Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Probing Non-Equilibrium Grain Boundary Dynamics with XPCS and Domain-Adaptive Machine Learning

Mingda Li

mingda@mit.edu

Massachusetts Institute of Technology https://orcid.org/0000-0002-7055-6368

Mouyang Cheng

Massachusetts Institute of Technology https://orcid.org/0009-0001-7014-2464

Bowen Yu

Massachusetts Institute of Technology

Chuliang Fu

Massachusetts Institute of Technology https://orcid.org/0000-0003-4626-9622

Nina Andrejevic

Argonne National Laboratory https://orcid.org/0000-0002-8648-5859

Matthias Agne

University of Oregon https://orcid.org/0000-0001-8270-5730

Riley Hanus

Northwestern University https://orcid.org/0000-0002-6242-0647

Qiwei Wan

Massachusetts Institute of Technology

Nathan Drucker

Harvard University

Thanh Nguyen

Massachusetts Institute of Technology

Andrei Fluerasu

Brookhaven National Laboratory https://orcid.org/0000-0003-4690-1675

Lutz Wiegart

Brookhaven National Laboratory

Xiaoqian Chen

Brookhaven National Laboratory

Daniel Pajerowski

Oak Ridge National Laboratory https://orcid.org/0000-0003-3890-2379

Yongqiang Cheng

Oak Ridge National Lab https://orcid.org/0000-0002-3263-4812

Joshua Turner

SLAC https://orcid.org/0000-0002-2106-7955

G. Jeffrey Snyder

Northwestern University https://orcid.org/0000-0003-1414-8682

Article

Keywords:

Posted Date: November 13th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-9870076/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: There is **NO** Competing Interest.

Probing Non-Equilibrium Grain Boundary Dynamics with XPCS and Domain-Adaptive Machine Learning

Mouyang Cheng[1,2,3,†,*], Bowen Yu[1,4,†], Chu-Liang Fu[1,5], Nina Andrejevic[6], Matthias T. Agne[7], Riley Hanus[8], Qiwei Wan[1,5], Nathan C. Drucker[1,9], Thanh Nguyen[1,5], Andrei Fluerasu[10], Lutz Wiegart[10], Xiaoqian M Chen[10], Daniel Pajerowski[11], Yongqiang Cheng[11], Joshua J Turner[12], G. Jeffrey Snyder[8], and Mingda Li[1,2,5,]**

[1]Quantum Measurement Group, MIT, Cambridge, MA 02139, USA
[2]Center for Computational Science and Engineering, MIT, Cambridge, MA 02139, USA
[3]Department of Materials Science and Engineering, MIT, Cambridge, MA 02139, USA
[4]Department of Physics, MIT, Cambridge, MA 02139, USA
[5]Department of Nuclear Science and Engineering, MIT, Cambridge, MA 02139, USA
[6]Center for Nanoscale Materials, Argonne National Laboratory, Lemont, IL 60439, USA
[7]Materials Science Institute, University of Oregon, Eugene, OR 97403, USA
[8]Department of Materials Science and Engineering, Northwestern University, Evanston, IL 60208, USA
[9]Department of Applied Physics, School of Engineering and Applied Sciences, Harvard University, Cambridge, MA 02138, USA
[10]National Synchrotron Light Source II, Brookhaven National Laboratory, Upton, NY 11973, USA
[11]Neutron Scattering Division, Oak Ridge National Laboratory, Oak Ridge, TN 37831, USA
[12]SLAC National Accelerator Laboratory, Menlo Park, CA 94025, USA
[†]These authors contributed equally.
[*]e-mail: vipandyc@mit.edu
[**]e-mail: mingda@mit.edu

ABSTRACT

Grain-boundary (GB) dynamics control the stability, mechanical, and functional response of nanocrystalline materials, but direct experimental access to their slow non-equilibrium motion has been limited. Here we establish X-ray photon correlation spectroscopy (XPCS), combined with domain-adaptive machine learning, as a quantitative probe of GB dynamics. Temperature- and grain-size-dependent two-time XPCS measurements in nanocrystalline silicon reveal pronounced departures from time-translation invariance, showing that GB relaxation can remain far from equilibrium over experimental timescales. However, direct extraction of quantitative physical information from these high-dimensional, noisy fluctuation maps faces a significant challenge. To overcome this barrier, we develop a semi-supervised learning framework that transfers physical parameter labels from continuum simulations to unlabeled experimental XPCS maps through domain-adaptive representation alignment. This AI-augmented approach enables the extraction of key kinetic parameters, including bulk diffusivity, GB stiffness, and effective GB concentration, directly from experimental XPCS measurements. Our results show how machine learning can transform indirect fluctuation signals into quantitative materials dynamics, providing a general route to study non-equilibrium defect motion in solids.

