Learning the Physics of Pattern Formation from Images

Hongbo Zhao,1 Brian D. Storey,2,3 Richard D. Braatz,1,4 and Martin Z. Bazant1,4

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

(Received 24 September 2019; revised manuscript received 9 December 2019; accepted 21 January 2020; published 14 February 2020)

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.

DOI: 10.1103/PhysRevLett.124.060201

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 parameterized 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].
\[
\frac{\partial c}{\partial t} = \nabla \cdot \left( D(c) \nabla \frac{\delta F}{\delta c} \right),
\]

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 \(R_0(c)\),

\[
\frac{\partial c}{\partial t} = -R_0(c) \frac{\delta F}{\delta c},
\]

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

\[
\frac{\delta F}{\delta c} = \mu_h(c) - \kappa \nabla^2 c,
\]

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 \([\mu_h(c) < 0, \text{where prime denotes differentiation}],\) 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,

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

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 \(p\) of the objective function, defined by the \(L_2\) norm of the difference between the data and the model prediction,

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

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 \(t_i\); \(c_{\text{data}}\) 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,

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

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 \(\mu_h(c) = \log c(1-c)^{-1} + 3(1-2c), \quad R_0(c) = c(1-c), \quad \text{and} \quad D(c) = 1-c\). 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 \(\kappa\) 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, \(Dd\mu_h/d\ln c\), 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
and 1(d), despite some discrepancies between the inferred regularization parameter. As shown in Figs. 1(a), 1(b), how the result depends on the level of noise and the true model, the objective function is effectively zero. Despite the penalty on higher order polynomials, $\mu_h(c)$ in the spinodal region is captured, while the less sensitive region. 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 $D(c)$. 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 $\delta F/\delta \eta = M\nabla^2(\delta F/\delta \eta)$, where the chemical potential in dimensionless form is

$$\frac{\delta F}{\delta \eta} = \mu_h(\eta) - \int d\mathbf{r} C_2(|\mathbf{r} - \mathbf{r}|)\eta(\mathbf{r}).$$

where $C_2(|\mathbf{r} - \mathbf{r}|)$ 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
\(\hat{C}_2(k) = h - [1 - (|k|/k_0)^2]^2\), 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,

\[
\hat{C}_2(k) = b_0 + b_1|k|^2 + b_2|k|^4,
\]

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 \(\hat{C}_2(k)\) simultaneously] from a nucleation event with

![Graph showing data and quartic approximation](image)

FIG. 2. Learning density functional theory (DDFT) from images. (a) Five images of the nucleation event used as training data. (b) Inversion of \(\mu_h(\eta)\) and \(\hat{C}_2(k)\). \(\hat{C}_2(k)\) 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 \(\hat{C}_2(k)\) 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 \(\hat{C}_2(k)\) using Eq. (9). The shaded region in the last image shows the estimated error of \(\hat{C}_2(k)\) inferred by adding \(e \sum_d d_i^2\) to the objective function normalize by area, where \(e = 0.01\), using 100 Hermite functions.

The dynamics in \((|k|/k_0)^2\) is the characteristic wave number. The resulting component \(\hat{C}_2(k)\) 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 \(\hat{C}_2(k)\) cannot be identified at large \(|k|\), since high wave number components decay rapidly in time due to the Laplacian operator in the PDE. Therefore, we emphasize that finding \(\hat{C}_2(k)\) over the entire range of \(k\) is an ill-posed problem.

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

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

where \(H_n\) is the Hermite polynomial. For a faster convergence, \(k\) 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 \(\hat{C}_2(k)\) confirms high sensitivity around \(k_0\) and high uncertainty at large \(|k|\).

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.

This work was supported by Toyota Research Institute through the D3BATT Center on Data-Driven-Design of Rechargeable Batteries.

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


