Compressive sensing as a paradigm for building physics models

Lance J. Nelson and Gus L. W. Hart

Department of Physics and Astronomy, Brigham Young University, Provo, Utah 84602, USA

Fei Zhou (周非) and Vidvuds Ozoliņš[*]

Department of Materials Science and Engineering, University of California, Los Angeles, California 90095, USA

The widely accepted intuition that the important properties of solids are determined by a few key variables underpins many methods in physics. Though this reductionist paradigm is applicable in many physical problems, its utility can be limited because the intuition for identifying the key variables often does not exist or is difficult to develop. Machine learning algorithms (genetic programming, neural networks, Bayesian methods, etc.) attempt to eliminate the *a priori* need for such intuition but often do so with increased computational burden and human time. A recently developed technique in the field of signal processing, compressive sensing (CS), provides a simple, general, and efficient way of finding the key descriptive variables. CS is a powerful paradigm for model building; we show that its models are more physical and predict more accurately than current state-of-the-art approaches and can be constructed at a fraction of the computational cost and user effort.

I. INTRODUCTION

Physical intuition and experience suggest that many important properties of materials are primarily determined by just a few key variables. For instance, the crystal structures of intermetallic compounds have been successfully classified into groups (so-called structure maps) according to the properties of the constituent atoms.[1–4] The widely known Miedema rules relate alloy formation energies to atomic charge densities and electronegativities.[5] Most magnets can be described using a Heisenberg model with only a few short-ranged exchange interactions,[6] and the formation energies of multicomponent alloys can be efficiently parameterized using generalized Ising models (cluster expansions) with a finite number of pair and multibody interactions.[7–9] In all these cases, enormous gains in efficiency and conceptual clarity are achieved by building models that express the quantity of interest (typically, total energy) in a simple, easy-to-evaluate functional form. These models can then be used to perform realistic simulations at finite temperatures, on large systems, and/or over long time scales, significantly extending the reach of current state-of-the-art quantum-mechanics based methods.

The conventional approach to model building starts by selecting a small, physically motivated basis set that describes the configuration of the system. The target properties are then expressed in terms of these basis functions and the unknown coefficients are determined by performing least-squares fits to the calculated or experimentally measured data. While conceptually simple, this method is often difficult to use in practice. First, the number of unknown coefficients has to be smaller than the number of data points, which precludes the use of very large basis sets. Second, least-squares fitting is susceptible to noise, and there is often the possibility of "overfitting"—the model is trained to reproduce the fitting data, but performs poorly in a predictive capacity. Finally, finding the optimal finite basis set is an *NP*-hard problem, i.e., the solution time increases faster than polynomial with the number of possible basis functions. To keep the number of coefficients smaller than the amount of data, one must

choose, based on physical intuition, which basis functions to keep. This physical intuition in many cases may be unavailable and/or difficult to develop; hence there is no clear path to achieve systematic improvement. Recent years have seen numerous attempts to use machine learning algorithms (genetic programming, neural networks, Bayesian methods, etc.) to decrease the role of intuition in model building.[10–17]

We show that a recently developed technique in the field of signal processing, compressive sensing (CS),[18] provides a simple, general, and efficient approach to model building.[19] Instead of attempting to develop physical intuition for which coefficients will be most relevant, the CS framework allows the inclusion of essentially all possible basis functions. Using very large basis sets eliminates the need to use physical intuition to construct smaller ones. Furthermore, CS is computationally efficient for very large problems, robust even for very noisy data, and its models predict more accurately than current state-of-the-art approaches.

II. COMPRESSIVE SENSING: AN ILLUSTRATION

Before demonstrating the power of compressive sensing for building physical models, we first illustrate the concept itself with a simple time series. Discussion of compressive sensing requires the definition of ℓ_p norms:

$$\|u\|_p = \left(\sum_i |u_i|^p \right)^{1/p}, \tag{1}$$

of which the ℓ_1 (taxicab or Manhattan distance) and ℓ_2 (Euclidean; subscript 2 often omitted) norms are special cases. The number of nonzero elements of $\vec{u}$ is often (improperly) referred to as the ℓ_0 "norm" even though it is not a norm in a strict mathematical sense.

In the signal processing community, compressive sensing is used to recover sparse signals *exactly* with far fewer samples than required by standard spectral techniques, such as the well-known Fourier and Laplace transforms. Consider a signal like

that shown in Fig. 1(a), which has the functional form

$$f(t) = \sum_{n=1}^{N} u_n e^{i2\pi nt}, \tag{2}$$

where most of the coefficients u_n are zero (i.e., the signal is sparse). The Fourier transform is mathematically equivalent to solving the matrix equation

$$\mathbb{A}\vec{u} = \vec{f}, \tag{3}$$

where the matrix $\mathbb{A}$ is formed by the values of the Fourier basis functions at the sampling times t_m, i.e., it consists of rows of n terms of the form $A_{mn} = e^{i2\pi nt_m}$, and $f_m \equiv f(t_m)$ is the sampled signal. The solution vector $\vec{u}$ contains the relative amounts of the different Fourier components, as shown in Fig. 1(b). Capturing all relevant frequency components of the signal using Fourier transform techniques requires the signal to be sampled regularly and at a frequency at least as high as the Nyquist frequency [shown as red points in Fig. 1(a)], a severe restriction stemming from the requirement that the linear system (3) should not be underdetermined.

However, the main idea of compressive sensing is that,when the signal is sparse, one should be able to recover the exact signal with a number of measurements that is proportional to the number of nonzero components, i.e., with far fewer samples than given by the Nyquist frequency. Conceptually, this could be done by searching for a solution that reproduces the measured time signal exactly *and* has the minimum number of nonzero Fourier components. Unfortunately, this formulation results in a discrete optimization problem, which cannot be solved in polynomial time. Compressive sensing recasts the problem as a simple minimization of the ℓ_1 norm of the solution, subject to the constraint given by Eq. (3) above:

$$\min_u \{\|\vec{u}\|_1 : \mathbb{A}\vec{u} = \vec{f}\}, \tag{4}$$

where $\|\vec{u}\|_1 = \sum_i |u_i|$ is the ℓ_1 norm defined in Eq. (1). In other words, one seeks to minimize the sum of the components of the solution vector $\vec{u}$ subject to the condition that the measured signal is reproduced exactly; this constitutes the so-called basis pursuit problem. Equation (4) is a convex optimization problem, which can be solved efficiently (see Sec. IV). We note here that optimization of the common sum-of-squares (ℓ_2) norm of $\vec{u}$ would result in a dense solution that may deviate considerably from the original signal.[18]

As a simple illustration, the exact decomposition of an example function, shown in Fig. 1, was possible via compressive sensing with only five random samples (black dots in Fig. 1) of the signal, instead of the 24 equally spaced samples (red dots in Fig. 1) needed for a discrete Fourier transform. Quite generally, a mathematical theorem proven by Candes, Romberg, and Tao[20] guarantees that with an overwhelming probability, any sparse signal with S nonzero components can be recovered from $M \sim S \ln N$ random measurements, where N is the total number of sensing basis functions. This very powerful result is the mathematical foundation of compressive sensing.

Another practically important feature of compressive sensing is the ability to tolerate noise in the input data and to deal with signals that are only approximately sparse, i.e.,
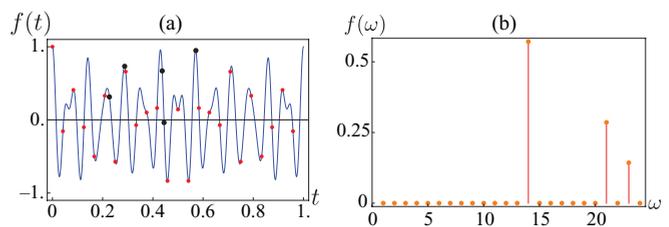

FIG. 1. (Color online) (a) A sparse signal (blue line) like that of Eq. (2), uniform samples of the signal at the Nyquist frequency (red dots), and a few random samples (black circles). The signal is composed of only three nonzero frequencies. (b) Exact recovery of the frequency components of the signal using compressive sensing.

are dominated by a few large terms, but also contain a large number of smaller contributions; this is the case in almost all physics applications. It has been proven that, if the sensing matrix $\mathbb{A}$ obeys the so-called *restricted isometry property* (RIP), an accurate reconstruction of the signal from highly under-sampled measurements can be achieved also in the presence of both random and systematic noise.[18,20] The RIP criterion is automatically satisfied if the measurements are chosen randomly (see Sec. IV D for a detailed discussion).