Introduction

Nanograined materials derive many of their remarkable properties from the dense grain-boundary (GB) networks that thread through their microstructure. In these materials, GBs are not static interfaces, but dynamic mesoscale objects whose migration, sliding, relaxation, and structural transformation can reshape materials over time. Stress-driven GB migration shows that GBs can move under shear rather than simply obstructing dislocations[1], while atomic-resolution studies reveal coupled sliding, atomic transfer, disconnection motion, and structural transformation during GB deformation[2,3]. In nanograined alloys, relaxation and stabilization of high-density GB networks can suppress diffusional creep and sustain high-temperature strength[4]; under irradiation, GB facet motion and interfacial defect

dynamics govern microstructural evolution[5]. Thus, nanograined solids are governed not only by GB structure, but by GB dynamics: how boundaries move, fluctuate, relax, and transform under external perturbations.

Despite this importance, experimental access to slow GB dynamics remains limited. Many relevant relaxation modes occur over minutes to hours and manifest as weak, spatially correlated fluctuations rather than abrupt morphological changes. High-energy diffraction microscopy can track grain-scale evolution[6], while four-dimensional electron microscopy[7,8], Bragg coherent diffraction imaging[9,10], and time-resolved X-ray tomography[11–13] provide complementary views of structure, strain, and mesoscale morphology. Yet none readily combines the temporal window, spatial sensitivity, and statistical averaging needed to quantify slow, non-equilibrium GB relaxation. Consequently, key dynamical laws, including curvature-driven migration, stress-assisted motion, thermal activation, impurity pinning, and noise-activated hopping, remain difficult to validate experimentally despite extensive theory[14–19].

In this work, we show that X-ray photon correlation spectroscopy (XPCS), when augmented by specialized machine learning, can directly reveal GB dynamics in nanocrystalline materials. XPCS has long been used to probe time-resolved nanoscale dynamics in soft matter, glasses, colloids, and other complex systems[20–23]. More recently, it has also been extended to probe slow fluctuations of ordered states in hard condensed matter systems[24–26]. However, extracting quantitative physical information from XPCS remains highly challenging because the measured two-time correlation maps are high-dimensional, noisy, and often far from equilibrium. Recent machine-learning approaches have enabled automated classification of relaxation dynamics from experimental XPCS data[27], but converting such measurements into material-specific kinetic parameters remains largely inaccessible.

Here, we overcome this challenge by designing a semi-supervised domain-adaptive learning framework that bridges continuum simulations and experimental XPCS measurements. Our model uses simulated GB dynamics to learn the relationship between two-time correlations and physical parameters, while experimental data constrain the learned representation through domain alignment and non-equilibrium consistency. This enables, for the first time, direct extraction of key GB-motion parameters, including bulk diffusivity D, GB stiffness Γ, and effective GB concentration λ_{GB}, from experimental XPCS maps. It also allows us to directly quantify how far nanograined silicon departs from equilibrium through the breakdown of time-translation invariance. By turning XPCS into a quantitative measurement of GB dynamics, our work establishes a new route for studying nanograined materials across long timescales and weak fluctuation regimes inaccessible to real-space imaging. More broadly, the framework provides a generally applicable strategy for synergizing experimental and computational data, marking a step toward applying AI directly to noisy, high-dimensional experimental observables in complex materials systems.

