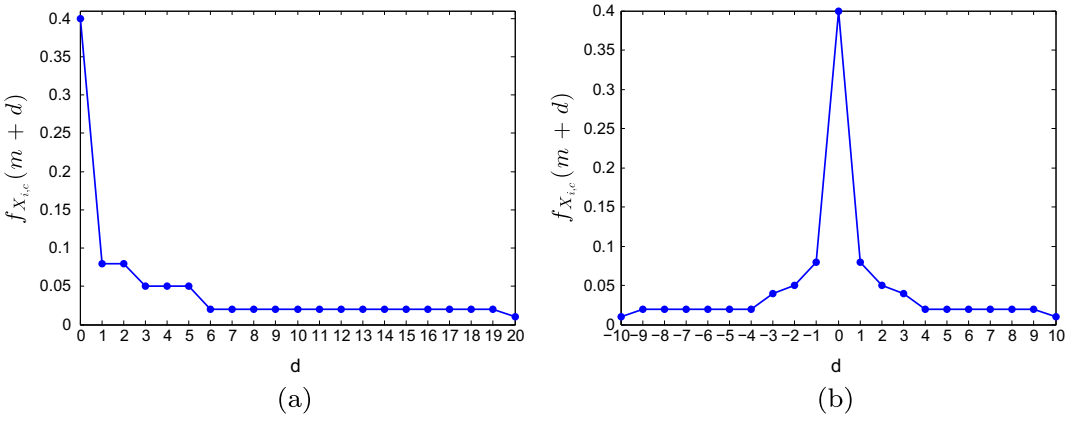

Fig. 5. Two different types of conditional distribution functions used for the outcomes of the measurements based on the energy dissipation value m obtained from the fitted model. d is the distance from the value obtained from the fitted model. (a) Conditional distribution function for $m \leqslant 10$. (b) Conditional distribution function for $m > 10$.

are optimal on average with respect to the prior (current state of knowledge):

$$i^{exploit}, c^{exploit} = \underset{1 \leqslant i \leqslant n,\, c \in C}{\arg\min}\, \mathrm{E}_{h_i,\sigma_i}[g(h_i, \sigma_i, c; \mathbf{b})] \tag{17}$$

We assume that one of the generated samples in the uncertainty class has the underlying true values for the dopant parameters. We perform three different experimental design approaches: proposed method, random selection, and pure exploitation. As explained in the beginning of this section, after each measurement, we report the energy dissipation of the robust material relative to the assumed true parameters. We repeat simulations for 10,000 different assumed true parameters chosen from the generated samples in the uncertainty class to obtain the average results for each experimental design approach.

Fig. 6 shows the average energy dissipation obtained after each simulation for the three selection approaches. We conduct up to 5 simulations for the proposed experimental design method and up to 10 simulations for the pure exploitation and random selection policies. Results are averaged over all 10,000 assumed true models. All three curves begin from the same point, which is the average energy dissipation obtained using the robust material before conducting any simulations. The black dashed line in the figure shows the average energy dissipation of the optimal material, which is for

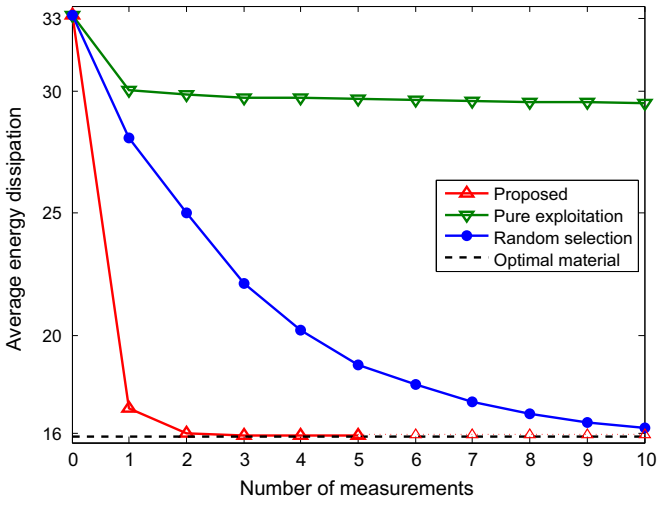

Fig. 6. Average energy dissipation resulting from applying the robust dopant and concentration obtained after different numbers of measurements. We made 5 measurements for the proposed approach and 10 measurements for pure exploitation and random selection policies.

the case that the true model is completely known. A significant difference between the performance of the proposed experimental design method and two other methods is evident in the figure. After only a few simulation runs chosen by the proposed experimental design method, we can achieve the average optimal energy dissipation. Getting optimal results after fewer measurements is especially crucial in materials discovery where measurements are expensive and time consuming.

In Fig. 7, the average probability of the true parameters is shown for different approaches when different numbers of simulations are conducted. Using the proposed method the probability of the true parameters increases much faster than for the two other selection approaches. Comparing this figure and the previous figure shows that, although even after five experiments based on our experimental design method the average probability of the true model is around 0.5, the optimal energy dissipation is almost reached after only two simulations. This is because MOCU measures the pertinent uncertainty that increases the operational cost. Thus, although there is still uncertainty in the model, more simulations are not needed. Another observation from this figure is

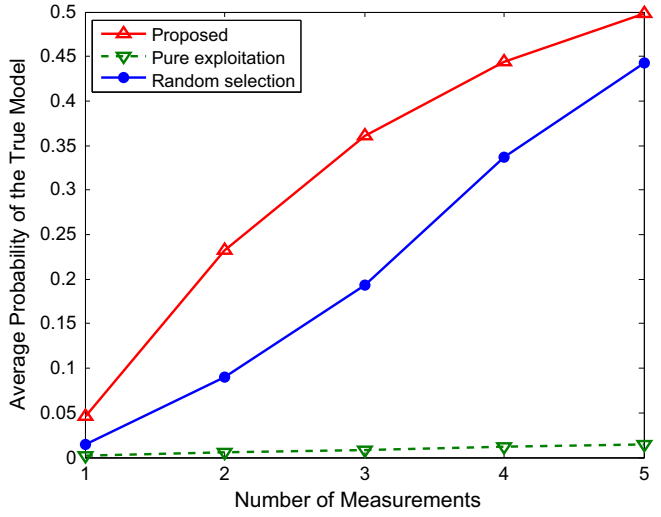

Fig. 7. Average probability of the true model after conducting different numbers of measurements based on different selection policies.