When applying compressive sensing to model building, two tasks must be accomplished: (i) a basis must be chosen and (ii) the coefficients associated with each basis function must be determined. Mathematically, the problem is analogous to the simple Fourier example considered above, with the sensing matrix $\mathbb{A}$ being determined by the values of the basis functions at the chosen measurement points. Below, we illustrate the use of compressive sensing on two cluster expansion (CE) models of configurational energetics:[7] (i) Ag-Pt alloys on a face-centered cubic (fcc) lattice and (ii) protein folding energies in the so-called zinc finger motif. CE is chosen as an example because it is conceptually simple, mathematically rigorous, and widely used in the materials community to calculate temperature-composition phase diagrams. Furthermore, CE is a stringent test case for compressive sensing because a significant amount of effort has been expended developing advanced model building techniques, which have been implemented in sophisticated general-purpose computer codes.[15–17,21–25]

III. CLUSTER EXPANSION

A. Energy model

Since a formal mathematical description of CE can be found in the literature, here we only restate its main features and refer the reader to Refs. 7–9 for detailed explanations. The CE method uses a complete set of discrete basis functions, defined over clusters of lattice sites, which describe the occupation of each site and thus the entire atomic configuration on the crystal. The total energy is given by

$$E(\sigma) = E_0 + \sum_f \bar{\Pi}_f(\sigma) J_f, \tag{5}$$

where f represents symmetrically distinct clusters of lattice sites (points, pairs, triplets, etc.), σ denotes the atomic configuration, usually expressed by a collection of pseudospin variables $\{S_i\}$ describing the type of atom at each lattice site, and the cluster correlations $\bar{\Pi}_f(\sigma)$ are formed as symmetrized

averages of products of these pseudospin variables. The key quantities in this approach are J_f, the effective cluster interactions (ECI's): given the ECI's, the energy of *any* atomic configuration on the lattice can be calculated rapidly from Eq. (5). Physical intuition based on the concept of "near-sightedness" of screened interatomic interactions suggests that only clusters within a limited range and involving a limited number of sites will have significant ECI's. The goal then is to determine which of the clusters f, out of the myriads of possible choices, contribute significantly to the total energy of the system and to calculate the values of these coefficients.

Currently, the most popular approaches are based on the so-called structure inversion method (SIM),[26] where a limited number of quantum-mechanics-based total energy calculations are used to determine $E(\sigma)$ on the left-hand side of Eq. (5). The cluster interactions J_f are truncated according to some recipe and their values are determined by least-squares fitting to the training set energies $E(\sigma)$. The accuracy of the resulting CE depends crucially on the chosen truncation method. Including too few interactions leads to poor predictive power because important interactions are not accounted for ("underfitting," while choosing too many parameters J_f results in spurious interactions and an associated decrease in predictive accuracy ("overfitting"). Use of the least-squares fitting necessarily requires that the number of structures must exceed the number of candidate ECIs, which is the CE analog of the Nyquist frequency in signal processing.

In modern practice, the trial ECI's are chosen by scanning over many possible sets of clusters while attempting to minimize the predictive error. Ideally, the predictive error should be calculated as the root mean square (RMS) deviation between the density functional theory (DFT) and CE-predicted energies over a separate "holdout" set of structures that are not used in fitting. This approach would require tens or hundreds of additional DFT calculations and is therefore seldom used in practice. Leave-one-out cross-validation (LOOCV) or k-fold cross-validation scores are commonly used as proxies for the predictive error since they do not require the construction of a separate holdout set.[21]

Starting from an initial set of ECI's (e.g., empty, point, and nearest-neighbor pair clusters), a typical procedure for improving the model attempts to add and/or substitute clusters into the current set, keeping changes if the predictive error is found to decrease. The procedure is terminated when none of the attempted changes produce an improvement in the predictive accuracy. Unfortunately, there is no guarantee that this process will truly result in the optimal set of ECI's because it is practically impossible to solve the *NP*-hard discrete optimization (DO) problem, especially, if the number of candidate ECI's is large, such as required for very accurate CE's or in situations of low symmetry (e.g., near defects, surfaces, nanoparticles). As a result, with DO, one may miss important clusters or add clusters that should not have been included in the model, which may result in nonphysical ECI values. Additionally, adding/removing clusters one by one is computationally expensive, requiring days to finish when considering very large pools of candidate clusters.

Genetic algorithms have been used with some success, but they also require large amounts of time to complete, especially for large cluster pools and fitting sets, and employ a host of tunable parameters.[23,27] Other methods for optimizing the CV score have been proposed[28,29] but are limited in the number of basis functions they can consider and require a heavy initial truncation of the basis.

Other researchers, in an attempt to avoid predictive errors associated with incomplete discrete optimization, have sought to devise direct minimization methods that automatically select ECIs only if they are required to reproduce the energies of the training set. Such an approach was proposed by Laks, Wei, and Zunger for pair interactions,[30] who added a distance-weighted ℓ_2 norm of the pair interactions to the objective function. However, this approach usually results in dense sets of long-ranged pair ECI's and, more importantly, is difficult to extend to multibody interactions.[31–33] Recently, a method based on Bayesian statistics was introduced to automatically estimate ECI's and shown to outperform several common methods in low-symmetry situations.[16] However, it makes use of physical intuition to construct informative prior distributions, which are required for estimating the ECI values. It is desirable to develop methods that avoid the use of intuition since heuristic rules, derived from experience with a few specific systems, may not be universally valid. Design of efficient, numerically robust and physically accurate methods for selecting the physically significant ECIs remains a challenging problem.

B. Compressive sensing cluster expansion (CSCE)

Here, we show that compressive sensing can be used to select the important ECI's and determine their values *in one shot*. The applicability of compressive sensing to CE is based on the mathematical theorem of Candes, Romberg, and Tao,[20] which guarantees that sparse ECI's can be recovered from a limited number of DFT formation energies given certain easy-to-satisfy properties of the matrix $\bar{\Pi}$ in Eq. (5). Adopting the common assumption that the "true" physical ECI's are approximately sparse, this theorem guarantees that a good approximation will be found even in cases when the data have both random and systematic noise, e.g., due to numerical errors in the DFT calculations or due to interactions beyond the chosen energy resolution, see Sec. IV E.

There are two possible formulations for compressive sensing cluster expansion (CSCE), both of which enforce the requirement that the cluster expansion should be as sparse as possible, while resulting in a certain level of accuracy for the training set. In the first approach, one may determine the optimal set of ECI's from

$$J = \arg\min_{J}\{||J||_1 : ||E - \bar{\Pi}J|| \leqslant \epsilon\}, \qquad (6)$$

where ℓ_1 norm of J's is used as a proxy for the number of nonzero ECI's. Solving Eq. (6), the least absolute shrinkage and selection operator (LASSO) problem,[34,35] offers a mathematically strict way of constructing a minimal cluster set that reproduces the training set with a given accuracy. Of course, over- (under-) fitting is still an issue if ϵ is chosen too small (large), but it is physically reasonable that, given the physical properties of the system and the size of the training set, an optimal ϵ always exists. Following common practice, optimal ϵ could be found either by minimizing the LOOCV score or the predictive rms error over a holdout set.

Since the inequality constraint is inconvenient to enforce during calculations,[34] here we follow common practice in signal processing and use an unconstrained approach that minimizes the sum of an ℓ_1 norm of the ECI's and a least-squares sum of the fitting errors:

$$J = \arg\min_J \mu ||J||_1 + \tfrac{1}{2}||E - \bar{\Pi} J||^2, \tag{7}$$

where μ is a parameter that controls the accuracy of the fit versus the sparseness of the solution: higher values of μ will result in sparser solutions and larger fitting errors (underfitting), while very small μ values will lead to dense solutions and degraded predictive accuracy (overfitting). It will be shown below in Sec. IV E that the optimal value of μ is proportional to the level of noise (random and systematic) in the calculated formation energies. Just like ϵ in Eq. (6), an optimal μ to avoid over- or underfitting can be chosen either by minimizing the LOOCV score or by minimizing the rms prediction error for a separate holdout set; it is shown below that both approaches result in very similar values of optimal μ. Furthermore, in Sec. V, we demonstrate that CSCE is not particularly sensitive to the precise value of μ and show that there is usually a range of μ's that give ECI's of similar predictive accuracy.

The main advantage of CSCE, Eqs. (6) and (7), over current CE methods is that the *NP*-hard discrete optimization of the truncated ECI set is replaced by convex optimization problems for which exact solutions may be found in polynomial time. Furthermore, the minimization of the ℓ_1 norm of the solution also serves to decrease the magnitude of the ECI's, leading to "smoother" ECI's, increased numerical stability with respect to the noise in the training data, and eventually more accurate predictions. In addition, the CSCE is simple to implement and use, which will facilitate its widespread adoption in solid state physics and other fields where configurational energetics play a role. In Sec. V below, we illustrate the superior performance of CSCE using examples from bulk alloys (Ag-Pt) and biology (protein folding energetics).

IV. PRACTICAL ASPECTS OF ℓ_1-BASED OPTIMIZATION

In what follows, we review methods for solving the unconstrained minimization problem given by Eq. (7), which we rewrite as

$$\min_u \mu ||\vec{u}||_1 + \tfrac{1}{2}||\mathbb{A}\vec{u} - \vec{f}||^2. \tag{8}$$

Equation (8) is referred to as the basis pursuit denoising problem. It has a tunable parameter, μ, which controls the sparseness of the solution: smaller (larger) values of μ produce less (more) sparse solutions.

