Learning the Physics of Pattern Formation from Images

Hongbo Zhao[1], Brian D. Storey,[2,3] Richard D. Braatz[1], and Martin Z. Bazant[1,4]

[1]*Department of Chemical Engineering, Massachusetts Institute of Technology 77 Massachusetts Avenue,*
Cambridge, Massachusetts 02139, USA
[2]*Toyota Research Institute, Cambridge, Massachusetts 02139, USA*
[3]*Olin College, Needham, Massachusetts 02492, USA*
[4]*Department of Mathematics, Massachusetts Institute of Technology,*
77 Massachusetts Avenue, Cambridge, Massachusetts 02139, USA

Using a framework of partial differential equation-constrained optimization, we demonstrate that multiple constitutive relations can be extracted simultaneously from a small set of images of pattern formation. Examples include state-dependent properties in phase-field models, such as the diffusivity, kinetic prefactor, free energy, and direct correlation function, given only the general form of the Cahn-Hilliard equation, Allen-Cahn equation, or dynamical density functional theory (phase-field crystal model). Constraints can be added based on physical arguments to accelerate convergence and avoid spurious results. Reconstruction of the free energy functional, which contains nonlinear dependence on the state variable and differential or convolutional operators, opens the possibility of learning nonequilibrium thermodynamics from only a few snapshots of the dynamics.

Pattern formation is widely observed in physical, chemical, and biological systems. One category of pattern-forming partial differential equations (PDEs) are those driven by gradients of a nonconvex energy landscape [1]. This category includes the well-known Allen-Cahn equation, Cahn-Hilliard equation [2,3], and Swift-Hohenberg equation [4], which have been extended and found wide applications in phase-field [5], phase-field-crystal [6], and nonequilibrium chemical thermodynamic [7,8] models. These PDEs homogenize the details of molecular degrees of freedom [9] and depend on parametrized constitutive relations such as free energy landscape, transport properties, and/or reaction kinetics, which encode spatiotemporal dynamics at the relevant time and length scales [10].

Existing methods to obtain constitutive relations face several challenges. Bottom-up approaches, such as semi-empirical microscopic models [11] and first-principles calculations [12–14], have difficulty spanning length and timescales, although recent scale-bridging methods [15–17] have shown promise. Top-down approaches based on fitting emergent properties, such as open-circuit voltage [18], are often infeasible or obscured by thermodynamic instabilities [19,20]. Empirical models fit to phase diagrams [21,22] may not be able to predict the dynamics far from equilibrium.

In this Letter we demonstrate the possibility of learning constitutive relations directly from experimental images, which has become increasingly relevant with the growing volume and quality of *in situ* and *in operando* images

[23–27]. Our approach extends the inversion problem based on still images [28–30] and low-dimensional representation of images [31,32] and makes use of all the pixels in 2D video data to infer the physics of pattern formation.

In the literature of PDE discovery, the model either comes from prior physical knowledge or is represented by a library of operators, identified using sparsity-promoting, nonparametric, and other techniques of dynamical systems identification [33–37]. The model is then trained by minimizing the residual of the PDE [31,32,34,38–41], the error of model prediction [42,43], or both [37,44–48].

In contrast to this literature, we constrain the form of the governing PDEs for physical interpretability, while allowing maximum flexibility by incorporating differential or convolutional operators and nonlinearity in the constitutive relationship. The methods above, which directly relate operators, rely on high resolution images for performing spatial and temporal derivatives, while it is often unknown *a priori* whether the given resolution is sufficient for taking the derivative due to the nonlinear nature of the PDE. Additionally, quantitative matching of the patterns is not guaranteed. To overcome these difficulties, we start from the given initial condition and evaluate the model on a sufficiently fine grid through adaptive time stepping and minimize the norm of the pixelwise error. The approach relies on the well-developed methods in PDE-constrained optimization [49–51].

Our canonical example is the Cahn-Hilliard equation, describing phase separation driven by chemical potential gradients [2,3],

$$\frac{\partial c}{\partial t} = \nabla \cdot \left(D(c) c \nabla \frac{\delta F}{\delta c} \right), \tag{1}$$

where $c(x, t)$ is a conserved concentration field, $D(c)$ is the diffusivity, and $\delta F / \delta c$ is the diffusive chemical potential or variational derivative of the free energy [7]. We also consider the Allen-Cahn equation for nonconservative order parameters with a state-dependent kinetic prefactor $R_0(c)$,

$$\frac{\partial c}{\partial t} = -R_0(c) \frac{\delta F}{\delta c}, \tag{2}$$

which is the linearized form of the general Allen-Cahn reaction model [7,8], where the reaction rate has a nonlinear dependence on the thermodynamic driving force (or chemical affinity). For both equations, the chemical potential is defined by

$$\frac{\delta F}{\delta c} = \mu_h(c) - \kappa \nabla^2 c, \tag{3}$$

where $\mu_h(c)$ is the homogeneous part of the chemical potential and $-\kappa \nabla^2 c$ is the gradient correction to the chemical potential and gives rise to a diffuse interface between phases.