Table 3
The percentage of chosen dopants for each measurement in a sequence of measurements based on the proposed experimental design.

	1st	2nd	3rd	4th	5th
Dopant 1	100	45.8	49.4	52.8	54.6
Dopant 2	0	12.3	7.4	6.1	5.8
Dopant 3	0	8.8	6.6	7.9	6.2
Dopant 4	0	0.7	5	4	3.8
Dopant 5	0	0.3	3	4.6	4.9
Dopant 6	0	20.8	18.4	16.7	15.3
Dopant 7	0	11.3	10.2	7.9	9.4

the poor performance of the pure exploitation approach in terms of gaining knowledge regarding the true model. Pure exploitation aims to work well relative to the prior distribution; however, as it does not produce a promising posterior distribution, which is obvious from Fig. 7, it is not able to improve the performance of the robust material after the simulation. That is why in Fig. 6 random selection outperforms pure exploitation.

Table 3 gives the percentage of the seven possible dopants chosen for each simulation according to the proposed experimental design. We can see that while dopant 1 is always suggested for the first simulation, for the next simulations different dopants are chosen. This is because the first simulation can result in different outcomes, and different outcomes yield different posterior distributions, which in turn affect the decision making process for the consequent measurements.

5. General framework for materials design

Here we outline the general approach we have employed so that it can be used for other problems and illustrate the algorithm via a simple, worked example. Let $\boldsymbol{\theta} = [\theta_1, \theta_2, \ldots, \theta_k]$ be a vector of k uncertain parameters in the model whose true values are unknown. The collection of all possible values for $\boldsymbol{\theta}$ is denoted by Θ and is called the *uncertainty class*. We assume a *prior probability distribution* over Θ with density function $f(\boldsymbol{\theta})$ that incorporates our prior knowledge, perhaps uniform if we have no knowledge regarding the distribution of probability mass over Θ. Let Ψ be the set of all possible operators, which in medicine might consist of a set of possible drug treatments and in materials and engineering a set of design options. The basic idea is that an operator is some transformation of an existing system, which is the ultimate objective of building a model, such as filtering, classification, network intervention, and doping process. $\zeta_{\boldsymbol{\theta}}(\psi)$ denotes the cost of applying operator $\psi \in \Psi$ to the model whose unknown parameters have value $\boldsymbol{\theta}$, where the cost is some measure of performance loss – it was energy dissipation in our previous example. The optimal model-specific operator $\psi(\boldsymbol{\theta})$ has the lowest cost relative to the specific model $\boldsymbol{\theta}$:

$$\psi(\boldsymbol{\theta}) = \arg\min_{\psi \in \Psi} \zeta_{\boldsymbol{\theta}}(\psi). \tag{18}$$

When dealing with uncertainty, the aim is to find the *robust operator* $\psi(\Theta)$ which performs optimally on average across the uncertainty class Θ:

$$\psi(\Theta) = \arg\min_{\psi \in \Psi} \mathrm{E}_{\boldsymbol{\theta}}[\zeta_{\boldsymbol{\theta}}(\psi)] = \arg\min_{\psi \in \Psi} \int \zeta_{\boldsymbol{\theta}}(\psi) f(\boldsymbol{\theta})\, d\boldsymbol{\theta}, \tag{19}$$

where the expectation is taken relative to $f(\boldsymbol{\theta})$. As seen from Eq. (19), $\psi(\Theta)$ has the lowest cost on average but not necessarily for any specific model $\boldsymbol{\theta}$. The *objective cost of uncertainty* relative to model $\boldsymbol{\theta}$ is defined by

$$\mathrm{U}(\boldsymbol{\theta}) = \zeta_{\boldsymbol{\theta}}(\psi(\Theta)) - \zeta_{\boldsymbol{\theta}}(\psi(\boldsymbol{\theta})). \tag{20}$$

$U(\boldsymbol{\theta})$ gives the differential cost of applying the robust operator $\psi(\Theta)$ instead of the optimal operator $\psi(\boldsymbol{\theta})$ to model $\boldsymbol{\theta}$. Taking the expectation of Eq. (20) over $f(\boldsymbol{\theta})$ yields the *mean objective cost of uncertainty* (MOCU),

$$M(\Theta) = E_{\theta}[\zeta_{\boldsymbol{\theta}}(\psi(\Theta)) - \zeta_{\boldsymbol{\theta}}(\psi(\boldsymbol{\theta}))], \tag{21}$$

which provides an objective-based uncertainty quantification [33]. MOCU measures the difference between the average cost of applying the robust operator $\psi(\Theta)$ and the average cost of applying the model-specific optimal operators $\psi(\boldsymbol{\theta})$ across the uncertainty class. The intuition behind MOCU is that it is not overall model uncertainty that is essential, but only that leading to increased operational cost.

Now assume that there are l possible experiments, $X_1, X_2, \ldots, X_l$, and the aim is to find which experiment is better to be conducted first in order to effectively reduce uncertainty. If the outcome of experiment X_i is x_i, then we define the remaining MOCU given the outcome x_i by

$$\begin{aligned} M(\Theta|x_i) &= E_{\theta|X_i=x_i}[\zeta_{\boldsymbol{\theta}}(\psi(\Theta|x_i)) - \zeta_{\boldsymbol{\theta}}(\psi(\boldsymbol{\theta}))] \\ &= \int (\zeta_{\boldsymbol{\theta}}(\psi(\Theta|x_i)) - \zeta_{\boldsymbol{\theta}}(\psi(\boldsymbol{\theta})))f(\boldsymbol{\theta}|X_i = x_i)\, d\boldsymbol{\theta}, \end{aligned} \tag{22}$$

where the expectation is taken over the conditional distribution $f(\boldsymbol{\theta}|X_i = x_i)$ and $\psi(\Theta|x_i)$ is the robust operator when the outcome x_i is observed which can be found similarly to Eq. (19) as

$$\psi(\Theta|x_i) = \arg\min_{\psi\in\Psi} E_{\theta|X_i=x_i}[\zeta_{\boldsymbol{\theta}}(\psi)]. \tag{23}$$

It should be noted that the relationship between parameter $\boldsymbol{\theta}$ and the experimental outcome $X_i = x_i$ is often available in the form of the conditional distribution $f_{X_i}(x_i|\boldsymbol{\theta})$. Therefore, we should obtain $f(\boldsymbol{\theta}|X_i = x_i)$ using Bayes' rule from $f_{X_i}(x_i|\boldsymbol{\theta})$. Also, the marginal distribution $f_{X_i}(x_i)$ of the outcomes of the experiment X_i can be obtained as $E_{\boldsymbol{\theta}}\left[f_{X_i}(x_i|\boldsymbol{\theta})\right]$.