Results

Workflow

We integrate experiment, theory, and AI to extract physically meaningful parameters from non-equilibrium XPCS measurements of GB motion, as is illustrated in Fig. 1. *Experimentally*, XPCS in a grazing-incidence small angle X-ray scattering (GISAXS) geometry probes near-surface mesoscopic structures by measuring coherent X-ray speckle patterns from the disordered GB network. Correlating these speckle patterns over time reveals the dynamics of the scatterers, yielding the two-time intensity-intensity correlation maps $g_2(\mathbf{q},t_1,t_2)$ across temperatures. *Theoretically*, the two-time correlation g_2 can be computed from atomic-scale dynamics, although directly resolving GB dynamics at that scale could be prohibitively expensive. Therefore, instead of tracking the dynamical trajectory of each silicon atom, we adopt a coarse-grained model, in which GB motion is described by a stochastic differential equation (SDE) for the interfacial height $z(x,t)$, with curvature-driven migration and thermal fluctuations governing its time evolution. Through this coarse-grained model, one can characterize the GB motion by a few controllable parameters, including the bulk diffusivity D, GB stiffness Γ, and the effective GB concentration λ_{GB} that contributes to the non-equilibrium motion. From these trajectories we compute synthetic $g_2(\mathbf{q},t_1,t_2)$ maps and their non-equilibrium measure, $\mathscr{S}_{\mathrm{noneq}}$ (see definition in Eq. 3) over a grid of parameters (D, Γ, λ_{GB}). Results show that, across a reasonable range of model parameters, the theory reproduces key experimental trends, including the g_2 patterns and the crossover from equilibrium to non-equilibrium behavior.

With experimental data and the theoretical model at hand, we aim to assign the theoretical parameters to experimental XPCS data. A natural thought is to train a neural network on simulated g_2 maps, use them as image-like input and optimize a regression loss $\mathscr{L}_y$ to predict (D, Γ, λ_{GB}). However, the domain gap between idealized simulations and experimental measurements, which could arise from background noise, instrumental response, and unmodeled microstructure, prevents direct application of such a model. This is further complicated by the fact that experimental data are scarce and unlabeled, making purely supervised regression infeasible. To address this, we adopt a semi-supervised learning framework, where both labelled simulated data and unlabelled experimental data are fully utilized. The net training objective is to minimize both the predictive loss $\mathscr{L}_y$ from the GB parameter predictor and the domain alignment loss $\mathscr{L}_d$ from the domain adaptor. Both experimental (orange) and simulated (blue) inputs are passed through the same *feature extractor* (with shared weights) to obtain their *feature vectors*. The regression loss $\mathscr{L}_y$ is applied only to simulated data, where ground-truth parameters are available. In parallel, a domain alignment loss $\mathscr{L}_d$ encourages the embeddings of simulated and experimental data to overlap, effectively bridging the domain gap. As a result, the model learns a shared representation that aligns experimental and theoretical data while preserving accurate parameter estimation from simulation. This enables direct and reliable inference of GB kinetic parameters from experimental XPCS measurements, effectively using AI to bridge theory and experiment within a unified framework. More details of the training setup are shown in the Methods section.