We seek to infer the chemical potential function over the entire range of composition, including within the spinodal region [$\mu_h'(c) < 0$, where prime denotes differentiation], from concentration profiles recorded far from equilibrium, since the spinodal region is unstable at equilibrium. Within the mean field approximation, a two-component mixture has an enthalpy of mixing with quadratic dependence on concentration [2]. More generally, the enthalpy can be expressed in terms of higher order polynomials [22]. If the concentration is normalized by its maximum value, the following numerically well-conditioned parametrization,

$$\mu_h(c) = \log \frac{c}{1 - c} + \sum_{n=1}^{N} a_n P_n(c), \tag{4}$$

uses the ideal entropy of mixing to bound the concentration $(0 < c < 1)$, where the P_n are Legendre polynomials defined on the interval [0,1], and a_n are the coefficients to be determined. We also represent $\log D(c)$ and $\log R_0(c)$ in a basis of Legendre polynomials.

The inverse problem is defined by the minimization with respect to parameters $\mathbf{p}$ of the objective function, defined by the L_2 norm of the difference between the data and the model prediction,

$$L(\mathbf{p}) = \frac{1}{2} \sum_{i=1}^{M} \int d\mathbf{r} [c(t_i, \mathbf{r}; \mathbf{p}) - c_{\text{data}}(t_i, \mathbf{r})]^2, \tag{5}$$

constrained by the PDE, where the integral is taken over the position $\mathbf{r}$ of the entire image; M is the total number of the training images taken at different time t_i; c_{data} is the concentration field converted from the gray values of the image pixels through calibration; and $c(t, \mathbf{r}; \mathbf{p})$ is the concentration field predicted by the PDE.

The gradient-based optimization of L can be performed with two approaches [52], forward sensitivity analysis (FSA), which computes the model sensitivity $\partial c / \partial \mathbf{p}$ along with the model evaluation, and adjoint sensitivity analysis (ASA), which obtains the gradient $\partial L / \partial \mathbf{p}$ by solving the adjoint linearized PDE [53–55]. FSA can also provide the Gauss-Newton approximation of the Hessian of the objective function [56], which becomes increasingly accurate as p approaches the truth,

$$\frac{\partial^2 L}{\partial p_j \partial p_k} \approx \sum_{i=1}^{M} \int d\mathbf{r} \frac{\partial c}{\partial p_j} \frac{\partial c}{\partial p_k}. \tag{6}$$

See the Supplemental Material [57] for the derivation of the algorithm. In our examples, we find that optimizing the three quantities simultaneously using the approximated Hessian [Eq. (6)] and trust-region algorithm [56] via FSA is faster.

To test this approach, we generate simulated data for the evolution of the Cahn-Hilliard and Allen-Cahn equations, for which the true model parameters are known, with the boundary conditions of zero normal gradient for both and zero flux for the former [7]. From the full simulation we take 5 high-resolution snapshots of the relaxation process and apply our method to learn the constitutive law. For all 4 cases (a)–(d) shown in Fig. 1, the enthalpic part of $\mu_h(c)$, $\log R_0(c)$, and $\log D(c)$ are represented by 10 Legendre polynomials, while the known truths are $\mu_h(c) = \log c(1 - c)^{-1} + 3(1 - 2c)$, $R_0(c) = c(1 - c)$, and $D(c) = 1 - c$. While the true $R_0(c)$ and $D(c)$ cannot be exactly recovered, they satisfy the thermodynamic constraint that they are positive [8,9]. The gradient energy coefficient κ is free to change in the optimization. The form of the equations shows that the constitutive quantities that uniquely determine the PDE can be obtained through proper scaling, that is, $D d\mu_h / d \ln c$, or $k\mu_h'(c)$ sets the timescale, while $\sqrt{\kappa / \mu_h'(c)}$ sets the length scale.