Taking the expectation of Eq. (22) with respect to the probability distribution $f_{X_i}(x_i)$ of the experiment outcomes yields

$$M(\Theta; i) = E_{x_i}\left[E_{\theta|X_i=x_i}[\zeta_{\boldsymbol{\theta}}(\psi(\Theta|x_i)) - \zeta_{\boldsymbol{\theta}}(\psi(\boldsymbol{\theta}))]\right], \tag{24}$$

which is referred to as the expected remaining MOCU given that experiment X_i has been conducted.

For experimental design, we define the *optimal experiment* X_{i^*}, the one that should be conducted first, as the experiment yielding the minimum expected remaining MOCU:

$$i^* = \arg\min_{i\in 1,\ldots,l} M(\Theta; i). \tag{25}$$

Potential experiments can be prioritized based on the value of their corresponding expected remaining MOCU, i.e., if $M(\Theta; 1') < M(\Theta; 2') < \cdots < M(\Theta; l')$, we can have $X_{1'} \prec X_{2'} \prec \cdots \prec X_{l'}$, meaning that the best experiment to be conducted first is $X_{1'}$; if conducting $X_{1'}$ is not possible, then $X_{2'}$ should be conducted, and so on. The minimization problem in Eq. (25) can be further simplified:

$$\begin{aligned} i^* &= \arg\min_{i\in 1,\ldots,l} M(\Theta; i) \\ &= \arg\min_{i\in 1,\ldots,l} E_{x_i}\left[E_{\theta|X_i=x_i}[\zeta_{\boldsymbol{\theta}}(\psi(\Theta|x_i)) - \zeta_{\boldsymbol{\theta}}(\psi(\boldsymbol{\theta}))]\right] \\ &= \arg\min_{i\in 1,\ldots,l} \left\{E_{x_i}\left[E_{\theta|X_i=x_i}[\zeta_{\boldsymbol{\theta}}(\psi(\Theta|x_i))]\right] - E_{x_i}\left[E_{\theta|X_i=x_i}[\zeta_{\boldsymbol{\theta}}(\psi(\boldsymbol{\theta}))]\right]\right\} \\ &= \arg\min_{i\in 1,\ldots,l} \left\{E_{x_i}\left[E_{\theta|X_i=x_i}[\zeta_{\boldsymbol{\theta}}(\psi(\Theta|x_i))]\right] - E_{\boldsymbol{\theta}}[\zeta_{\boldsymbol{\theta}}(\psi(\boldsymbol{\theta}))]\right\} \\ &= \arg\min_{i\in 1,\ldots,l} E_{x_i}\left[E_{\theta|X_i=x_i}[\zeta_{\boldsymbol{\theta}}(\psi(\Theta|x_i))]\right], \end{aligned} \tag{26}$$

where the fourth line follows because $\zeta_{\boldsymbol{\theta}}(\psi(\boldsymbol{\theta}))$ is not a function of x_i and therefore we just need to take the expectation with respect to $f(\boldsymbol{\theta})$, and the last line is obtained because the second term in the fourth line is not a function of the optimization parameter. Note that Eq. (8) used in our specific experimental design method for achieving low energy dissipation has been obtained due to Eq. (26).

Algorithm 2 provides a summary of the proposed experimental design method in its general form. For the proposed experimental design framework, we need to define a prior probability distribution $f(\boldsymbol{\theta})$ governing the uncertainty class Θ, a set of all possible operators Ψ, a set of possible experiments $X_i, 1 \leqslant i \leqslant l$, and a conditional distribution of the experimental outcomes $f_{X_i}(x_i|\boldsymbol{\theta})$ to be able to find the best experiment X_{i^*}.

Algorithm 2. Proposed experimental design method

1: **input**: $f(\boldsymbol{\theta}), \Psi, \{X_1, \ldots, X_l\}, f_{X_i}(x_i|\boldsymbol{\theta})$
2: **output**: Optimal experiment X_{i^*}
3: **for all** $X_i, 1 \leqslant i \leqslant l$ **do**
4: **for all** possible outcomes x_i of X_i **do**
5: Compute conditional expectation $f(\boldsymbol{\theta}|X_i = x_i)$ using Bayes' rule
6: Find the robust operator $\psi(\Theta|X_i = x_i)$ using Eq. (23)
7: $\omega(x_i) \leftarrow E_{\theta|X_i=x_i}[\zeta_{\boldsymbol{\theta}}(\psi(\Theta|X_i = x_i))]$ (obtained from step 6)
8: Compute $E_{x_i}[\omega(x_i)]$
9: $i^* \leftarrow \arg\min_{i=1,\ldots,l} E_{x_i}[\omega(x_i)]$ (optimization in Eq. (26))
10: **return** i^* (optimal experiment X_{i^*})

5.1. Illustrative example of Algorithm 2 with a hypothetical cost function

To clarify the above, consider the problem of minimizing the cost function $\zeta_{\theta}(\psi) = \max(\psi^2 + 2\theta\psi + 10, 0)$, where ψ plays the role of the operator and the dopant variables i and c in our previous dissipation problem also map to ψ. The one unknown parameter is θ and assume probabilities $f(\theta = 1) = 1/3$ and $f(\theta = 3) = 2/3$ corresponding to the prior distribution $f(\theta)$. Then the model-specific optimal operator $\psi(\theta)$ (Eq. (18)) has values $\psi(\theta = 1) = -1$ and $\psi(\theta = 3) = -3$ with corresponding optimal costs $\zeta_{\theta=1}(\psi(\theta = 1)) = 9$ and $\zeta_{\theta=3}(\psi(\theta = 3)) = 1$. The robust operator from Eq. (19) is then $\Psi(\Theta) = \arg\min_{\psi\in\Psi} E_{\theta}[\zeta_{\theta}(\psi)] = -7/3$.