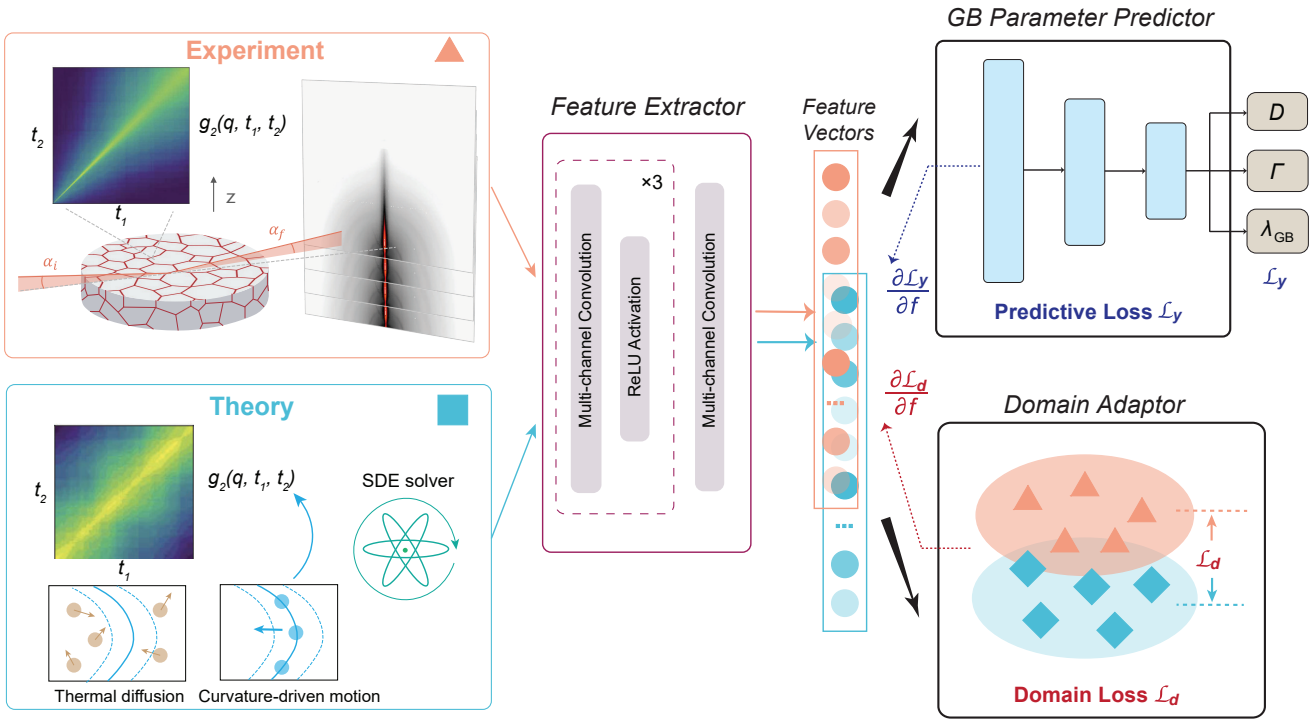


Figure 1. Integrated workflow with experiment, theory, and AI for grain-boundary dynamics. Experimentally, X-ray photon correlation spectroscopy (XPCS) measures the two-time correlation function $g_2(\mathbf{q},t_1,t_2)$, capturing non-equilibrium dynamics of the grain boundary. Theoretically, stochastic differential equation (SDE)-based simulations incorporate thermal diffusion and curvature-driven grain-boundary migration to generate theoretical $g_2(\mathbf{q},t_1,t_2)$ maps. To bridge the domain gap between theory and experiment, we employ a semi-supervised domain adaptation framework with two coupled losses: a regression loss $\mathscr{L}_y$ that trains theoretical g_2 maps to predict $(D,\Gamma,\lambda_{\mathrm{GB}})$, and a domain alignment loss $\mathscr{L}_d$ that constrains experimental (orange triangles) and theoretical (blue diamonds) embeddings with matched non-equilibrium measures into a shared feature space. This alignment enables direct and robust inference of GB kinetic parameters from experimental XPCS measurements.

XPCS measurement

The XPCS is measured for nanostructured bulk silicon, which is synthesized according to the methods described by Ref. [28]. More details about the XPCS measurements are shown in Methods. Several representative XPCS measurements of the two-time correlation function $g_2(t_1,t_2)$, acquired over an observation window $0 \le t_1,t_2 \le \mathscr{T} = 2,500$ s at selected temperatures T on the same sample, are shown in Fig. 2a. Here, $g_2(t_1,t_2)$ denotes $g_2(\mathbf{q},t_1,t_2)$ evaluated at an in-plane momentum transfer $q_r = 0.045$ Å^{-1} and an out-of-plane component $q_z = 0.275$ Å^{-1}, with the full scattering vector $\mathbf{q}$ omitted for brevity. At 299 K, the g_2 map is close to time-translation invariant (TTI): intensity is organized mainly along the $\tau = t_2 - t_1$ diagonals and is consistent with near-equilibrium dynamics, where $g_2(t_1,t_2)$ may be approximated as $g_2(\tau)$. Upon heating, this diagonal organization is progressively lost. At 371 K, the ridge broadens and bends into extended off-diagonal regions for $t_1,t_2 \ge 1,250$ s, demonstrating a shift from near-equilibrium dynamics towards non-equilibrium. At 421 K and 471 K, the maps develop complex, highly non-equilibrium structures, including pronounced block-like and banded wedges, which reveals strongly history-dependent, non-equilibrium structural dynamics. In particular, the box-like structures at 471 K suggest intermittent non-stationary dynamics, in which the system remains correlated within a finite time window and then decorrelates sharply across rearrangement events.