We use the first image as the initial condition for the simulation. The images used in case (b) are those for case (a) with 10% added Gaussian noise. The image at $t = 0$ is too noisy to be useful as the initial condition so we start from $t = 0.015$ for case (b). For case (a)–(c), we impose the known boundary conditions while for case (d) where only a part of the data within a larger domain is captured, the value c and the normal gradient $\mathbf{n} \cdot \nabla c$ at the boundaries of the images are smoothly interpolated in time and directly imposed as the time dependent boundary conditions of the Cahn-Hilliard model. Starting from a non-pattern-forming initial guess for the parameters, qualitative
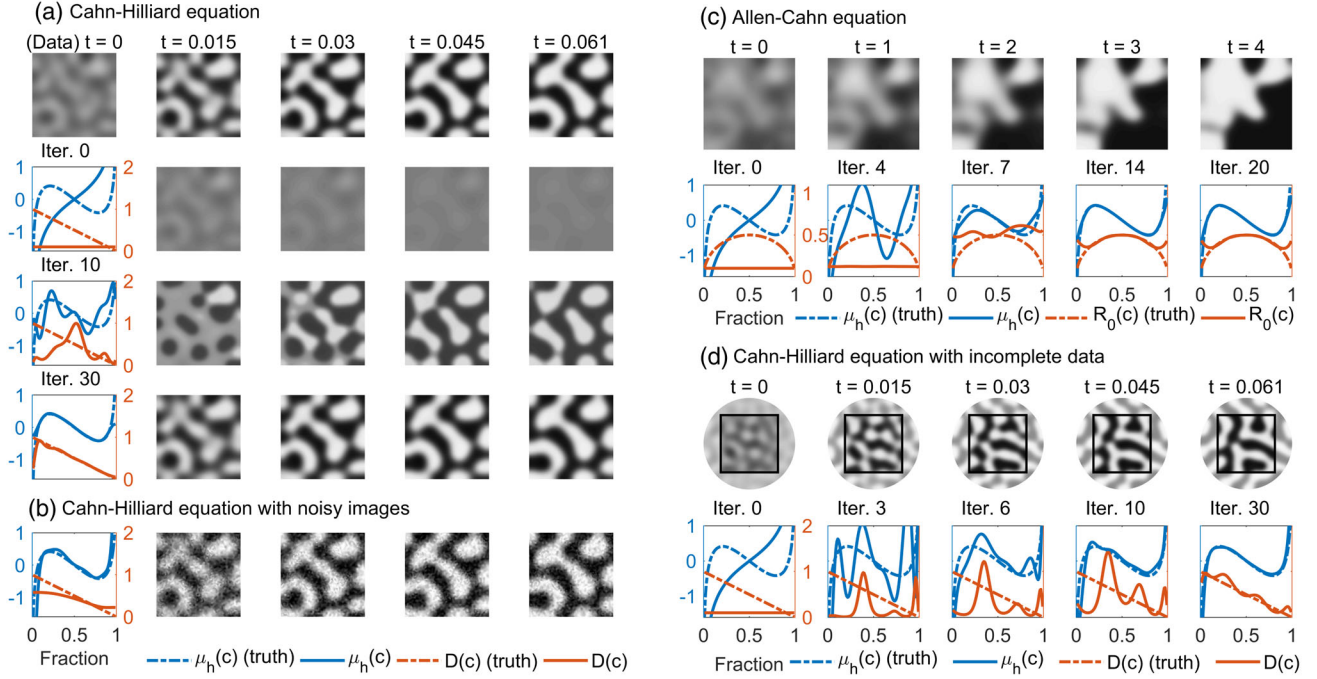
FIG. 1.   Learning phase-field models from images. (a) Inversion of chemical potential $\mu_h(c)$ and diffusivity $D(c)$ for the Cahn-Hilliard equation. At the top row are the 5 images provided for training. The following rows show $\mu_h(c)$ and $D(c)$ and the model prediction at each iteration of the optimization. Iteration 0 is the initial guess. (b) Inversion of $\mu_h(c)$ and $D(c)$ for the Cahn-Hilliard equation using 4 noisy images shown. (c) Inversion of chemical potential and kinetic prefactor $R_0(c)$ for the Allen-Cahn equation. (c) Inversion of $\mu_h(c)$ and $D(c)$ for the Cahn-Hilliard equation based on images taken from a part of a larger domain. Except for iteration 0, the functions shown at other iterations are scaled such that κ is the same as the known truth for comparison.

agreement for the images is achieved within only a few iterations, and convergence to the known truth is achieved within tens of iterations. For instance, for case (a), from the 20th iteration onward, the solver makes minor adjustments to $\mu_h(c)$ and $D(c)$ near $c = 0$ and $c = 1$, which only changes the near-zero objective function slightly as shown in Fig. S3 of the Supplemental Material [57].