A. Fixed-point continuation

The fixed-point continuation (FPC) method of Hale, Yin, and Zhang[36] is an iterative algorithm that starts from $\vec{u}^0 = \mathbf{0}$ and attempts to improve the objective function by following the gradient of the ℓ_2 term:

$$\vec{g}^k = \mathbb{A}^T(\mathbb{A}\vec{u}^k - \vec{f}), \tag{9}$$

$$u_n^{k+1} = \mathrm{shrink}(u_n^k - \tau g_n^k, \mu\tau), \tag{10}$$

where $k = 0, 1, 2, \dots$ is the iteration number and the shrinkage operator is defined as

$$\mathrm{shrink}(y, \alpha) := \mathrm{sign}(y)\max(|y| - \alpha, 0). \tag{11}$$

In other words, shrinkage decreases the absolute magnitude of y by α and sets y to zero if $|y| \leqslant \alpha$. The iterations are stopped when the ℓ_∞ norm, or maximum component value, of the gradient drops below the shrinkage threshold,

$$\frac{1}{\mu}||\vec{g}||_\infty - 1 < \delta_g, \tag{12}$$

and the change in the solution vector is sufficiently small,

$$\frac{||\vec{u}^{k+1} - \vec{u}^k||}{||\vec{u}^k||} < \delta_u. \tag{13}$$

The sensing matrix should be normalized in such a way that the largest eigenvalue α_A of $\mathbb{A}^T\mathbb{A}$ is less than or equal to 1; this is easily accomplished by dividing both $\mathbb{A}$ and $\vec{f}$ by $\sqrt{\alpha_A}$. The step size τ in Eq. (10) is given by

$$\tau = \min\left(1.999, -1.665\frac{M}{N} + 2.665\right), \tag{14}$$

where M and N are the number of equations and the number of expansion coefficients, respectively.

B. Bregman iteration

While FPC algorithm is generally applicable to any problem of type Eq. (8) and is guaranteed to converge, in practice it has a serious shortcoming: very small values of μ are needed to recover the exact solution to the basis pursuit problem without noise, Eq. (4), which cause an associated increase in the number of FPC iterations. To alleviate the need to use small μ's, Yin *et al.*[37] proposed an efficient iterative denoising algorithm for finding the solution to Eq. (8), which has the additional benefit of yielding the exact solution to the basis pursuit problem Eq. (4) for zero noise. This so-called Bregman iteration involves the following two-step cycle:

$$\vec{f}^{k+1} = \vec{f} + (\vec{f}^k - \mathbb{A}\vec{u}^k), \tag{15}$$

$$\vec{u}^{k+1} = \arg\min_u \mu||\vec{u}||_1 + \tfrac{1}{2}||\mathbb{A}\vec{u} - \vec{f}^{k+1}||^2, \tag{16}$$

starting from $\vec{f}^0 = \mathbf{0}$ and $\vec{u}^0 = \mathbf{0}$. A key feature of the algorithm is that the residual after iteration k is added back to the residual vector $\vec{f}^{k+1}$ for the next iteration, resulting in efficient denoising and rapid convergence.[37] Each minimization in Eq. (16) can be performed using the fixed-point continuation (FPC) method proposed by Hale, Yin, and Zhang.[36] The main advantages of the Bregman iteration are faster convergence and the ability to use μ values that are several orders of magnitude larger than those required for direct application of the FPC method.

C. Split Bregman iteration

For very large problems (i.e., large sensing matrices $\mathbb{A}$), the FPC optimization steps in the Bregman iterative method progress very slowly. The problem becomes severe when the condition number computed from the nonzero eigenvalues of $\mathbb{A}^T\mathbb{A}$ becomes large. Indeed, FPC is essentially a steepest

descent method combined with an ℓ_1 shrinkage step, and the number of required steepest descent iterations increases linearly with the ratio of the largest-to-smallest eigenvalues of $\mathbb{A}^T \mathbb{A}$.[38] An improved Bregman algorithm, which eliminates the hard-to-solve mixed ℓ_1 and ℓ_2 minimization problem in Eq. (16), was proposed by Goldstein and Osher.[39] It carries the name of "split Bregman" iteration because it splits off the ℓ_1 norm of the solution from the objective function and replaces it with the variable $\vec{d}$, which is designed to converge towards the ℓ_1 term, $\lim_{k\to\infty}(\vec{d}^k - \mu\vec{u}^k) = \mathbf{0}$. A least-squares ℓ_2 term is added to the objective function to ensure that $\vec{d} = \mu\vec{u}$ in the limit:

$$\vec{u} = \arg\min_{u,d} \|\vec{d}\|_1 + \tfrac{1}{2}\|\mathbb{A}\vec{u} - \vec{f}\|^2 + \tfrac{\lambda}{2}\|\vec{d} - \mu\vec{u}\|^2. \quad (17)$$

A key advantage of this formulation is that the minimization involving the quadratic form $\|\mathbb{A}\vec{u} - \vec{f}\|^2$ does not contain ℓ_1 terms and can be performed efficiently using standard convex optimization techniques, such as Gauss-Seidel or conjugate gradients (CG),[38] while the ℓ_1 minimization with respect to $\vec{d}$ at a fixed $\vec{u}$ contains an ℓ_2 term that is diagonal in the components of $\vec{d}$ and can be solved easily (see below). The full split Bregman iterative algorithm proceeds as follows:

$$\vec{u}^{k+1} = \arg\min_u \tfrac{1}{2}\|\mathbb{A}\vec{u} - \vec{f}\|^2 + \tfrac{\lambda}{2}\|\vec{d}^k - \mu\vec{u} - \vec{b}^k\|^2, \quad (18)$$

$$\vec{d}^{k+1} = \arg\min_d \|\vec{d}\|_1 + \tfrac{\lambda}{2}\|\vec{d} - \mu\vec{u}^{k+1} - \vec{b}^k\|^2, \quad (19)$$

$$\vec{b}^{k+1} = \vec{b}^k + \mu\vec{u}^{k+1} - \vec{d}^{k+1}, \quad (20)$$

starting from $\vec{d}^0 = \mathbf{0}, \vec{b}^0 = \mathbf{0}$, and $\vec{u}^0 = \mathbf{0}$. We use the conjugate gradient method to perform the ℓ_2 minimization in Eq. (18). The second step, Eq. (19), separates into individual vector components and can be solved explicitly using shrinkage as

$$d_n^{k+1} = \text{shrink}\left(\mu u_n^{k+1} + b_n^k, 1/\lambda\right). \quad (21)$$

The final step of the split Bregman cycle, Eq. (20), adds back the residual deficit in the ℓ_1 term, in complete analogy with the Bregman iteration Eq. (15). The results do not depend on the value of the parameter λ, although an unsuitable choice will lead to very slow or failed convergence. We find that in practice an optimal λ can easily be found from a few trial runs at a fixed value of μ, and then kept fixed for any μ. Just like FPC, the split Bregman iteration provides an exact solution to the basis pursuit denoising problem Eq. (8), but in contrast to the Bregman approach of Sec. IV B, small values of μ may be needed to solve the noiseless basis pursuit problem Eq. (4). In practice, we find that the convergence rate of the split Bregman method is almost always faster than those of the Bregman or FPC algorithms, and greatly so for large, ill-conditioned sensing matrices.

D. Choice of structures for CSCE

An important practical question regards the best strategy for choosing structures σ to include in the training set. Mathematical theorems from compressive sensing provide a definite answer to this question. The key idea is the notion of coherence between the measurement and representation basis. The representation basis $\Phi = \{\phi_j\}$ is used to express the signal as a sparse series expansion (e.g., plane waves form the

representation basis for the Fourier series), while the measurement basis $\Psi = \{\psi_k\}$ contains all possible measurements. For the Fourier example in Sec. II, the measurement basis is given by δ functions, i.e., signal values at certain points in time. Assuming that both ψ_j and ϕ_k are normalized and orthogonal, the coherence is defined as the maximum overlap between them:[40]

$$\nu(\Phi, \Psi) = \sqrt{N} \max_{j,k} |\langle\phi_j, \psi_k\rangle|. \quad (22)$$

In the Fourier example of Sec. II, the scalar products are all $|\langle\phi_j, \psi_k\rangle| = N^{-\frac{1}{2}}$, which corresponds to the lowest possible coherence, $\nu = 1$. In contrast, the highest possible value $\nu = \sqrt{N}$ would be obtained by directly measuring the amplitudes of the individual sinusoidal components of the signal, i.e., if plane waves were chosen as the measurement basis functions. Coherence is key in determining the number of measurements required to recover a given sparse signal with S nonzero components: the higher the coherence, the higher the required number of measurements. More quantitatively, the probability of correct signal recovery from M measurements exceeds $1 - \delta$ if the number of measurements satisfies $M \geqslant C\nu^2(\Phi, \Psi)S\log(N/\delta)$, where C is a constant and S is the number of nonzero components;[41,42] a similar result holds for compressive sensing in the presence of noise. This expression shows that the worst possible strategy for recovering sparse signals is to choose the same measurement basis as the one used in sparse representation $[\nu(\Phi, \Psi) \approx \sqrt{N}]$, since this would require a number of measurements equal to the number of unknown coefficients, N.