To quantify the non-equilibrium measure for each sample, we utilize the TTI property for thermal-equilibrated dynamics, where g_2 depends only on the lag $\tau = t_2 - t_1$. Deviations from TTI therefore provide a direct, model-agnostic measure of non-equilibrium. The best TTI "baseline" is first obtained by averaging g_2 along lines of constant τ (the diagonals of the two-time plane as shown in Fig. 2a). Denoting $g_2(t_1,t_2) = g_2(t_1,t_1+\tau)$, we define

$$g_2^{\mathrm{TTI}}(\tau) \equiv \frac{1}{N(\tau)} \int_{\max(0,-\tau)}^{\min(\mathscr{T},\mathscr{T}-\tau)} g_2(t,t+\tau)\,dt, \quad N(\tau) = \mathscr{T} - |\tau| \tag{1}$$

which performs the diagonal average along the g_2 map within the appropriate finite-window limits. The non-equilibrium drift $S_{\mathrm{noneq}}(t_1,t_2)$ for each time pair (t_1,t_2) is then defined as

$$S_{\mathrm{noneq}}(t_1,t_2) = g_2(t_1,t_2) - g_2^{\mathrm{TTI}}(t_2-t_1) \tag{2}$$

By construction, $S_{\mathrm{noneq}}(t_1,t_2) \equiv 0$ if the dynamics is fully equilibrated. Building upon this, we can further define the global non-equilibrium measure $\mathscr{S}_{\mathrm{noneq}}$ as

$$\mathscr{S}_{\mathrm{noneq}} = \frac{\left\langle S_{\mathrm{noneq}}(t_1,t_2)^2 \right\rangle_{t_1,t_2}}{\left\langle \left(g_2(t_1,t_2) - \left\langle g_2 \right\rangle_{t_1,t_2} \right)^2 \right\rangle_{t_1,t_2}} \in [0,1] \tag{3}$$

i.e., the fraction of the variance of g_2 that arises from non-equilibrium effects. The brackets $\langle\cdot\rangle_{t_1,t_2}$ denote time averaging over all (t_1,t_2) pairs within the observation window. $\mathscr{S}_{\mathrm{noneq}}$ is naturally normalized between 0 and 1: it is 0 for fully equilibrated dynamics, grows with the strength of non-equilibrium, and is capped at 1 where g_2 contains no time-translation-invariant component τ and all variance arises from explicit (t_1,t_2) dependence. For better visualization of non-equilibrium effects, we plot $f(\tau;t_1) = g_2(t_1,t_1+\tau)$ as functions of τ for different t_1 in Fig. 2b. For the spectrum close to TTI at 299 K, the correlation slices $f(\tau;t_1)$ nearly coincides as a single decay, consistent with a low non-equilibrium measure $\mathscr{S}_{\mathrm{noneq}} = 0.097$. For the highly non-equilibrated spectra at 421 K and 471 K, correlation slices $f(\tau;t_1)$ separate significantly according to t_1, which correspond to their large non-equilibrium measures $\mathscr{S}_{\mathrm{noneq}} = 0.771$ and $\mathscr{S}_{\mathrm{noneq}} = 0.837$. Additional XPCS measurement data for other nanocrystalline silicon samples are shown in Supplementary Information 2.1.

Simulation of XPCS signal from GB motion

To gain physical insight into the non-equilibrium dynamics observed experimentally, we develop a theoretical model where we attribute the leading contribution to GB dynamics. This assumption is justified by the long timescales involved (> 1,000 s, see Fig. 2a), which are naturally associated with processes occurring over large spatial scales, such as GB motion. We start from the Siegert relation[21,29,30], approximating g_2 as