Suppose now we have two hypothetical experiments X_1 and X_2 to be carried out to determine the parameter θ: (a) Experiment X_1 (whose outcome is denoted by x_1) with conditional probabilities $f_{X_1}(1|\theta = 1) = 0.9$ and $f_{X_1}(3|\theta = 3) = 0.7$ ($f_{X_1}(3|\theta = 1) = 0.1$ and $f_{X_1}(1|\theta = 3) = 0.3$), and (b) Experiment X_2 (whose outcome is denoted by x_2) with conditional probabilities $f_{X_2}(1|\theta = 1) = 0.8$ and $f_{X_2}(3|\theta = 3) = 0.9$. Hence $f_{X_2}(3|\theta = 1) = 0.2$ and $f_{X_2}(1|\theta = 3) = 0.1$.

As per our Algorithm 2 (step 4), for experimental design we need to take into account each possible outcome and for this we need to find the probability distribution $f_{X_i}(x_i)$. Using the above, we then have $f_{X_1}(1) = 0.5$ (thus $f_{X_1}(3) = 0.5$), and $f_{X_2}(1) = 0.333$ (thus $f_{X_2}(3) = 0.667$).

Step 5 of Algorithm 2 requires the conditional probability $f(\theta|X_i = x_i)$ for the unknown parameter θ for each of the outcomes. We can compute this using Bayes' rule and f_{X_1}, f_{X_2}. That is, for experiment X_1,

$$f(\theta = 1|X_1 = 1) = \frac{f_{X_1}(1|\theta = 1)f(\theta = 1)}{f_{X_1}(1)} = \frac{3}{5}.$$

Similarly, $f(\theta = 1|X_1 = 3) = \frac{1}{15}$. For experiment X_2, we obtain $f(\theta = 1|X_2 = 1) = 0.8$ and $f(\theta = 1|X_2 = 3) = 0.1$.

In step 6, we need to calculate the robust operator $\psi(\Theta|X_i = x_i)$ for each outcome of X_1 and X_2. This leads to $\psi(\Theta|X_1 = 1) = \arg\min_{\psi\in\Psi} \mathrm{E}_{\theta|X_1=1}[\zeta_\theta(\psi)] = -9/5, \psi(\Theta|X_1 = 3) = -43/15$, $\psi(\Theta|X_2 = 1) = -1.4$ and $\psi(\Theta|X_2 = 3) = -2.8$.

In order to calculate Eq. (26), we first carry out step 7 which involves taking the expectation with respect to the conditional probability distribution $f(\theta|X_i = x_i)$. This gives $\omega(X_1 = 1) = \mathrm{E}_{\theta|X_1=1}[\zeta_\theta(\psi(\Theta|X_1 = 1))] = 6.76$. Similarly, $\omega(X_1 = 3) = \mathrm{E}_{\theta|X_1=3}[\zeta_\theta(\psi(\Theta|X_1 = 3))] = 1.78$ and $\omega(X_2 = 1) = 8.04$, $\omega(X_2 = 3) = 2.16$.

Finally, step 8 gives us the value of the expression being minimized in Eq. (26) for each experiment. We obtain $\mathrm{E}_{x_1}[\omega(x_1)] = 4.27$ and $\mathrm{E}_{x_2}[\omega(x_2)] = 4.12$. So that according to step 9 as $\mathrm{E}_{x_2}[\omega(x_2)] = 4.12 < 4.27 = \mathrm{E}_{x_1}[\omega(x_1)]$, and the optimal experiment that should be conducted first is X_2.

5.2. Illustrative example of Algorithm 2 with three particles or switches and two states with well defined cost function

Consider the toy network in Fig. 8 composed of three binary switches, S_1, S_2, and S_3, and suppose it is part of a large system. The state of the network is determined by the state vector $[S_1, S_2, S_3]$. Suppose that when switch $S_3 = 1$ (ON), a signal is sent to activate various components of the system and that under present conditions this is undesirable, for instance, it can lead to system failure. Therefore, we desire an intervention for this network to prevent states with $S_3 = 1$. In order to do so, we can block or delete one of the edges or regulators in the network. Therefore, we have two different options for intervention: ψ_1 blocking edge θ_1 and ψ_2 blocking edge θ_2, i.e., $\Psi = \{\psi_1, \psi_2\}$. The cost after applying the intervention is the probability of having $S_3 = 1$, $\zeta(\psi) = f_{S_3}(1)$, $f_{S_3}(1)$ being the probability that $S_3 = 1$. This is precisely the kind of situation one faces in genomics where the switches are genes and $S_3 = 1$ initiates a cascade of downstream genes whose activation is associated with an undesirable phenotype, such as cancer [43]. The experimental design question is then which of two experiments X_1 and X_2 determining θ_1 and θ_2 values for the edges or regulators, respectively, should be carried out first to minimize the cost of the intervention. In an analogy with the materials problem, the switches can represent occupation of different elements or structural instabilities. For example, if we consider ferroelectric ABO_3 perovskites, the binary switches S_1 and S_2 can be the occupation of Ba or Ca at the A site and Ti or Zr at the B site, and they can regulate the outcome at S_3, which can be a ferroelectric instability arising from some external stimuli such as the temperature.

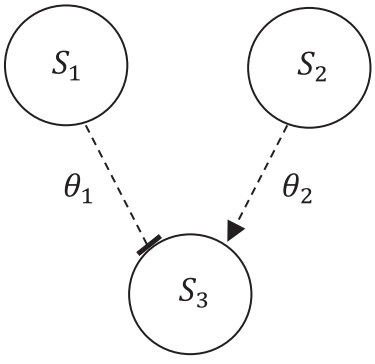

Fig. 8. The toy network used for illustrative example.