Figure 1(b) is an example of inversion based on noisy images. To obtain the result, we added regularization [52] to suppress the high-frequency oscillation that would otherwise appear in the overfitted $\mu_h(c)$ and $\log D(c)$. Despite the penalty on higher order polynomials, $\mu_h(c)$ in the spinodal region is captured, while the less sensitive $D(c)$ has a dampened variation in c. See Fig. S7 [57] for how the result depends on the level of noise and the regularization parameter. As shown in Figs. 1(a), 1(b), and 1(d), despite some discrepancies between the inferred and true model, the objective function is effectively zero. The model is insensitive to $D(c)$ and $R_0(c)$ near $c = 0$ and $c = 1$, where the system is too close to equilibrium to reveal its dynamic properties. The sensitivity analysis in the Supplemental Material [57] also confirms that the spinodal pattern are mainly determined by $d\mu_h/dc$ in the spinodal region.

Since real images may be noisy, incomplete, or low resolution, we more fully analyze the robustness and scaling of the inversion method in the Supplemental

Material [57]. It is found that constitutive laws can still be obtained even when only the initial and final frames or severely down-sampled images are provided as training data.

Next we apply the inversion method to a more general, nonlocal form for the free energy functional with long-range interactions, which is used in dynamical density functional theory (DDFT) to describe crystallization at the atomic scale [6]. For small variations around a uniform concentration c_0, we neglect the state-dependent mobility and expand the free energy functional in terms of the normalized order parameter $\eta = c/c_0 - 1$ to arrive at $\partial\eta/\partial t = M\nabla^2(\delta F/\delta\eta)$, where the chemical potential in dimensionless form is

$$\frac{\delta F}{\delta\eta} = \mu_h(\eta) - \int d\tilde{\mathbf{r}} C_2(|\mathbf{r} - \tilde{\mathbf{r}}|)\eta(\tilde{\mathbf{r}}), \qquad (7)$$

where $C_2(|\mathbf{r} - \tilde{\mathbf{r}}|)$ is the direct correlation function. The ideal entropy (or its Taylor expansion) is typically used for $\mu_h(\eta)$, but we keep a general form for the purpose of extracting it through inversion. A uniform field becomes linearly unstable to perturbations when $\sigma = -M|\mathbf{k}|^2(\mu_h'(c) - \hat{C}_2(|\mathbf{k}|)) > 0$, where $\hat{C}_2(\mathbf{k})$ is the Fourier transform of $C_2(\mathbf{r})$. The lowest order, isotropic, and analytic form that satisfies this condition and $\hat{C}_2(\infty) \leq 0$ is

$\hat{C}_2(\mathbf{k}) = h - [1 - (|\mathbf{k}|/k_0)^2]^2$, where h is a positive constant and k_0 is the characteristic wave number. The resulting form is mathematically equivalent to the Swift-Hohenberg free energy [4]. To capture the evolution of an isotropic pattern (e.g., in crystallization or Rayleigh-Bénard instability) from an unknown direct correlation function, we use the quartic approximation,

$$\hat{C}_2(\mathbf{k}) = b_0 + b_1|\mathbf{k}|^2 + b_2|\mathbf{k}|^4, \qquad (8)$$

with $b_2 < 0$, which corresponds to $b_0 - b_1\nabla^2 + b_2\nabla^4$ in real space. This approximation is also known as the phase field crystal model.

Figure 2(a) shows the inversion of the free energy [$\mu_h(\eta)$ and $\hat{C}_2(\mathbf{k})$ simultaneously] from a nucleation event with

(a) Data

(b) Quartic approximation

(c) Quartic approximation result

(d) Hermite function basis function

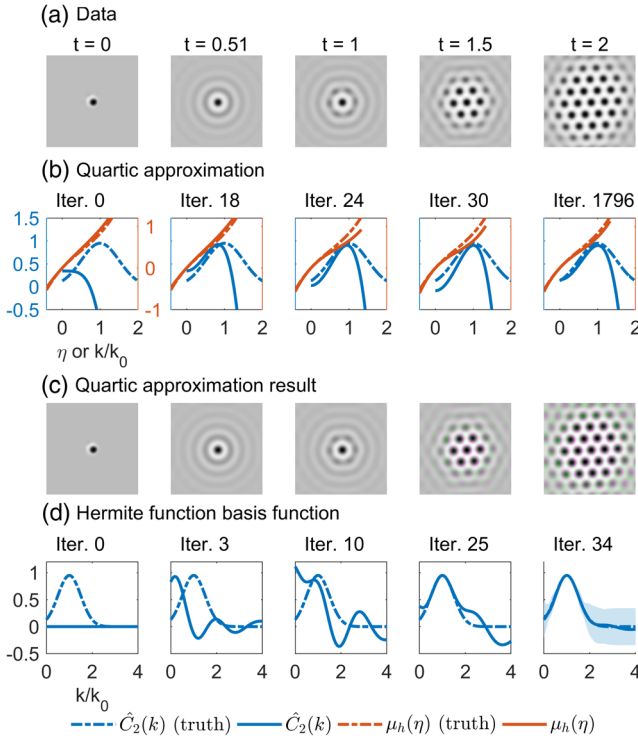- - - - $\hat{C}_2(k)$ (truth)   —— $\hat{C}_2(k)$ - - - - $\mu_h(\eta)$ (truth)   —— $\mu_h(\eta)$