$$g_2(\mathbf{q},t,t') = 1 + \beta |g_1(\mathbf{q},t,t')|^2 \tag{4}$$

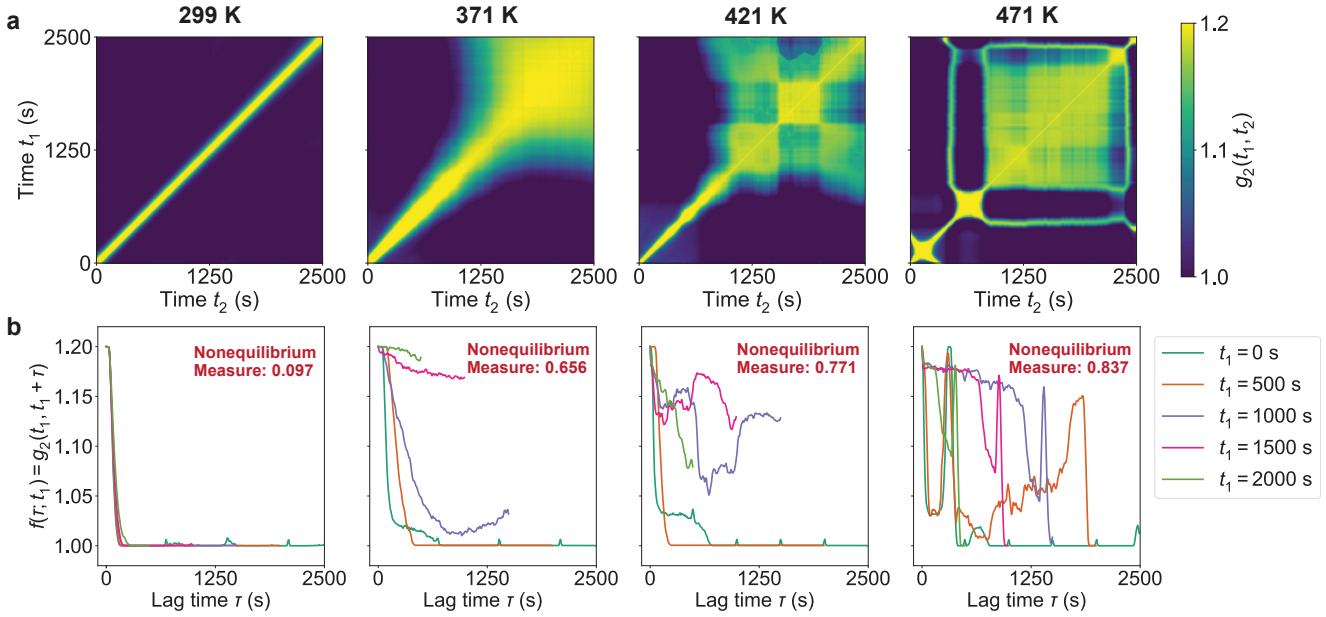


Figure 2. Experimental temperature-dependent XPCS dynamics and non-equilibrium behavior. a. Two-time correlation functions $g_2(t_1,t_2)$ measured at 299 K, 371 K, 421 K, and 471 K for the same nano-crystalline silicon sample. The results reveal increasingly non-equilibrium features in the dynamics as temperature rises. Here, $g_2(t_1,t_2)$ denotes $g_2(\mathbf{q},t_1,t_2)$ evaluated with in-plane momentum transfer $q_r = 0.045\text{Å}^{-1}$ and out-of-plane $q_z = 0.275\text{Å}^{-1}$. **b.** Corresponding intensity-intensity correlation slices $f(\tau;t_1) = g_2(t_1,t_1+\tau)$ at different times t_1. As temperature increases, $f(\tau;t_1)$ demonstrates systematic deviations between different t_1, indicating a transition from near-equilibrium to non-equilibrium GB dynamics. A more quantitative non-equilibrium measure following the definition in Eq. 3 is also shown for each temperature.