Suppose the state of S_3 is determined from S_1 and S_2 by evaluating the following expression:

$$R(S_3) = \theta_1 S_1 + \theta_2 S_2.$$

Using $R(S_3)$, we define:

$$S_3 = \begin{cases} 1 & \text{if } R(S_3) > 0 \\ 0 & \text{if } R(S_3) < 0 \\ 1(\text{with probability } 0.5) & \text{if } R(S_3) = 0. \end{cases}$$

We substitute 1 for edge θ_i in $R(S_3)$ if it is shown by a normal arrow (an activating edge), and -1 if it is shown by a blunt arrow (a suppressive edge). Now assume that θ_1 and θ_2 are unknown. Each uncertain parameter (edge) can take two values: 1 and -1. The uncertainty class Θ, shown in Fig. 9, contains four networks:

$$\Theta = \{\boldsymbol{\theta}^1 : (\theta_1 = -1, \theta_2 = -1); \boldsymbol{\theta}^2 : (\theta_1 = -1, \theta_2 = 1); \boldsymbol{\theta}^3 : (\theta_1 = 1, \theta_2 = -1); \boldsymbol{\theta}^4 : (\theta_1 = 1, \theta_2 = 1)\}.$$

We assume the following probabilities for S_1 and S_2:

$$\begin{cases} f_{S_1,S_2}(0,0) = 0.1 \\ f_{S_1,S_2}(0,1) = 0.2 \\ f_{S_1,S_2}(1,0) = 0.3 \\ f_{S_1,S_2}(1,1) = 0.4 \end{cases}.$$

We also consider the following prior distribution for the unknown edges θ_1 and θ_2:

$$\begin{cases} f(\boldsymbol{\theta}^1) = f_{\theta_1,\theta_2}(-1,-1) = 0.35 \\ f(\boldsymbol{\theta}^2) = f_{\theta_1,\theta_2}(-1,1) = 0.3 \\ f(\boldsymbol{\theta}^3) = f_{\theta_1,\theta_2}(1,-1) = 0.25 \\ f(\boldsymbol{\theta}^4) = f_{\theta_1,\theta_2}(1,1) = 0.1 \end{cases}.$$

We find the cost of applying each intervention in Ψ to each network inside the uncertainty class. Considering $\boldsymbol{\theta}^1$, after applying intervention ψ_1 that blocks regulation θ_1, we have $R(S_3) = -S_2$. Therefore, when $S_2 = 1$, S_3 becomes 0, and when $S_2 = 0$, S_3 becomes 1 with probability 0.5. Hence,

$$\begin{aligned} \zeta_{\boldsymbol{\theta}^1}(\psi_1) &= f_{S_3}(1) = 0.5 \times f_{S_2}(0) \\ &= 0.5 \times \left(f_{S_1,S_2}(0,0) + f_{S_1,S_2}(1,0)\right) = 0.2. \end{aligned}$$

Similarly regarding $\zeta_{\boldsymbol{\theta}^1}(\psi_2)$, we have $R(S_3) = -S_1$. Therefore,

$$\begin{aligned} \zeta_{\boldsymbol{\theta}^1}(\psi_2) &= 0.5 \times f_{S_1}(0) \\ &= 0.5 \times \left(f_{S_1,S_2}(0,0) + f_{S_1,S_2}(0,1)\right) = 0.15. \end{aligned}$$

Therefore, the optimal intervention $\psi(\boldsymbol{\theta}^1)$ for network $\boldsymbol{\theta}^1$ is ψ_2. The optimal interventions for other networks in the uncertainty class can be found similarly:

$$\boldsymbol{\theta}^2 : \begin{cases} \zeta_{\boldsymbol{\theta}^2}(\psi_1) = f_{S_2}(1) + 0.5 \times f_{S_2}(0) = 0.8 \\ \zeta_{\boldsymbol{\theta}^2}(\psi_2) = 0.5 \times f_{S_1}(0) = 0.15 \end{cases} \Rightarrow \psi(\boldsymbol{\theta}^2) = \psi_2,$$

$$\boldsymbol{\theta}^3 : \begin{cases} \zeta_{\boldsymbol{\theta}^3}(\psi_1) = 0.5 \times f_{S_2}(0) = 0.2 \\ \zeta_{\boldsymbol{\theta}^3}(\psi_2) = f_{S_1}(1) + 0.5 \times f_{S_1}(0) = 0.85 \end{cases} \Rightarrow \psi(\boldsymbol{\theta}^3) = \psi_1,$$

$$\boldsymbol{\theta}^4 : \begin{cases} \zeta_{\boldsymbol{\theta}^4}(\psi_1) = f_{S_2}(1) + 0.5 \times f_{S_2}(0) = 0.8 \\ \zeta_{\boldsymbol{\theta}^4}(\psi_2) = f_{S_1}(1) + 0.5 \times f_{S_1}(0) = 0.85 \end{cases} \Rightarrow \psi(\boldsymbol{\theta}^4) = \psi_1.$$

To find the robust intervention $\psi(\Theta)$, we need to find the average cost of each intervention across the uncertainty class (Eq. (19)). It can be seen that $\mathrm{E}_\theta[\zeta_\theta(\psi_1)] = \sum_{i=1}^4 f(\boldsymbol{\theta}^i)\zeta_{\boldsymbol{\theta}^i}(\psi_1) = 0.44$ and $\mathrm{E}_\theta[\zeta_\theta(\psi_2)] = 0.3950$. Hence, $\psi(\Theta) = \psi_2$. Now that the robust intervention $\psi(\Theta)$ and the optimal interventions $\psi(\boldsymbol{\theta}^i)$ have been derived, we can compute the objective cost of uncertainty relative