In cluster expansion, the representation basis are formed by symmetry-distinct cluster types and the measurements are represented by structures σ. The corresponding representation basis functions are Kroenecker δ's, $\phi_g(f) = \delta_{fg}$, where f and g are cluster numbers. The measurements are represented by symmetry-inequivalent structures σ, and the corresponding basis functions are given by normalized rows of the cluster correlation matrix, i.e., $\psi_\sigma(f) = \bar{\Pi}_f(\sigma)/\sqrt{\sum_{f'} \bar{\Pi}_{f'}(\sigma)^2}$. The coherence is given by the maximum scalar product between the two, which is

$$\nu(\Phi, \Psi) = \sqrt{N} \max_{\sigma, f} \frac{|\bar{\Pi}_f(\sigma)|}{\sqrt{\sum_{f'} \bar{\Pi}_{f'}(\sigma)^2}}. \quad (23)$$

Because random matrices with independent identically distributed (i.i.d.) entries are incoherent with almost any representation basis, they occupy a special place in compressive sensing. If the possible measurements are designed by selecting N uniformly distributed random vectors on the unit sphere, followed by subsequent orthogonalization, the coherence between Φ and Ψ is on the order of $\sqrt{2 \ln N}$.[18] This suggests the following simple strategy for selecting structures for CSCE: (1) generate M uniformly distributed random vectors $\psi_\sigma(f)$ on the unit sphere ($\sigma = 1, \ldots, M$), (2) orthogonalize $\psi_\sigma(f)$, and (3) match each $\psi_\sigma(f)$ onto a real structure σ with normalized correlations $\bar{\Pi}_f(\sigma)/\sqrt{\sum_{f'} \bar{\Pi}_{f'}(\sigma)^2}$ approximating $\psi_\sigma(f)$ as closely as possible.

The last step can be conveniently performed by enumerating all possible ordered structures up to a certain size of the unit cell using the methods of Refs. 43 and 44 and then

choosing the best matches from this list. We stress that the somewhat counterintuitive strategy of selecting random structures follows from the general mathematical properties of ℓ_1-based compressive sensing and represents the best possible method for choosing structure sets for CSCE.

Parenthetically, we note that selecting structures at random makes for a remarkably simple approach to generating input data. The "structure selection" problem, that is, deciding which structure to use to train the model, has been a vexing problem in the cluster expansion community since cluster expansions first began to be trained with first-principles data. At first, structures that were easy to calculate (few atoms per unit cell) were selected. In later years, more sophisticated approaches came to be used,[21,45,46] but a simple, easy-to-implement solution has remained elusive. Compressive sensing not only solves the "cluster selection" problem (because it makes unbiased selections from a huge set of clusters) but also overcomes the structure selection problem because it dictates that the best strategy is to select ordered structures with pseudorandom correlations.

E. Effect of noise and its relation to optimal μ

The lone adjustable parameter, μ, should be chosen to achieve the optimal balance between the sparseness of the ECI's and the RMS fitting error for the training set. The effect of μ on the calculated ECI's is most transparently seen by analyzing the FPC equations (9) and (10), which show that μ controls the energy cutoff for the gradient of the ℓ_2 norm of the residuals: components of $\vec{g}$ with absolute values $|g_f| \leqslant \mu$ will be set to zero by the shrinkage operator and therefore will be excluded from the model. In what follows, we show that the optimal value for μ is proportional to the level of noise (random and systematic) in the training data.

We first consider the relation between the normalized sensing matrix $\mathbb{A}$ in Eq. (9) and the CSCE correlation matrix $\bar{\Pi}$: they are related by $\mathbb{A} = \bar{\Pi}/\sqrt{\alpha_{\bar{\Pi}}}$, where $\alpha_{\bar{\Pi}}$ is the largest eigenvalue of $\bar{\Pi}\bar{\Pi}^T$. The corresponding relation for the measurement vectors is $\vec{f} = E/\sqrt{\alpha_{\bar{\Pi}}}$. The distributions of the extremal eigenvalues for ideal random matrices are known from the theory of principal component analysis.[47] However, it is not immediately clear that the eigenvalue distributions found for i.i.d. random matrices will be directly applicable to the CSCE correlation matrices $\bar{\Pi}$ because the correlation values for real structures are neither independent nor identically distributed, and hence the entries of $\bar{\Pi}$ are only approximately i.i.d. We have numerically calculated the distribution of the largest eigenvalue of $\bar{\Pi}\bar{\Pi}^T$ using subsets of $1 \leqslant M \leqslant 500$ fcc-based ordered structures with 12 or fewer atoms in the unit cell.[43] We considered $N = 986$ correlations (up to six-body terms) and averaged the calculated eigenvalues over 1000 subsets randomly drawn from the above list of 10 850 structures. We find that, for a fixed N, the average value of $\alpha_{\bar{\Pi}}$ increases linearly with the number of structures M. Therefore $\mathbb{A} \propto \bar{\Pi}/\sqrt{M}$.

a. Random noise. Here, we demonstrate that CSCE is not only stable with respect to noise in the input data, but that it can also filter out the effects of noise on the calculated ECI's. We assume that the DFT formation energies $E(\sigma)$ contain random noise, which is represented by a vector $\vec{\eta}_\epsilon$ of length M and

i.i.d random components with variance $\epsilon_{\mathrm{rand}}^2$. The contribution of $\vec{\eta}_\epsilon$ to the FPC gradient in Eq. (9) is given by

$$\delta g_f \propto -\frac{1}{M}\sum_{\sigma=1}^{M}\bar{\Pi}_f(\sigma)\eta_\epsilon(\sigma), \qquad (24)$$

where the factor $1/M$ comes from the fact that both the sensing matrix $\mathbb{A}$ and the measurement vector $\vec{f}$ are related to the correlation matrix $\bar{\Pi}$ and input energies E by a normalization factor $1/\sqrt{\alpha_{\bar{\Pi}}}$. If the structures are chosen randomly according to the prescription outlined in Sec. IV D, then $\bar{\Pi}_f(\sigma) \in [-1,1]$ are approximately i.i.d. Hence the individual terms under the summation sign in Eq. (24) will be randomly distributed with a mean of zero and a variance proportional to $\epsilon_{\mathrm{rand}}^2$. To deduce the behavior of δg_f in the limit of large M, one can apply the central limit theorem (CLT) of classical statistics, which states that the average of M random terms is normally distributed with a variance that is given by the variance of the individual terms divided by M, i.e., the variance of δg_f is proportional to $\epsilon_{\mathrm{rand}}^2/M$. It then follows from the properties of the normal distribution that the average ℓ_1 norm of the noise term in the gradient decreases with the size of the training set as

$$||\delta g_f||_1 \propto \frac{\epsilon_{\mathrm{rand}}}{\sqrt{M}}. \qquad (25)$$

This relation demonstrates an important noise-tolerance aspect of CSCE, which guarantees that the true physical ECI's will be recovered even if the training data sets contains uncorrelated random noise of arbitrary magnitude, provided that the number of data points is sufficiently large. The practical significance of this feature cannot be overstated: not only is CSCE stable with respect to random noise, but an absolute numerical accuracy in the DFT energies is not even needed to recover the correct ECI's![48] Equation (25) also offers guidance for choosing μ to smooth the effect of random noise: as long as $\mu \simeq ||\delta g_f||_1$, the contribution of noise to the gradient will be zeroed out in the shrinkage step [see Eq. (10)] and will not affect the calculated ECI's. In practice, however, the optimal value of μ is difficult to determine using Eq. (25) because the level of noise in the DFT formation energies is not known *a priori,* and approaches based on optimizing the predictive error are more practical.

b. Systematic noise. We next consider the effect of systematic noise due to errors in the ECI's, which we denote by δJ_f. These errors contribute a term $\delta E(\sigma) = \sum_f \bar{\Pi}_f(\sigma)\delta J_f$ to the residual, and the corresponding error in the FPC gradient is given by

$$\delta g_f \propto -\sum_{f'}\langle\bar{\Pi}_f\bar{\Pi}_{f'}\rangle\delta J_{f'}, \qquad (26)$$

where we have introduced a correlation matrix for cluster correlations $\bar{\Pi}$ calculated over the training set:

$$\langle\bar{\Pi}_f\bar{\Pi}_{f'}\rangle = \frac{1}{M}\sum_{\sigma=1}^{M}\bar{\Pi}_f(\sigma)\bar{\Pi}_{f'}(\sigma). \qquad (27)$$

This matrix is of fundamental importance for CSCE because it describes how the value of one ECI is affected by errors in the other ECI's, or the degree of cross contamination between systematic ECI errors. Minimum sensitivity to cross contamination is achieved when $\langle\bar{\Pi}_f\bar{\Pi}_{f'}\rangle$ is diagonal, but the

latter case is impossible to realize in practice due to the fact that there are rather pronounced correlations between the cluster averages in real structures. In the best case scenario, the correlation matrix $\langle \bar{\Pi}_f \bar{\Pi}_{f'} \rangle$ will be approximately diagonal if the training set structures are chosen randomly according to the algorithm proposed in Sec. IV D. Indeed, if the average cluster correlations $\bar{\Pi}_f(\sigma)$ are approximately i.i.d., the off-diagonal elements of the correlation matrix $\langle \bar{\Pi}_f \bar{\Pi}_{f'} \rangle$ tend to zero with increasing M, while the diagonal elements remain $O(1)$:

$$\langle \bar{\Pi}_f \bar{\Pi}_{f'} \rangle = \begin{cases} \langle \bar{\Pi}_f^2 \rangle & \text{for } f = f' \\ O\left(\frac{1}{\sqrt{M}}\right) & \text{for } f \neq f' \end{cases} . \qquad (28)$$

Hence, in the limit of large M, CSCE based on a randomly chosen training set cleanly separates the contributions of the systematic ECI errors to the gradient, i.e., the ECI error for cluster f only affects the component f of the gradient, enabling accurate recovery of the correct solution. This is an important feature for any physics model-building approach because it guarantees the stability of the solution with respect to the interactions that are not represented within the chosen basis set. Furthermore, these considerations offer another insight into the physical meaning of the parameter μ: it can be used to filter out the cross-contamination due to effects of systematic noise if chosen as

$$\mu \sim \frac{||\delta \vec{J}||_\infty}{\sqrt{M}}, \qquad (29)$$

where $||\delta \vec{J}||_\infty$ is the magnitude of the largest error in the cluster interactions. Since the diagonal contribution to the gradient remains constant with increasing M, successively smaller ECI's can be extracted by increasing the size of the training set M and simultaneously decreasing the value of μ according to Eq. (29). Unfortunately, the practical value of this expression is limited because the ECI errors are not known, and approaches based on minimizing the prediction errors or CV scores should be used instead.