FIG. 2. Learning dynamic density functional theory (DDFT) from images. (a) Five images of the nucleation event used as training data. (b) Inversion of $\mu_h(\eta)$ and $\hat{C}_2(|\mathbf{k}|)$. $\hat{C}_2(|\mathbf{k}|)$ is represented by a quartic approximation [Eq. (8)]. $\mu_h(\eta)$ is represented by up to third order polynomials. The linear term in $\mu_h(\eta)$ and the constant term in $\hat{C}_2(|\mathbf{k}|)$ are indistinguishable. Hence, they are adjusted such that $\mu_h'(0) = 1$ for comparison with the known truth. (c) The model prediction at the final iteration in (b) using the quartic approximation. The last two images are anaglyphs that highlight their difference with the data. Magenta indicates the model value is greater than data while green means the opposite. (d) Inversion of $\hat{C}_2(|\mathbf{k}|)$ using Eq. (9). The shaded region in the last image shows the estimated error of $\hat{C}_2(|\mathbf{k}|)$ inferred by adding $\epsilon \sum_i d_i^2$ to the objective function normalize by area, where $\epsilon = 0.01$, using 100 Hermite functions.

periodic boundary condition. The true model [58] that is known to give rise to hexagonal lattice is $\mu_h(\eta) = \eta - \eta^2/2 + \eta^3/3$ and $\hat{C}_2(\mathbf{k}) = 0.95 \exp[-(|\mathbf{k}| - k_0)^2/2\alpha^2]$, where $\alpha/k_0 = 0.5$. Despite the large difference between the true $\hat{C}_2(\mathbf{k})$ and the quartic approximation shown in Fig. 2(b), especially at high wave numbers, the dynamics and hexagonal pattern are retained. The early-stage pattern is nearly identical. The late-stage pattern is more diffuse than the data [as shown by the colored difference in the last two image of Fig. 2(c)] because high wave number components decay faster in the quartic approximation.

Since the early-stage pattern grows exponentially at $e^{\sigma t}$, previous findings [59,60] and the result above show that $\hat{C}_2(\mathbf{k})$ is described sufficiently well around the peak by k_0, $\hat{C}_2(k_0)$, and $\hat{C}_2''(k_0)$. The sensitivity analysis in the Supplemental Material [57] confirms that $\hat{C}_2(\mathbf{k})$ cannot be identified at large $|\mathbf{k}|$, since high wave number components decay rapidly in time due to the Laplacian operator in the PDE. Therefore, we emphasize that finding $\hat{C}_2(\mathbf{k})$ over the entire range of $\mathbf{k}$ is an ill-posed problem.

From liquid state theory [61], the structure factor $\hat{S}(\mathbf{k})$ approaches 1 and $\hat{C}_2(\mathbf{k})$ approaches zero as $|\mathbf{k}| \to \infty$, where $1/\hat{S}(\mathbf{k}) = 1 - \hat{C}_2(\mathbf{k})$. Since $\hat{C}_2(\mathbf{k})$ at large $|\mathbf{k}|$ is not informed by data, we include the prior knowledge by choosing basis functions that approach zero as $|\mathbf{k}| \to \infty$ [62],

$$\hat{C}_2(\mathbf{k}) = \sum_{n=0}^{N} d_n (2^n n! \sqrt{\pi})^{-1/2} e^{-|\mathbf{k}|^2/2} H_n(|\mathbf{k}|), \qquad (9)$$

where H_n is the Hermite polynomial. For a faster convergence, $\mathbf{k}$ is scaled by the wave number k_0, which can be measured from the pattern. Figure 2(d) shows that using Eq. (9) ($N = 10$), converges more quickly to the correct $\hat{C}_2$ [Here we fix $\mu_h(\eta)$. See Fig. S9 [57] for solving for $\mu_h(\eta)$ and $\hat{C}_2$ simultaneously using Eq. (9) and the corresponding model prediction]. In comparison, b_1 and b_2 are highly coupled in the quartic approximation, and the convergence is very slow. Ill posedness is resolved here using a limited number of the basis functions. The shaded region in Fig. 2(d), or the estimated uncertainty in $\hat{C}_2(\mathbf{k})$ confirms high sensitivity around k_0 and high uncertainty at large $|\mathbf{k}|$.