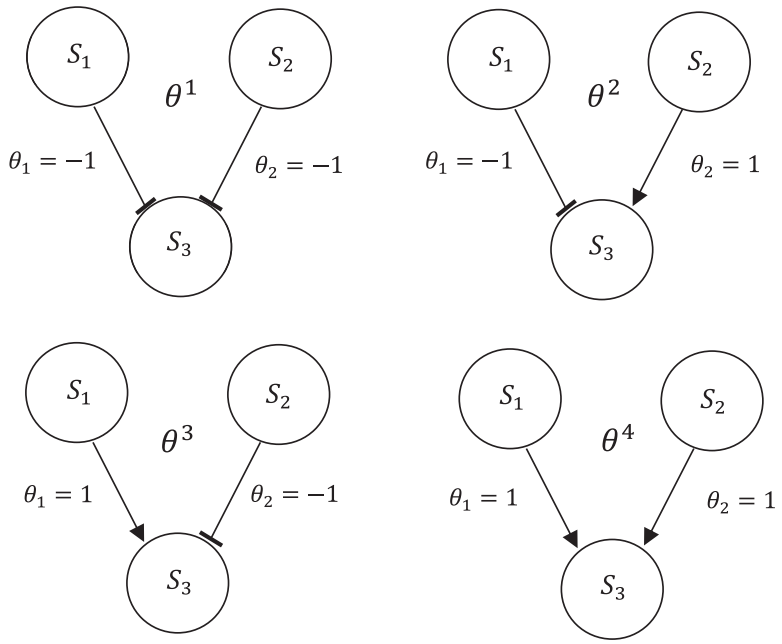

Fig. 9. The uncertainty class Θ for three binary switches S_1, S_2, S_3, which contains all possible networks.

to each $\boldsymbol{\theta}^i$ according to Eq. (20). For example, $U(\boldsymbol{\theta}^1) = 0$ or $U(\boldsymbol{\theta}^3) = 0.65$. Also, using Eq. (21), MOCU is obtained as: $M(\Theta) = E_{\boldsymbol{\theta}}[U(\boldsymbol{\theta})] = 0.1675$.

Now assume that there are two experiments X_1 and X_2 determining θ_1 and θ_2, respectively. The aim of experimental design is to find out which experiment is better to be carried out first. Suppose that the following conditional probabilities regarding the experiment outcomes are given:

$$\begin{cases} f_{X_1}(-1|\theta_1 = -1) = 0.6 \\ f_{X_1}(1|\theta_1 = 1) = 0.8 \\ f_{X_2}(-1|\theta_2 = -1) = 0.8 \\ f_{X_2}(1|\theta_2 = 1) = 0.6 \end{cases}.$$

One can verify that $f_{X_1}(1) = 0.54$ and $f_{X_2}(1) = 0.36$.

In order to carry out step 5 of Algorithm 2, we need to compute the conditional probabilities of the unknown regulations for each possible experiment outcome. For $X_1 = -1$:

$$f_{\theta_1,\theta_2}(-1,-1|X_1 = -1) = \frac{f_{X_1}(-1|\theta_1 = -1, \theta_2 = -1) f_{\theta_1,\theta_2}(-1,-1)}{f_{X_1}(-1)} = 0.4565$$
$$f_{\theta_1,\theta_2}(-1,1|X_1 = -1) = 0.3913$$
$$f_{\theta_1,\theta_2}(1,-1|X_1 = -1) = 0.1087$$
$$f_{\theta_1,\theta_2}(1,1|X_1 = -1) = 0.0435.$$

Note that $f_{X_1}(-1|\theta_1 = -1, \theta_2 = -1) = f_{X_1}(-1|\theta_1 = -1)$. Also, for $X_1 = 1$ we have:

$$f_{\theta_1,\theta_2}(-1,-1|X_1 = 1) = 0.2593$$
$$f_{\theta_1,\theta_2}(-1,1|X_1 = 1) = 0.2222$$
$$f_{\theta_1,\theta_2}(1,-1|X_1 = 1) = 0.3704$$
$$f_{\theta_1,\theta_2}(1,1|X_1 = 1) = 0.1481.$$

We also find the conditional probabilities for $X_2 = -1$, $f_{\theta_1,\theta_2}(\pm 1, \pm 1|X_2 = -1)$, and for $X_2 = 1$, $f_{\theta_1,\theta_2}(\pm 1, \pm 1|X_2 = 1)$.

According to step 6 of the algorithm, we need to find the robust operator $\psi(\Theta|X_i = x_i)$ for each possible outcome. To find $\psi(\Theta|X_1 = -1)$, first we need to find $E_{\theta|X_1=-1}[\zeta_\theta(\psi)]$ for each intervention. Considering intervention ψ_1 we have:

$$\begin{aligned} E_{\theta|X_1=-1}[\zeta_\theta(\psi_1)] = {} & f_{\theta_1,\theta_2}(-1,-1|X_1 = -1)\zeta_{\boldsymbol{\theta}^1}(\psi_1) \\ & + f_{\theta_1,\theta_2}(-1,1|X_1 = -1)\zeta_{\boldsymbol{\theta}^2}(\psi_1) \\ & + f_{\theta_1,\theta_2}(1,-1|X_1 = -1)\zeta_{\boldsymbol{\theta}^3}(\psi_1) \\ & + f_{\theta_1,\theta_2}(1,1|X_1 = -1)\zeta_{\boldsymbol{\theta}^4}(\psi_1) \\ = {} & 0.4609. \end{aligned}$$

Similarly, we have $E_{\theta|X_1=-1}[\zeta_\theta(\psi_2)] = 0.2565$. Therefore, $\psi(\Theta|X_1 = -1) = \psi_2$. We also find the robust interventions for the other outcomes: $\psi(\Theta|X_1 = 1) = \psi_1$ with $E_{\theta|X_1=1}[\zeta_\theta(\psi_1)] = 0.4222$, $\psi(\Theta|X_2 = -1) = \psi_1$ with $E_{\theta|X_2=-1}[\zeta_\theta(\psi_1)] = 0.35$, and $\psi(\Theta|X_2 = 1) = \psi_2$ with $E_{\theta|X_2=1}[\zeta_\theta(\psi_2)] = 0.3638$. The next step is to implement step 7 of the algorithm for each possible outcome x_i : $\omega(X_1 = -1) = E_{\theta|X_1=-1}[\zeta_\theta(\psi_2)] = 0.2565$, $\omega(X_1 = 1) = 0.4222$, $\omega(X_2 = -1) = 0.35$, and $\omega(X_2 = 1) = 0.3638$.