The preceding analysis shows that μ can be interpreted as a parameter controlling the filtering of the noise in the calculated energies, including both random noise due to numerical errors in the DFT formation enthalpies and systematic noise due to cluster interactions that are not recoverable using the given structure set. Expressing the total noise level as a sum of random and systematic contributions, $\epsilon^2 = \epsilon_{\text{rand}}^2 + \epsilon_{\text{syst}}^2$, the effect of both is expected to decrease the inverse of the size of the training set, and the optimal value of μ is expected to vary as $1\sqrt{M}$. We note here that the Bregman and split Bregman iterations contain additional noise-filtering steps [see Eqs. (15) and (20)] that add back the residual to the residual of the next iteration. As a result, the optimal value of μ will in general vary between the different ℓ_1 optimization approaches, even though the solutions and the predictive errors are practically the same.[37,39]

V. APPLICATIONS

A. Short-ranged pair model with noise

We first work with an *ad hoc* cluster expansion example where we choose a set of sparse coefficients and then use them to compute the energies of various crystal structures for use as input to CSCE. The advantage of this approach is that knowing the exact solution *a priori* allows us to easily determine the accuracy of the solution found by CSCE and determine how numerical noise influences the performance of the algorithm. While this example is certainly not representative of any real alloy system, it clearly illustrates some key features of the method, particularly how CS performs with noisy data.

Using the UNCLE[23] framework the following clusters on an fcc lattice were enumerated: 141 pairs, 293 triplets, 241 four-bodies, 87 five-bodies, and 222 six-bodies (986 clusters in total, including the onsite and empty clusters). The coefficients of the three shortest nearest-neighbor pairs were chosen as 10, 4, and 1, respectively; all other coefficients were set to zero. Uniformly distributed random noise equal to $\sim 10\%$, 20%, and 50% of the noiseless energies was added to the computed energies $E(\sigma)$. We emphasize that these noise levels *significantly* exceed typical numerical errors in the calculated formation enthalpies from state-of-the-art quantum mechanics codes.[49]

The values of each of the 986 basis functions were computed for all structures in the training set, thus forming the sensing matrix, $\mathbb{A}$. The rows of the sensing matrix $\mathbb{A}$, which each represent a training set structure, were constructed by drawing randomly from a uniform distribution on $[-1,1]$. For real systems, such as Ag-Pt in the next section, these rows should be mapped onto real crystallographic configurations as described in Sec. IV D. However, since the quality of the fit for the short-ranged pair case was found to be unaffected by this mapping, either favorably or adversely, we chose to simply use the random vectors themselves in order to simplify computations.

Figures 2(a) and 2(b) illustrate the performance of CS by showing two quantities: (1) the ℓ_1 norm of the difference between the exact and fitted coefficients ($\|J_{\text{exact}} - J_{\text{fit}}\|_1$) and (2) the number of nonzero coefficients (ℓ_0 norm of the solution, $\|J_{\text{fit}}\|_0$). We varied μ to investigate its optimal values for a given noise level. Each data point in Fig. 2 was obtained by averaging over approximately 100 different sets, each of size $M = 200$ or 400.

The curves in Fig. 2 exhibit a series of plateaus, each one indicating a region over which the extracted solution remains practically unchanged. Notice, for example, the plateau located between $\log_{10} \mu = -0.75$ and $\log_{10} \mu = -0.4$ in the $\|J_{\text{fit}}\|_0$ versus μ curve for $M = 200$ and the lowest noise content (circle markers). This plateau indicates that CSCE has extracted three nonzero coefficients. Furthermore, the value of $\|J_{\text{exact}} - J_{\text{fit}}\|_1$ drops close to zero in this range, indicating that CSCE has found essentially the exact answer. Using values of μ below the optimal range results in sharp increases in both the number of nonzero coefficients and in the error $\|J_{\text{exact}} - J_{\text{fit}}\|_1$, indicating overfitting.

Conversely, μ values above the optimal range result in fewer nonzero coefficients and an incremental increase in $\|J_{\text{exact}} - J_{\text{fit}}\|_1$, probably indicating underfitting. As a function of increasing μ, one first obtains a plateau where the CS reproduces the two largest expansion coefficients (10 and 4), followed by another plateau where only the largest coefficient is reproduced. This example illustrates the important point that CS is largely *insensitive* to the choice of μ—the ability to recover the correct solution does not depend on the exact
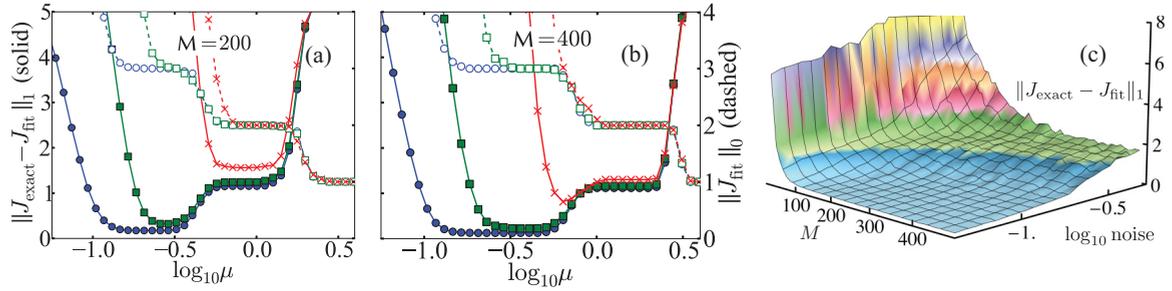
FIG. 2. (Color online) $\|J_{\text{exact}} - J_{\text{fit}}\|_1$(solid) and $\|J_{\text{fit}}\|_0$(dashed) vs $\log_{10}\mu$ for the short-ranged pair model with $M = 200$ (a) and $M = 400$ (b). Random uniform noise of $\sim 10\%$(blue circles), 20%(green squares), and 50% (red "x" s) of the noiseless energies was added to the fitting structures. (c) $\|J_{\text{exact}} - J_{\text{fit}}\|_1$ vs the number of fitting structures and the noise level. Each point represents an average over ~ 100 different subsets of M structures.

value of μ, as long as it lies within an optimal, but broad, range.

Upon increasing the noise in the fitting data at a fixed data set size [compare the curves marked by circles and squares in Fig. 2(a)], the plateaus in $\|J_{\text{fit}}\|_0$ versus μ become narrower until the highest plateau, corresponding to full recovery of the true solution, disappears completely ("x" markers in Fig. 2). At the same time, the minimum in the error $\|J_{\text{exact}} - J_{\text{fit}}\|_1$ versus μ is increasing incrementally. This displays the robustness and stability of CS—even at a very high noise level we are able to recover the majority of the signal content.

The shift towards higher values of optimal μ upon increasing noise level in Fig. 2 is consistent with the physical interpretation of μ as the threshold for noise filtering given in Sec. IV E. We also note that an increase in the number of structures M tends to slightly lower the optimal μ, which can be attributed to a fuller recovery of the correct solution and an associated decrease in the systematic noise.

Figure 2(c) displays $\|J_{\text{exact}} - J_{\text{fit}}\|_1$, averaged over approximately 100 random subsets, as a function of M, the number of fitting structures, and the noise level. Here, we see the same plateau structure found in Fig. 2(a), with the lower (blue) plateau indicating essentially an exact fit. This plot demonstrates that, for all noise levels considered (up to as high as 50% of the noiseless energies!), there remains a training set size for which the exact solution will be recovered.