Equation (7) is a generic form for the gradient of a Lyapunov function. Our method could be applied to learning a wide variety of other physical models, such as the nonlocal Cahn-Hilliard equation [63], the Hopfield model [64], and the BSK model of electrostatic correlations [65–68]. In the strain-induced anisotropic spinodal decomposition [20,69], while the displacement field is not directly observed, it is known to be coupled to the concentration field through mechanical equilibrium. The inversion technique used in this Letter can be extended to finding these

hidden variables, as well as models with generalized nonlocal constitutive relations [70].

We envision that the inversion method can be used to infer thermodynamic and kinetic properties from *in operando* images of phase-separating materials [71]. *In situ* measurement of crystallization in reciprocal space [72,73], together with phase information [74], also provides valuable information about the intermolecular interaction in conjunction with the DDFT model. With experimental data, Bayesian inference through Monte Carlo methods can quantify the uncertainty[75,76]. As shown in the Supplemental Material [57], the inversion result may be influenced by noise, initial conditions, resolution and whether the particular frames are informative, so to avoid suboptimal solutions, a global or stochastic optimization scheme may also be considered.

In summary, we have demonstrated that multiple constitutive relations embedded in pattern-forming PDE models can be extracted from a small set of images. Compared to other methods of learning PDEs, physics can be learned from fewer low-resolution and noisy images; quantitative matching of the patterns can be achieved; and the PDEs can include nonlinear or nonlocal constitutive relationships. By inverting image data to obtain the kernel of a convolution, we show that the learning of differential operators can also be extended to general integrodifferential operators. The ill posedness of learning integral operators can be regularized by symmetry, convolution, or isotropy via physical arguments.

[1] M. C. Cross and P. C. Hohenberg, Rev. Mod. Phys. **65**, 851 (1993).

[2] J. W. Cahn and J. E. Hilliard, J. Chem. Phys. **28**, 258 (1958).

[3] J. W. Cahn, J. Chem. Phys. **42**, 93 (1965).

[4] J. Swift and P. C. Hohenberg, Phys. Rev. A **15**, 319 (1977).

[5] L.-Q. Chen, Annu. Rev. Mater. Res. **32**, 113 (2002).

[6] H. Emmerich, H. Löwen, R. Wittkowski, T. Gruhn, G. I. Tóth, G. Tegze, and L. Gránásy, Adv. Phys. **61**, 665 (2012).

[7] M. Z. Bazant, Acc. Chem. Res. **46**, 1144 (2013).

[8] M. Z. Bazant, Faraday Discuss. **199**, 423 (2017).

[9] H. B. Callen, *Thermodynamics and an Introduction to Thermostatistics* (John Wiley & Sons, New York, 1985).

[10] B. B. Machta, R. Chachra, M. K. Transtrum, and J. P. Sethna, Science **342**, 604 (2013).

[11] I. Ho and G. B. Stringfellow, Appl. Phys. Lett. **69**, 2701 (1996).

[12] G. Ceder and A. Van der Ven, Electrochim. Acta **45**, 131 (1999).

[13] F. Zhou, T. Maxisch, and G. Ceder, Phys. Rev. Lett. **97**, 155704 (2006).

[14] R. Malik, F. Zhou, and G. Ceder, Nat. Mater. **10**, 587 (2011).

[15] R. Gómez-Bombarelli, J. N. Wei, D. Duvenaud, J. M. Hernández-Lobato, B. Sánchez-Lengeling, D. Sheberla, J. Aguilera-Iparraguirre, T. D. Hirzel, R. P. Adams, and A. Aspuru-Guzik, ACS Cent. Sci. **4**, 268 (2018).

[16] L. Zhang, J. Han, H. Wang, R. Car, and W. E, Phys. Rev. Lett. **120**, 143001 (2018).

[17] G. H. Teichert, A. R. Natarajan, A. Van der Ven, and K. Garikipati, Comput. Methods Appl. Mech. Eng. **353**, 201 (2019).

[18] H. Lukas, S. G. Fries, and B. Sundman, *Computational Thermodynamics* (Cambridge University Press, Cambridge, England, 2007).

[19] T. R. Ferguson and M. Z. Bazant, Electrochim. Acta **146**, 89 (2014).

[20] D. A. Cogswell and M. Z. Bazant, ACS Nano **6**, 2215 (2012).