Corresponding to step 8 of the algorithm, we compute the following quantities for each experiment: $E_{x_1}[\omega(x_1)] = \sum_{x_1 \in \{-1,1\}} f_{X_1}(x_1)\omega(X_1 = x_1) = 0.3460$ and similarly $E_{x_2}[\omega(x_2)] = 0.3550$. Finally, based on step 9, as $E_{x_1}[\omega(x_1)] = 0.3460 < 0.3550 = E_{x_2}[\omega(x_2)]$. Thus, the optimal experiment which needs to be carried out first is X_1 to determine θ_1. We emphasize that experimental design determines which experiment, X_1 (aimed at estimating θ_1) or X_2(aimed at estimating θ_2) is better to be carried out first, but the criterion for choosing the first experiment is that which is more likely to yield a better intervention (deleting edge). As we do not know the outcomes ahead of time, we do not know which intervention will result after conducting each experiment. Therefore, we take a probabilistic approach to finding the experiment resulting in an intervention with lower cost.

6. Conclusion

Experimental design has the potential to play a major role in materials science as one of the challenges is to find materials with targeted properties. This is especially the case as computational models typically have considerable uncertainty and experimental

measurements can be expensive and time consuming, In this paper, we propose a general experimental design framework which can be used for a vast majority of materials design and discovery problems. The method is based on the concept of mean objective cost of uncertainty, which captures the pertinent uncertainty in the model. Provided we have a mathematical model, can assign appropriate prior distributions to the uncertainties in the model, and define the objective from the model, the proposed approach, owing to optimization, is guaranteed to achieve on average the best results given the prior distribution and the data.

We have illustrated our approach using a Landau mesoscale model for SMAs with the objective of minimizing the number of "experiments" required to find the material features leading to the lowest dissipation. We find that our MOCU design strategy, which strives to minimize the uncertainty in the model pertaining to the design objective, leads to the features with optimal dissipation after only two experiments. The pure exploitation approach, which is customarily used in materials science, performs relatively poorly as it does not produce a promising posterior distribution.

Acknowledgements

The authors thank the Laboratory Directed Research and Development (LDRD) program at Los Alamos National Laboratory (project number 20140013DR) for support.