B. Actual alloy example: Ag-Pt

Having explained the basic properties of CSCE for a model system, we now test its performance on real DFT data for binary Ag-Pt alloys on a face-centered cubic (fcc) lattice. Ag-Pt was chosen due to a report of unusual ordering tendencies[50] which are nontrivial to reproduce with current state-of-the-art CE methods. The energies of more than 1100 Ag-Pt fcc-based crystal structures[51] were calculated from the density-functional theory (DFT) using the VASP software.[52,53] We used projector-augmented-wave (PAW) potentials[54] and the generalized gradient approximation (GGA) to the exchange-correlation functional proposed by Perdew, Burke, and Ernzerhof.[55] To reduce random numerical errors, equivalent k-point meshes were used for Brillioun zone integration.[56] Optimal choices of the unit cells, using a Minkowski reduction algorithm, were adopted to accelerate the convergence of

the calculations.[57] The effect of spin-orbit coupling was not included in our calculations because it's effect was shown to be a simple tilt of the calculated energies, as explained in Ref. 58.

Out of a total of approximately 1100 structure energies calculated for this system, 250 were chosen at random to be held out of the fitting process and used for prediction. This "holdout" set remained unchanged for all fitting sets chosen. Of the remaining 850 data points available for fitting, subsets of up to $N = 400$ were chosen to be used as CSCE training data.

We start by illustrating the performance of two different methods for selecting the optimal value of μ. First, we varied μ and calculated the standard LOOCV score over ten different randomly drawn subsets of M structures; the results are shown by blue curves in Fig. 3. It is seen that the LOOCV scores reach their minima at $\mu \approx 4$ and 2 meV/atom for $M = 100$ and 400, respectively, which we interpret as the optimum μ's providing maximum predictive power. Second, we calculated the average prediction errors for all structures left out of the fitting set,
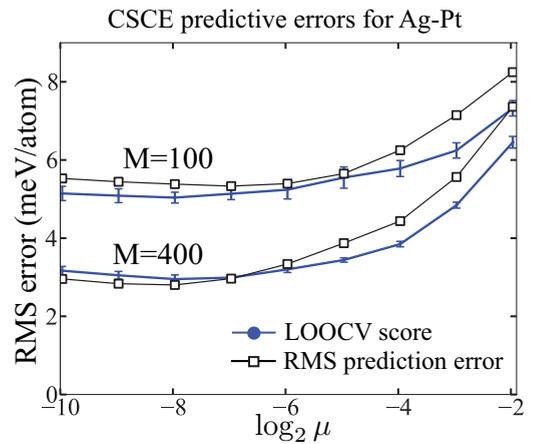

FIG. 3. (Color online) Root-mean-square errors for the prediction set (black line with empty squares) and the leave-one out cross-validation score (LOOCV, solid blue line) as functions of the parameter μ. LOOCV has been averaged over ten randomly drawn sets of 100 (400) structures, and the error bars were calculated from the variance in the predicted LOOCV scores over these sets. Predictive errors for the holdout set and the fitting errors for the training set were averaged over 500 different sets of 100 (400) structures; the corresponding error bars are smaller than the size of the symbols.

which are represented by the black dotted lines in Fig. 3. We see that the RMS errors for the prediction set largely follow the same behavior as the LOOCV scores, reaching minima at nearly identical μ values.

As expected, fitting errors for the training set (not shown here) decrease monotonically with decreasing μ and are significantly smaller than either the LOOCV scores or prediction errors for the holdout set. The leveling off in both the prediction errors and the LOOCV score at small values of μ can be explained by noting that CSCE fits the training set perfectly and further decrease of μ does not bring about noticeable changes in the calculated ECI's. We note that this behavior is different from the short-ranged pair model in the previous section, where decreasing μ below the optimal range caused a rapid deterioration in the accuracy of the calculated ECI's. We attribute this difference to the lower level of noise in the Ag-Pt case, so that the range of μ's that leads to acceptable ECI's is much wider than at the 20–50% noise level for the short-ranged pair model.

To compare the performance of CSCE with other established methods, a discrete optimization (DO) scheme as implemented in the state-of-the-art ATAT software package,[21,22] was used. Note that the ATAT program is capable of employing advanced algorithms beyond minimization of the LOOCV score to ensure that the ground-state line is reproduced correctly and to determine which structures should be used as input. In order to make a straightforward comparison between CSCE and DO and to ensure a reasonable fit construction time for this problem, we only used the LOOCV-based DO functionality of ATAT. Since the DO method for $N = 986$ clusters on a training set of a few hundred structures takes several days to complete, averages were taken over only ten training sets of size M (except for $M = 400$ when we used 42 different training sets to perform statistical analysis of the calculated ECI's). In order to simulate building a complicated unknown model, we deliberately avoided applying physical intuition (e.g., picking short-range interactions) and simply performed the optimizations with minimal restrictions. The maximum number of reported ECI's was capped to $M/4$ for ATAT-based DO. For CSCE, we used a fixed $\mu = 8$ meV/atom and computed solutions for 500 randomly chosen training sets of M structures.

Figure 4 shows a box and whisker plot of the RMS errors over the prediction set for CS solutions and the mean RMS values for the DO solutions (box-and-whiskers were not used for DO solutions due to the small number of DO fits). Each box and whisker represents RMS values for approximately 500 different fits. We see that CSCE achieves an RMS error value much lower (2.8 meV/atom) than LOOCV-based DO (6.8 meV/atom). Furthermore, Fig. 4 shows that the ℓ_1 norm of the solution increases almost linearly for the DO fit, while it levels off for the CSCE fit, indicating that the latter is converging towards a stable solution, while the former keeps adding large ECI's, a behavior suggestive of overfitting.

C. Statistical analysis of Ag-Pt ECI's

Because CSCE is fast, thousands of fits for many different training sets can be computed in a few minutes. The results of all these fits can be analyzed statistically to determine which
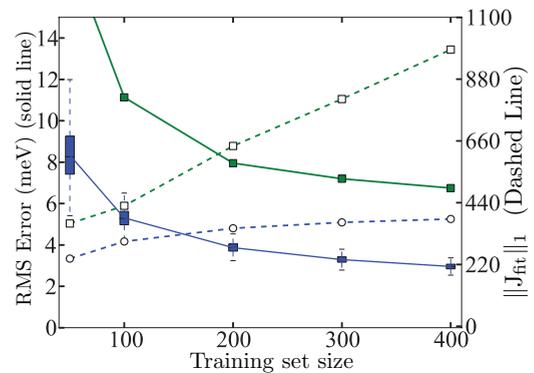

FIG. 4. (Color online) Results from compressive sensing and leave-one-out cross validation for the fcc-based, Ag-Pt alloy system. The solid line gives the root-mean-square (RMS) errors for predictions made on a constant holdout set for CS(box and whisker) and leave-one-out cross validation (squares). The dashed lines give the ℓ_1-norm of the solution vector for both methods.

coefficients are consistently identified as contributors and to eliminate artifacts due to a particular choice of the training set. This functionality, the ability to gather enough data in a reasonable amount of time to perform statistical analyses, is a significant advantage of CSCE over (slower) DO methods that can be used to gain insight into the probability distributions for the cluster interactions. These distributions can be used to quantify the uncertainty in the CSCE predictions for physical properties that go beyond a simple LOOCV score or an RMS prediction error. For instance, one can draw ECI's from the calculated distributions and generate ground state convex hulls with statistical error bars on each structure, quantifying the uncertainty in the predicted $T = 0$ K phase diagrams.

CSCE fits for 500 different fitting set choices were computed for Ag-Pt. Most of the resulting distributions had only one sharp peak at zero, indicating that, independently of the choice of the training set, they were almost never selected by CSCE and therefore should be set to zero. Several ECI's exhibited a unimodal distribution with nonzero mean, which were interpreted as strongly significant nonzero interactions. Finally, a fraction of the ECI's showed bimodal distributions with two peaks of comparable weight and one of the peaks centered at zero energy. Since the latter ECI's were selected by CSCE with an approximately 50% probability, they belong to the class of "marginal" interactions, which were counted as significant only if their distribution mean was greater than one standard deviation. To make a fair comparison between CSCE and the DO methods implemented in the ATAT program, the same statistical criteria for determining relevant coefficients was used for the DO fits, even though data for only 42 fits were available.