[21] T. Cool, A. Bartol, M. Kasenga, K. Modi, and R. E. García, CALPHAD: Comput. Coupling Phase Diagrams Thermochem. **34**, 393 (2010).

[22] O. Redlich and A. T. Kister, Ind. Eng. Chem. **40**, 345 (1948).

[23] J. N. Weker and M. F. Toney, Adv. Funct. Mater. **25**, 1622 (2015).

[24] F. Lin, Y. Liu, X. Yu, L. Cheng, A. Singer, O. G. Shpyrko, H. L. Xin, N. Tamura, C. Tian, T. C. Weng, X. Q. Yang, Y. S. Meng, D. Nordlund, W. Yang, and M. M. Doeff, Chem. Rev. **117**, 13123 (2017).

[25] A. V. Bayles, C. S. Valentine, T. Überrück, S. P. O. Danielsen, S. Han, M. E. Helgeson, and T. M. Squires, Phys. Rev. X **9**, 011048 (2019).

[26] Y. Guo, R. B. Smith, Z. Yu, D. K. Efetov, J. Wang, P. Kim, M. Z. Bazant, and L. E. Brus, J. Phys. Chem. Lett. **7**, 2151 (2016).

[27] S. V. Kalinin, B. G. Sumpter, and R. K. Archibald, Nat. Mater. **14**, 973 (2015).

[28] R. Seemann, S. Herminghaus, and K. Jacobs, Phys. Rev. Lett. **86**, 5534 (2001).

[29] S. Herminghaus, K. Jacobs, K. Mecke, J. Bischot, A. Fery, M. Ibn-Elhaj, and S. Schlagowski, Science **282**, 916 (1998).

[30] A. Y. Borisevich, A. N. Morozovska, Y.-M. Kim, D. Leonard, M. P. Oxley, M. D. Biegalski, E. A. Eliseev, and S. V. Kalinin, Phys. Rev. Lett. **109**, 065702 (2012).

[31] H. U. Voss, P. Kolodner, M. Abel, and J. Kurths, Phys. Rev. Lett. **83**, 3422 (1999).

[32] K. H. Andersen, M. Abel, J. Krug, C. Ellegaard, L. R. Søndergaard, and J. Udesen, Phys. Rev. Lett. **88**, 234302 (2002).

[33] S. L. Brunton, J. L. Proctor, J. N. Kutz, and W. Bialek, Proc. Natl. Acad. Sci. U.S.A. **113**, 3932 (2016).

[34] S. H. Rudy, S. L. Brunton, J. L. Proctor, and J. N. Kutz, Sci. Adv. **3**, e1602614 (2017).

[35] Z. Wang, X. Huan, and K. Garikipati, Comput. Methods Appl. Mech. Eng. **356**, 44 (2019).

[36] S. Maddu, B. L. Cheeseman, I. F. Sbalzarini, and C. L. Müller, arXiv:1907.07810.

[37] H. U. Voss, J. Timmer, and J. Kurths, Int. J. Bifurcation Chaos Appl. Sci. Eng. **14**, 1905 (2004).

[38] S. Rudy, A. Alla, S. L. Brunton, and J. N. Kutz, SIAM J. Appl. Dyn. Syst. **18**, 643 (2019).

[39] S. L. Brunton and J. N. Kutz, J. Phys. Mater. **2,** 044002 (2019).

[40] H. Voss, M. J. Bünner, and M. Abel, Phys. Rev. E **57,** 2820 (1998).

[41] M. Bär, R. Hegger, and H. Kantz, Phys. Rev. E **59,** 337 (1999).

[42] Z. Long, Y. Lu, X. Ma, and B. Dong, in *Proceedings of the 35th International Conference on Machine Learning, PMLR 80:3208-3216* (2018) [arXiv:1710.09668].

[43] J. M. Fullana, M. Rossi, and S. Zaleski, Physica (Amsterdam) **103D,** 564 (1997).

[44] R. Tipireddy, P. Perdikaris, P. Stinis, and A. Tartakovsky, arXiv:1904.04058.

[45] M. Raissi, P. Perdikaris, and G. E. Karniadakis, J. Comput. Phys. **348,** 683 (2017).

[46] M. Raissi and G. E. Karniadakis, J. Comput. Phys. **357,** 125 (2018).

[47] M. Raissi, P. Perdikaris, and G. E. Karniadakis, J. Comput. Phys. **378,** 686 (2019).

[48] T. G. Müller and J. Timmer, Int. J. Bifurcation Chaos Appl. Sci. Eng. **14,** 2053 (2004).