where β is an experiment-dependent constant and $g_1(\mathbf{q},t,t') = \frac{1}{N}\sum_{i,j}\langle e^{i\mathbf{q}\cdot[\mathbf{r}_i(t)-\mathbf{r}_j(t')]}\rangle$ is the coherent intermediate scattering function, with N the total number of atoms, $\mathbf{r}_i(t)$ the position of atom i at time t, and $\langle\cdot\rangle$ denoting an ensemble average. The intermediate scattering function $g_1(\mathbf{q},t,t')$ can, in principle, be computed directly from atomistic simulations such as molecular dynamics (MD)[31]. However, MD is limited to timescales on the order of nanoseconds, far shorter than the microsecond-to-second dynamics probed by XPCS, making a direct comparison infeasible. To facilitate direct comparison with experimental data, we find it sufficient to adopt a continuum model with a mixture of non-equilibrium GB motion and equilibrium diffusion. Specifically, we split the atoms in the system into two parts: I. the atoms proximal to the grain boundary (GB); II. the atoms away from the GB. We assume that all atoms not on the GB are at equilibrium and the non-equilibrium XPCS signals are mainly attributed to the motion of GB, where the GB configuration is not fully relaxed, possibly due to low annealing temperature. Neglecting the cross terms, we approximate $g_1(\mathbf{q},t,t')$ as

$$g_1(\mathbf{q},t,t')/g_0 = \lambda_{\text{GB}}\left\langle e^{i\mathbf{q}\cdot(\mathbf{r}_i(t)-\mathbf{r}_j(t'))}\right\rangle_{\text{GB}} + (1-\lambda_{\text{GB}})\left\langle e^{i\mathbf{q}\cdot(\mathbf{r}_i(t)-\mathbf{r}_j(t'))}\right\rangle_{\text{eq}} \tag{5}$$

Here, g_0 is a normalization constant and $\lambda_{\text{GB}} \in [0,1]$ is the effective ratio of atoms proximal to the GB to the total atoms. Larger grain size means lower concentration of GB, thereby smaller λ_{GB}. The equilibrium dynamics are assumed to be governed by fast diffusive motion, and we approximate the corresponding decorrelation by a Brownian form,

$$\left\langle e^{i\mathbf{q}\cdot\left(\mathbf{r}_i(t)-\mathbf{r}_j(t')\right)}\right\rangle_{\text{eq}} = \left\langle e^{i\mathbf{q}\cdot\left(\mathbf{r}_i(t)-\mathbf{r}_j(t)\right)}\right\rangle e^{-\mathbf{q}^2D(t'-t)} = S(\mathbf{q})e^{-\mathbf{q}^2D(t'-t)}, \quad t' > t. \tag{6}$$

where $S(\mathbf{q})$ is the structure factor of the sample, and D is the bulk diffusivity of the atoms in the sample. For a given momentum transfer $\mathbf{q}$, $S(\mathbf{q})$ is time-independent over the XPCS observation window, and it can be absorbed as a

prefactor into λ_{GB} as well.