Figure 5 gives a comparison of the CS-determined coefficients and those found by DO. The upper pane compares a typical DO fit with a typical CSCE fit, while the lower pane gives a comparison of statistically relevant ECI's from both methods. The CSCE-derived ECI's appear to evolve towards one specific solution as the size of the fitting set increases, indicating convergence of the solution. Notice also that the magnitudes of the CSCE coefficients decrease as the spatial extent of the cluster increases and as the number of cluster
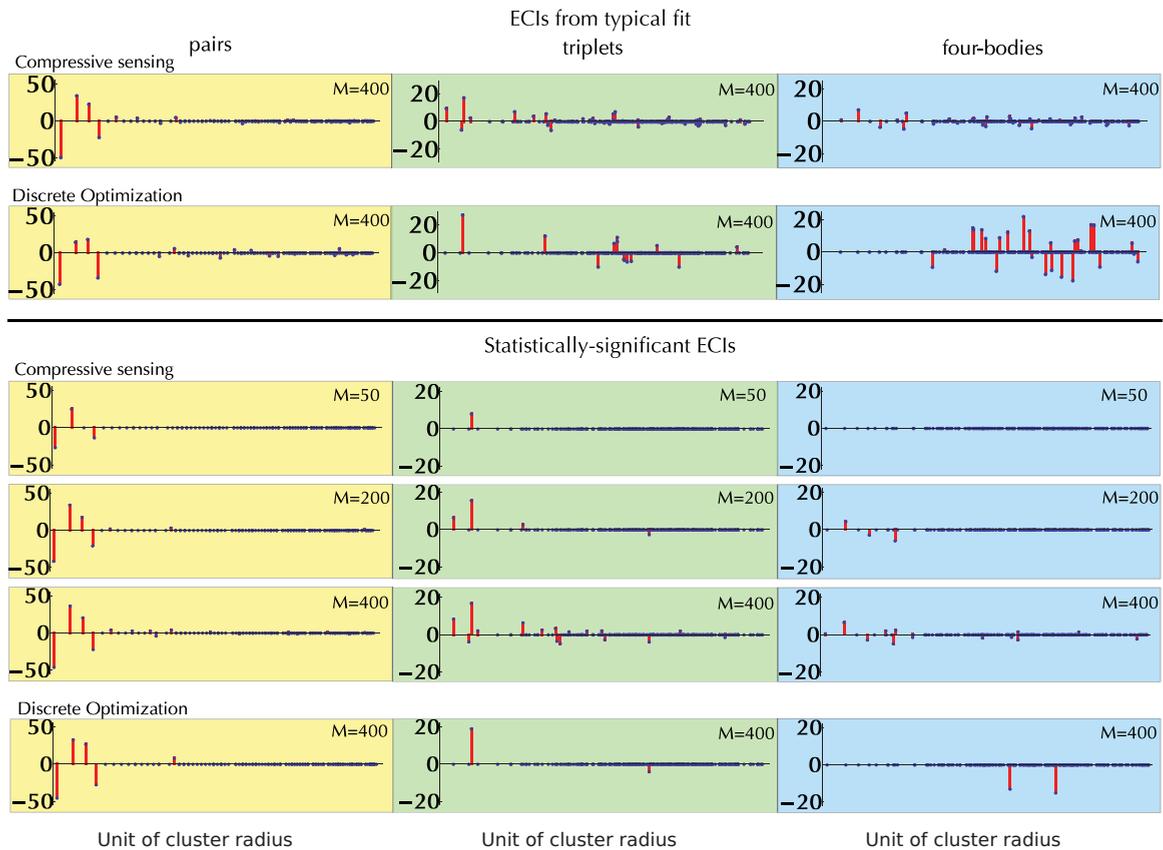
FIG. 5. (Color online) Comparison of the interaction coefficients found using the DO method implemented in ATAT software and compressive sensing. The upper pane shows a comparison of two typical fits from CS and ATAT. The lower pane shows the coefficients that were found to be statistically relevant from both methods. The x axis is the cluster radius, which is defined as the average distance from the center of mass of all cluster vertices, given as a fraction of the lattice parameter. The cluster interactions are given in meV/cluster. (Blue dots were placed on the x axis even for clusters not found to be relevant to help the reader know the ordinal number of the relevant clusters.) Physical intuition suggests that shorter-radius, fewer-vertex clusters are the most important contributors in alloy energetics. Pair interaction coefficients found by both methods are similar. As the number of vertices increases, CS finds coefficients in harmony with physical intuition, while DO finds spurious, long-ranged three- and four-body interactions. CS solutions also demonstrate a convergence to one specific solution as the size of the fitting set increases. (Note: triplets and quadruplets are shown on a scale from -20 to 20 meV, different from the scale used for the pairs.)

vertices increases (note that triplets and quadruplets are shown on a scale from -20 to 20 meV, as opposed to -50 to 50 meV for pairs). This is in harmony with long-standing claims in the CE community, and it confirms that a stable solution has been found. DO-determined clusters follow this pattern for pair clusters only. At higher vertex numbers, a typical DO fit finds nonphysical, spurious coefficients for three- and four-body interactions. The set of statistically-relevant DO coefficients appear to be lacking several important interactions, specifically short-ranged three- and four-body interactions. This indicates that (i) current DO methods are much too slow to be able to gather enough statistics to do a meaningful statistical analysis and/or (ii) current DO methods are very sensitive to the choice of the training set and fall short in their ability to identify physically relevant interactions without user guidance.

Note that the mathematical framework of CS has no knowledge of the spatial extent or geometry of the cluster functions. Remarkably, the dominant expansion coefficients, regardless of spatial extent, are efficiently retrieved using CS. In cases where a purely real-space cluster expansion fails to converge, CS may fail to construct a suitable model, but it

could be combined (as has been done with other approaches) with reciprocal-space fomulations.[30,31,59,60]

Figure 6 shows the results of a ground-state search performed by using the statistically significant $M = 400$ coefficients to predict the energies of all fcc-based superstructures up to 12 atoms. Error bars were calculated from randomly drawn sets of $M = 400$ structures. The ground-state line in this figure is consistent with first-principles data for this system, which finds the same ground states as in Fig. 6, with a few degenerate structures lying on the convex hull between $c = 0.4$ and 0.5.

This example shows that, in comparison with traditional cluster selection methods, CS is not only simpler and faster (less than a minute on a single CPU for CS versus *days* for LOOCV at $M = 400$), but also produces more physical solutions that result in a significant improvement in physical accuracy.

D. Protein folding application

We now turn to a technically much more challenging case— that of protein design in biology. Modeling the protein folding
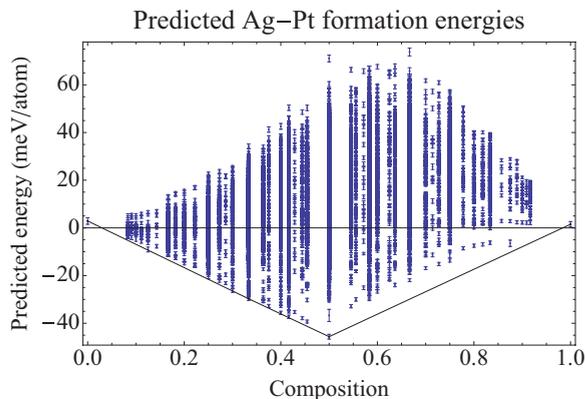
FIG. 6. (Color online) Predicted CSCE formation energies obtained using the ECI's shown in Fig. 5; error bars are standard deviation due to different random choices of $\leqslant 400$ structure subsets. Black solid line denotes the convex hull calculated from the average energies; only Ag, Ca_7Ge-type Ag_7Pt (barely, with a depth of less than 1 meV/atom), $L1_1$ AgPt, and Pt are predicted to be $T = 0$ K ground states.

energies in the zinc-finger motif represents a technically difficult test case with applications in biology.[61,62] One of the key problems in protein design is to find the sequence of amino acids (AAs), which stabilizes a particular 3D structure, or *folding*. Physics-based energy functionals are considered to be some of the most-promising methods in protein design since they link the stability of the folded 3D structure to the total free energy, accurately accounting for electrostatics, van der Waals interactions, and solvation effects. However, their use is problematic due to the astronomical number of possible AA sequences for even very short proteins. It was shown[61,62] that the CE model can be generalized to describe protein energetics, allowing very fast direct evaluation of the protein energy as a function of its sequence.

Here, we use the data from Ref. 61 for the so-called zinc-finger protein fold and closely follow exactly the same computational procedures as employed in that study. The fitting is done using a basis of approximately 76 000 clusters and energies of 60 000 AA sequences; a separate set of 4,000 AA sequence energies is used to test the predictive power of the CE model. The very large size of the problem presents a severe test to the conventional LOOCV-based model building approach, requiring running times of several weeks on parallel computers with user-supervised partial optimization.[61] We chose the highly efficient split Bregman iteration[39] for solving the basis pursuit denoising problem in Eq. (8), which allows us to perform a full optimization in approximately 30 minutes on a single 2.4 GHz Intel Xeon E5620 processor. Figure 7 shows that for the physically important negative-energy configurations, we are able to achieve an rms predictive error of 2.1 kcal/mol with 3100 model parameters, significantly better than the rms error of 2.7 kcal/mol with approximately 6000 parameters obtained using the LOOCV method in Ref. 61. Since the predictive errors are Gaussian-distributed with a mean of zero, the statistical uncertainty in the predictive error due to the finite size of the prediction set (>1000 negative-energy structures) can be calculated using standard statistical formulas for the χ^2
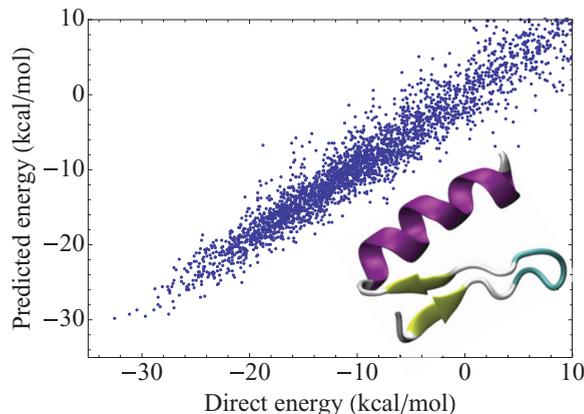

FIG. 7. (Color online) Predictive performance of CS for protein energetics in the zinc-finger structure (shown in the inset).

distribution; they are found to be less than 1% of the calculated RMSE. These results show that the computational efficiency, conceptual simplicity and physical accuracy of the ℓ_1-based minimization shows promise for future applications in protein design.