References

[1] White House Office of Science and Technology Policy, Materials Genome Initiative for Global Competitiveness, 2011. <https://www.whitehouse.gov/sites/default/files/microsites/ostp/materials_genome_initiative-final.pdf>.
[2] T. Lookman, P. Balachandran, D. Xue, G. Pilania, T. Shearman, J. Theiler, J. Gubernatis, J. Hogden, K. Barros, E. BenNaim, et al., A perspective on materials informatics: state-of-the-art and challenges, in: Information Science for Materials Discovery and Design, Springer, 2016, pp. 3–12.
[3] S. Curtarolo, G.L. Hart, M.B. Nardelli, N. Mingo, S. Sanvito, O. Levy, The high-throughput highway to computational materials design, Nat. Mater. 12 (3) (2013) 191–201.
[4] R. Gautier, X. Zhang, L. Hu, L. Yu, Y. Lin, T.O. Sunde, D. Chon, K.R. Poeppelmeier, A. Zunger, Prediction and accelerated laboratory discovery of previously unknown 18-electron ABX compounds, Nat. Chem. 7 (4) (2015) 308–316.
[5] H. Koinuma, I. Takeuchi, Combinatorial solid-state chemistry of inorganic materials, Nat. Mater. 3 (2004) 429–438, http://dx.doi.org/10.1038/nmat1157.
[6] D. Xue, P.V. Balachandran, J. Hogden, J. Theiler, D. Xue, T. Lookman, Accelerated search for materials with targeted properties by adaptive design, Nat. Commun. 7 (2016) 11241.
[7] C. Kim, G. Pilania, R. Ramprasad, From organized high-throughput data to phenomenological theory using machine learning: the example of dielectric breakdown, Chem. Mater. 28 (5) (2016) 1304–1311.
[8] P.V. Balachandran, S.R. Broderick, K. Rajan, Identifying the "inorganic gene" for high–temperature piezoelectric perovskites through statistical learning, Proc. Roy. Soc. A: Math. Phys. Eng. Sci. 467 (2132) (2011) 2271–2290, http://dx.doi.org/10.1098/rspa.2010.0543.
[9] P.V. Balachandran, J. Theiler, J.M. Rondinelli, T. Lookman, Materials prediction via classification learning, Sci. Rep. 5 (2015) 13285.
[10] T.D. Sparks, M.W. Gaultois, A. Oliynyk, J. Brgoch, B. Meredig, Data mining our way to the next generation of thermoelectrics, Scripta Mater. 111 (2016) 10–15.
[11] A. Seko, T. Maekawa, K. Tsuda, I. Tanaka, Machine learning with systematic density-functional theory calculations: application to melting temperatures of single- and binary-component solids, Phys. Rev. B 89 (2014) 054303, http://dx.doi.org/10.1103/PhysRevB.89.054303.
[12] O. Isayev, D. Fourches, E.N. Muratov, C. Oses, K. Rasch, A. Tropsha, S. Curtarolo, Materials cartography: representing and mining materials space using structural and electronic fingerprints, Chem. Mater. 27 (3) (2015) 735–743, http://dx.doi.org/10.1021/cm503507h.
[13] B. Meredig, A. Agrawal, S. Kirklin, J.E. Saal, J.W. Doak, A. Thompson, K. Zhang, A. Choudhary, C. Wolverton, Combinatorial screening for new materials in unconstrained composition space with machine learning, Phys. Rev. B 89 (2014) 094104.
[14] A. Seko, A. Togo, H. Hayashi, K. Tsuda, L. Chaput, I. Tanaka, Prediction of low-thermal-conductivity compounds with first-principles anharmonic lattice-dynamics calculations and Bayesian optimization, Phys. Rev. Lett. 115 (2015) 205901.
[15] K. Rajan, Materials informatics: the materials "gene" and big data, Annu. Rev. Mater. Res. 45 (1) (2015) 153–169.
[16] P.V. Balachandran, D. Xue, J. Theiler, J. Hogden, T. Lookman, Adaptive strategies for materials design using uncertainties, Sci. Rep. 6 (2016) 19660.
[17] P.V. Balachandran, D. Xue, T. Lookman, Structure-Curie temperature relationships in $BaTiO_3$-based ferroelectric perovskites: anomalous behavior of $(Ba,Cd)TiO_3$ from DFT, statistical inference, and experiments, Phys. Rev. B 93 (2016) 144111, http://dx.doi.org/10.1103/PhysRevB.93.144111.
[18] L.M. Ghiringhelli, J. Vybiral, S.V. Levchenko, C. Draxl, M. Scheffler, Big data of materials science: critical role of the descriptor, Phys. Rev. Lett. 114 (2015) 105503, http://dx.doi.org/10.1103/PhysRevLett.114.105503.
[19] S. Broderick, K. Rajan, Informatics derived materials databases for multifunctional properties, Sci. Technol. Adv. Mater. 16 (1) (2015) 013501.
[20] M.S. Esfahani, E.R. Dougherty, An optimization-based framework for the transformation of incomplete biological knowledge into a probabilistic structure and its application to the utilization of gene/protein signaling pathways in discrete phenotype classification, IEEE/ACM Trans. Comput. Biol. Bioinf. 12 (6) (2015) 1304–1321.
[21] M.S. Esfahani, E.R. Dougherty, Incorporation of biological pathway knowledge in the construction of priors for optimal Bayesian classification, IEEE/ACM Trans. Comput. Biol. Bioinf. 11 (1) (2014) 202–218.
[22] L.A. Dalton, E.R. Dougherty, Bayesian minimum mean-square error estimation for classification error—Part I: definition and the Bayesian MMSE error estimator for discrete classification, IEEE Trans. Signal Process. 59 (1) (2011) 115–129.
[23] L.A. Dalton, E.R. Dougherty, Bayesian minimum mean-square error estimation for classification error—Part II: linear classification of Gaussian models, IEEE Trans. Signal Process. 59 (1) (2011) 130–144.
[24] L.A. Dalton, E.R. Dougherty, Optimal classifiers with minimum expected error within a Bayesian framework—Part I: discrete and Gaussian models, Pattern Recogn. 46 (5) (2013) 1288–1300.
[25] L.A. Dalton, E.R. Dougherty, Optimal classifiers with minimum expected error within a Bayesian framework—Part II: properties and performance analysis, Pattern Recogn. 46 (5) (2013) 1288–1300.
[26] T.D. Huan, A. Mannodi-Kanakkithodi, R. Ramprasad, Accelerated materials property predictions and design using motif-based fingerprints, Phys. Rev. B 92 (1) (2015) 014106.
[27] A. Jain, S.P. Ong, G. Hautier, W. Chen, W.D. Richards, S. Dacek, S. Cholia, D. Gunter, D. Skinner, G. Ceder, et al., Commentary: the materials project: a materials genome approach to accelerating materials innovation, Apl Mater. 1 (1) (2013) 011002.
[28] D. Xue, Y. Zhou, X. Ding, T. Lookman, J. Sun, X. Ren, Aging and deaging effects in shape memory alloys, Phys. Rev. B 86 (18) (2012) 184109.
[29] L. Brinson, R. Lammering, Finite element analysis of the behavior of shape memory alloys and their applications, Int. J. Solids Struct. 30 (23) (1993) 3261–3280.
[30] F. Auricchio, R.L. Taylor, J. Lubliner, Shape-memory alloys: macromodelling and numerical simulations of the superelastic behavior, Comput. Methods Appl. Mech. Eng. 146 (3) (1997) 281–312.
[31] J. Van Humbeeck, S. Kustov, Active and passive damping of noise and vibrations through shape memory alloys: applications and mechanisms, Smart Mater. Struct. 14 (5) (2005) S171.
[32] R. Wang, C. Cho, C. Kim, Q. Pan, A proposed phenomenological model for shape memory alloys, Smart Mater. Struct. 15 (2) (2006) 393.
[33] B.-J. Yoon, X. Qian, E. Dougherty, Quantifying the objective cost of uncertainty in complex dynamical systems, IEEE Trans. Signal Process. 61 (9) (2013) 2256–2266.
[34] R. Dehghannasiri, B.-J. Yoon, E. Dougherty, Optimal experimental design for gene regulatory networks in the presence of uncertainty, IEEE/ACM Trans. Comput. Biol. Bioinf. 12 (4) (2015) 938–950.
[35] R. Dehghannasiri, B.-J. Yoon, E.R. Dougherty, Efficient experimental design for uncertainty reduction in gene regulatory networks, BMC Bioinform. 16 (Suppl 13) (2015) S2.
[36] D. Mohsenizadeh, R. Dehghannasiri, E. Dougherty, Optimal objective-based experimental design for uncertain dynamical gene networks with experimental error, IEEE/ACM Trans. Comput. Biol. Bioinf. (2016).
[37] I. Shmulevich, E.R. Dougherty, S. Kim, W. Zhang, Probabilistic boolean networks: a rule-based uncertainty model for gene regulatory networks, Bioinformatics 18 (2) (2002) 261–274.
[38] A.C. Atkinson, A.N. Donev, Optimum Experimental Designs, Oxford University Press, Oxford, 1992.
[39] H. Raiffa, R. Schlaifer, Applied Statistical Decision Theory, Division of Research, Graduate School of Business Administration, Harvard University, 1961.
[40] V.V. Fedorov, Theory Of Optimal Experiments, Elsevier Science, 1972.
[41] E.T. Jaynes, Prior probabilities, IEEE Trans. Syst. Sci. Cyber. 4 (3) (1968) 227–241.
[42] K.W. Ng, G.-L. Tian, M.-L. Tang, Dirichlet and Related Distributions: Theory, Methods and Applications, vol. 888, John Wiley & Sons, 2011.
[43] X. Qian, E.R. Dougherty, Effect of function perturbation on the steady-state distribution of genetic regulatory networks: optimal structural intervention, IEEE Trans. Signal Process. 56 (10) (2008) 4966–4976.