[49] M. Hinze, R. Pinnau, M. Ulbrich, and S. Ulbrich, *Optimization with PDE Constraints*, Mathematical Modelling: Theory and Applications Vol. 23 (Springer Netherlands, Dordrecht, 2009).

[50] S. Molesky, Z. Lin, A. Y. Piggott, W. Jin, J. Vucković, and A. W. Rodriguez, Nat. Photonics **12,** 659 (2018).

[51] A. Jain, J. A. Bollinger, and T. M. Truskett, AIChE J. **60,** 2732 (2014).

[52] A. Tarantola, *Inverse Problem Theory: Methods for Data Fitting and Model Parameter Estimation* (SIAM, Philadelphia, 2005).

[53] T. Maly and L. R. Petzold, Applied Numerical Mathematics **20,** 57 (1996).

[54] Y. Cao, S. Li, and L. Petzold, J. Comput. Appl. Math. **149,** 171 (2002).

[55] Y. Cao, S. Li, L. Petzold, and R. Serban, SIAM J. Sci. Comput. **24,** 1076 (2003).

[56] R. Fletcher, *Practical Methods of Optimization* (John Wiley & Sons, Ltd, Chichester, West Sussex England, 2000).

[57] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevLett.124.060201 for more details on the numerical methods for solving inverse problems, specifically the equations for forward sensitivity analysis and adjoint sensitivity analysis of PDEs. Analytical formulas of the PDE sensitivity are presented to analyze the identifiability of the unknown constitutive relations.

Additionally, we demonstrate the robustness of the inversion method with more examples of various magnitudes of noise, temporal and spatial resolutions and boundary conditions.

[58] M. Greenwood, N. Provatas, and J. Rottler, Phys. Rev. Lett. **105,** 045702 (2010).

[59] A. Jaatinen, C. V. Achim, K. R. Elder, and T. Ala-Nissila, Phys. Rev. E **80,** 031602 (2009).

[60] C. M. Bender and S. A. Orszag, *Advanced Mathematical Methods for Scientists and Engineers I* (Springer New York, New York, NY, 1999).

[61] J.-P. Hanse and I. R. McDonald, *Theory of Simple Liquids* (Elsevier, New York, 2013).

[62] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions: With Formulas, Graphs, and Mathematical Tables* (United States Department of Commerce, National Institute of Standards and Technology, 1965).

[63] R. Choksi, M. A. Peletier, and J. F. Williams, SIAM J. Appl. Math. **69,** 1712 (2009).

[64] J. J. Hopfield, Proc. Natl. Acad. Sci. U.S.A. **81,** 3088 (1984).

[65] M. Z. Bazant, B. D. Storey, and A. A. Kornyshev, Phys. Rev. Lett. **106,** 046102 (2011).

[66] R. P. Misra, J. P. de Souza, D. Blankschtein, and M. Z. Bazant, Langmuir **35,** 11550 (2019).

[67] J. P. de Souza and M. Z. Bazant, arXiv:1902.05493.

[68] M. McEldrew, Z. A. H. Goodwin, A. A. Kornyshev, and M. Z. Bazant, J. Phys. Chem. Lett. **9,** 5840 (2018).

[69] A. G. Khachaturyan, *Theory of Structural Transformations in Solids* (John Wiley & Sons, New York, 2013).

[70] J. Słomka and J. Dunkel, Proc. Natl. Acad. Sci. U.S.A. **114,** 2119 (2017).

[71] J. Lim, Y. Li, D. H. Alsem, H. So, S. C. Lee, P. Bai, D. A. Cogswell, X. Liu, N. Jin, Y. S. Yu, N. J. Salmon, D. A. Shapiro, M. Z. Bazant, T. Tyliszczak, and W. C. Chueh, Science **353,** 566 (2016).

[72] M. C. Weidman, D. M. Smilgies, and W. A. Tisdale, Nat. Mater. **15,** 775 (2016).

[73] J. J. Geuchies, C. Van Overbeek, W. H. Evers, B. Goris, A. De Backer, A. P. Gantapara, F. T. Rabouw, J. Hilhorst, J. L. Peters, O. Konovalov, A. V. Petukhov, M. Dijkstra, L. D. Siebbeles, S. Van Aert, S. Bals, and D. Vanmaekelbergh, Nat. Mater. **15,** 1248 (2016).

[74] J. Miao, T. Ishikawa, I. K. Robinson, and M. M. Murnane, Science **348,** 530 (2015).

[75] T. Cui, K. J. Law, and Y. M. Marzouk, J. Comput. Phys. **304,** 109 (2016).

[76] A. M. Stuart, Acta Numer. **19,** 451 (2010).