VI. CONCLUSION

In conclusion, compressive sensing can be straightforwardly adopted to build physical models that are dominated by a relatively small number of contributions drawn from a much larger underlying set of basis functions. Compressive sensing is applicable to any "sparse" basis-expansion problem, a broad class of problems in physics, chemistry and materials science. Compressive sensing allows the identification of relevant parameters from a large pool of candidates using a small number of experiments or calculations—a real paradigm shift from traditional techniques. Furthermore, many other scientific problems that do not appear to be a basis pursuit problem may be recast as one in which case CS could efficiently provide accurate and robust solutions with relatively little user input. With the huge amount of experimental and computational data in physical sciences, compressive sensing techniques represent a promising avenue for model building on many fronts including structure maps, empirical potential models, tight-binding methods, and cluster expansions for configurational energies, thermodynamics, and kinetic Monte Carlo methods.

In the arena of cluster expansion, compressive sensing provides a simple solution to two challenges: "cluster selection" and "structure selection." Cluster selection is effectively solved because compressive sensing can select clusters efficiently from a very large set (thousands or tens of thousands). Essentially, it allows the user to specify a cluster set so large that it encompasses every physically-conceivable interaction. The second challenge, structure selection, is overcome by the fact that compressive sensing *requires* that input structures simply be chosen randomly from configuration space.

---

*vidvuds@ucla.edu

[1] A. Zunger, Phys. Rev. B **22**, 5839 (1980).

[2] P. Villars, J. Less-Common Met. **92**, 215 (1983).

[3] D. Pettifor, Solid State Commun. **51**, 31 (1984).

[4] D. Pettifor, J. Phys. C **19**, 285 (1986).

[5] A. Miedema, R. Boom, and F. R. De Boer, J. Less-Common Met. **41**, 283 (1975).

[6] F. Körmann, A. Dick, B. Grabowski, B. Hallstedt, T. Hickel, and J. Neugebauer, Phys. Rev. B **78**, 033102 (2008).

[7] J. Sanchez, F. Ducastelle, and D. Gratias, Phys. A **128**, 334 (1984).

[8] D. Fontaine, Solid State Phys. **47**, 33 (1994).

[9] A. Zunger, *First-Principles Statistical Mechanics of Semiconductor Alloys and Intermetallic compounds*, Vol. 319 of NATO Advanced Study Institute, Series B: Physics, edited by P. Turchi and A. Gonis (Plenum, New York, 1994), pp. 361–419.

[10] C. C. Fischer, K. J. Tibbetts, D. Morgan, and G. Ceder, Nat. Mater. **5**, 641 (2006).

[11] J. C. Schön, K. Doll, and M. Jansen, Phys. Status Solidi B **247**, 23 (2010).

[12] T. R. Munter, D. D. Landis, F. Abild-Pedersen, G. Jones, S. Wang, and T. Bligaard, Comput. Sci.: Discovery **2**, 015006 (2009).

[13] W. Setyawan, R. M. Gaume, S. Lam, R. S. Feigelson, and S. Curtarolo, ACS Combinatorial Science **13**, 382 (2011).

[14] S. M. Woodley and R. Catlow, Nat. Mater. **7**, 937 (2008).

[15] A. P. J. Jansen and C. Popa, Phys. Rev. B **78**, 085404 (2008).

[16] T. Mueller and G. Ceder, Phys. Rev. B **80**, 024103 (2009).

[17] E. Cockayne and A. van De Walle, Phys. Rev. B **81**, 012104 (2010).

[18] E. Candès and M. Wakin, IEEE Signal Process. Mag. **25**, 21 (2008).

[19] M. AlQuraishi and H. McAdams, Proc. Natl. Acad. Sci. USA **108**, 14819 (2011).

[20] E. Candès, J. Romberg, and T. Tao, IEEE Trans. Inf. Theory **52**, 489 (2006).

[21] A. van de Walle, M. Asta, and G. Ceder, Calphad **26**, 539 (2002).

[22] A. van de Walle, Calphad **33**, 266 (2009).

[23] D. Lerch, O. Wieckhorst, G. L. W. Hart, R. W. Forcade, and S. Müller, Modell. Simul. Mater. Sci. Eng. **17**, 055003 (2009).

[24] N. A. Zarkevich and D. D. Johnson, Phys. Rev. Lett. **92**, 255702 (2004).

[25] T. Mueller and G. Ceder, Phys. Rev. B **82**, 184107 (2010).

[26] J. W. D. Connolly and A. R. Williams, Phys. Rev. B **27**, 5169 (1983).

[27] V. Blum, G. L. W. Hart, M. J. Walorski, and A. Zunger, Phys. Rev. B **72**, 165113 (2005).

[28] R. Drautz and A. Díaz-Ortiz, Phys. Rev. B **73**, 224207 (2006).

[29] A. Díaz-Ortiz, H. Dosch, and R. Drautz, J. Phys.: Condens. Matter **19**, 406206 (2007).

[30] D. B. Laks, L. G. Ferreira, S. Froyen, and A. Zunger, Phys. Rev. B **46**, 12587 (1992).

[31] R. P. Kurta, V. N. Bugaev, and A. Díaz-Ortiz, Phys. Rev. Lett. **104**, 085502 (2010).

[32] L. Thuinet and R. Besson, Appl. Phys. Lett. **100**, 251902 (2012).

[33] O. Shchyglo, A. Díaz-Ortiz, A. Udyansky, V. N. Bugaev, H. Reichert, H. Dosch, and R. Drautz, J. Phys.: Condens. Matter **20**, 045207 (2008).

[34] R. Tibshirani, J. Roy. Stat. Soc. Ser. B **58**, 267 (1996).

[35] S. Chen, D. Donoho, and M. Saunders, SIAM J. Sci. Comput. **20**, 33 (1998).

[36] E. Hale, W. Yin, and Y. Zhang, CAAM TR07-07, Rice University (2007).

[37] W. Yin, S. Osher, D. Goldfarb, and J. Darbon, SIAM J. Imaging Sci. **1**, 143 (2008).

[38] S. Boyd and L. Vandenberghe, *Convex Optimization* (Cambridge University Press, Cambridge, UK, 2004).

[39] T. Goldstein and S. Osher, SIAM J. Imaging Sci. **2**, 323 (2009).

[40] D. Donoho and X. Huo, IEEE Trans. Inf. Theory **47**, 2845 (2001).

[41] E. Candes and J. Romberg, Inverse Probl. **23**, 969 (2007).

[42] E. Candes, J. Romberg, and T. Tao, Commun. Pure Appl. Math. **59**, 1207 (2006).

[43] G. L. W. Hart and R. W. Forcade, Phys. Rev. B **77**, 224115 (2008).

[44] G. L. W. Hart and R. W. Forcade, Phys. Rev. B **80**, 014120 (2009).

[45] A. Seko, Y. Koyama, and I. Tanaka, Phys. Rev. B **80**, 165122 (2009).

[46] Teck Tan and Duane Johnson recently developed an as-yet-unpublished method for structure selection that applies fractional factorial design to the structure selection problem (private communication).

[47] I. Johnstone, Annal. stat. **29**, 295 (2001).

[48] This applies only to random numerical errors in the DFT formation energies and excludes systematic errors, such as those due to the approximate nature of the exchange-correlation functionals.

[49] In our estimation, numerical errors in the calculated DFT formation energies are only a few milli-electron-volts per atom for the case of Ag-Pt compounds considered in Sec. V B.

[50] P. Durussel and P. Feschotte, J. Alloys Compound. **239**, 226 (1996).

[51] Such a large number of structures was only chosen to test the performance of different CE methods and is several times larger than typical training set sizes used in state-of-the-art CE methods.

[52] G. Kresse and D. Joubert, Phys. Rev. B **59**, 1758 (1999).

[53] G. Kresse and J. Furthmüller, Comp. Mat. Sci. **6**, 15 (1996).

[54] P. E. Blöchl, Phys. Rev. B **50**, 17953 (1994).

[55] J. P. Perdew, K. Burke, and M. Ernzerhof, Phys. Rev. Lett. **77**, 3865 (1996).

[56] S. Froyen, Phys. Rev. B **39**, 3168 (1989).

[57] P. Q. Nguyen and D. Stehlé, ACM Trans. Algorithms **5**, 1 (2009).

[58] L. J. Nelson, S. Curtarolo, and G. L. W. Hart, Phys. Rev. B **85**, 054203 (2012).

[59] O. Shchyglo, A. Diaz-Ortiz, A. Udyansky, V. N. Bugaev, H. Reichert, H. Dosch, and R. Drautz, J. Phys.: Condens. Matter **20**, 045207 (2008).

[60] L. Holliger and R. Besson, Phys. Rev. B **83**, 174202 (2011).

[61] F. Zhou, G. Grigoryan, S. Lustig, A. Keating, G. Ceder, and D. Morgan, Phys. Rev. Lett. **95**, 148103 (2005).

[62] G. Grigoryan, A. Reinke, and A. Keating, Nature (London) **458**, 859 (2009).