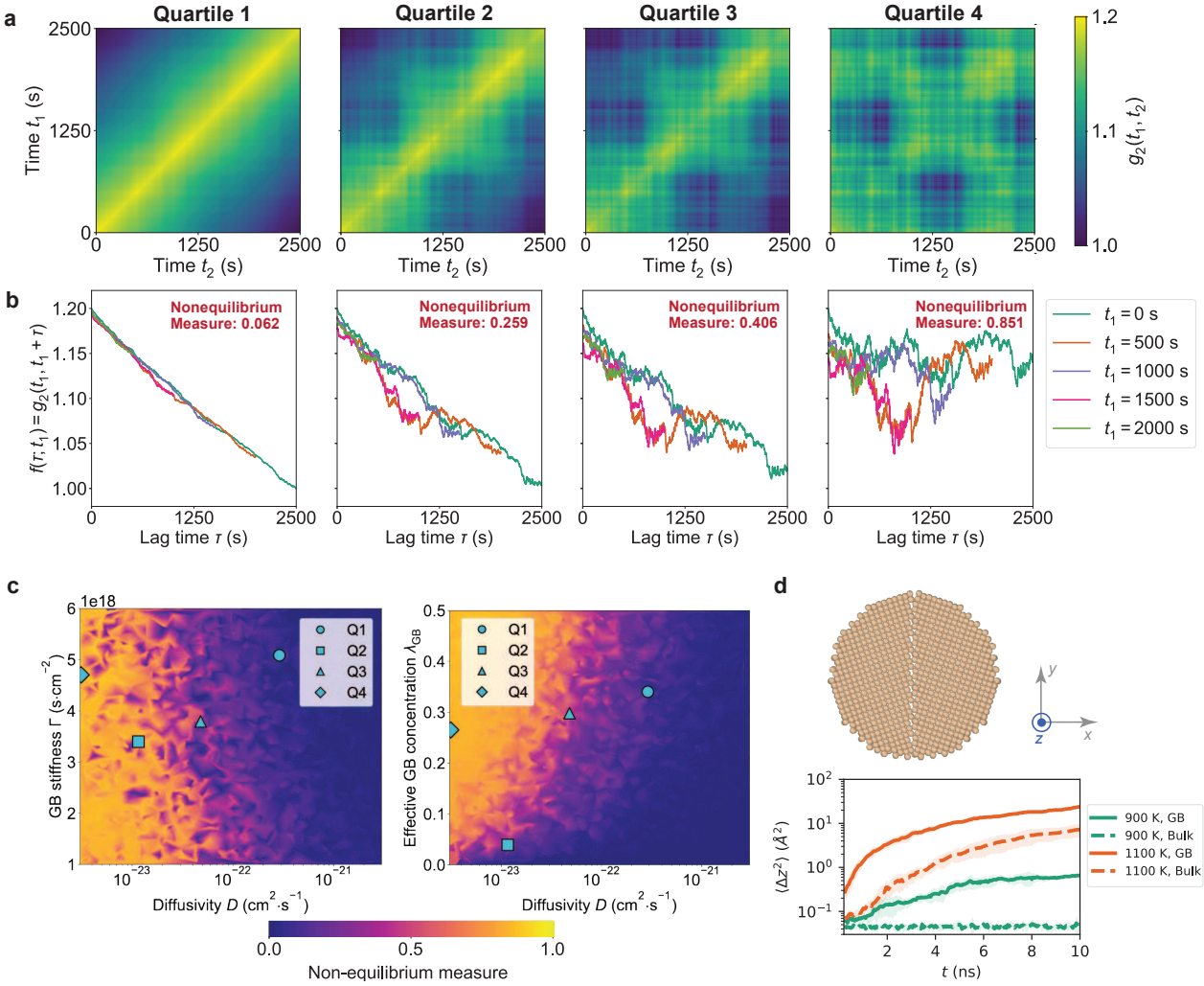


Figure 3. Theoretical SDE predictions and mapping of non-equilibrium dynamics across parameter space. a. Simulated two-time correlation functions $g_2(t_1,t_2)$ for four representative parameter sets, showing the progression from nearly time-translation-invariant dynamics to strongly non-equilibrium behavior. Here $g_2(t_1,t_2)$ denotes $g_2(\mathbf{q},t_1,t_2)$ with in-plane $q_r = 0.045\text{Å}^{-1}$. **b.** Corresponding intensity-intensity correlation slices $f(\tau;t_1) = g_2(t_1,t_1+\tau)$ at different starting times t_1, with non-equilibrium measures $\mathscr{S}_{\mathrm{noneq}} = 0.062$, 0.259, 0.406, and 0.851 for the four cases in **a**. **c.** Heatmap visualization of the predicted non-equilibrium measure across pairwise parameter spaces involving bulk diffusivity D, GB stiffness Γ, and effective GB concentration λ_{GB}; markers indicate the four representative cases shown in **a** and **b**. **d.** Classical MD of a finite cylindrical $\Sigma5(210)[001]$ Si bicrystal, with free lateral surfaces and a periodic z direction parallel to the GB plane. Axial mean-square displacement, $\langle \Delta z^2 \rangle$, is shown versus time at 900 and 1100 K for atoms initially proximal to the GB (solid) and other interior atoms, classified as bulk (dashed). Curves are means over three independent seeds, with shading indicating the corresponding standard deviation.

To compute the contribution from GB motion, we assume that the grain boundary (GB) lies in the x-y plane, with out-of-plane deviations described by a displacement field $z = z(x)$, i.e. the morphology varies only along the x-direction and is uniform along y. Thus, the classical equation of motion (EOM) for the GB in the continuum

limit[32] is reduced to a stochastic differential equation (SDE):

$$v(x,t) = \frac{\partial z(x,t)}{\partial t} = M\left[\Gamma \frac{\partial^2 z(x,t)}{\partial x^2} + \xi(x,t)\right] \tag{7}$$

where $v(x,t)$ is the z-directional velocity, M is the mobility constant of the GB, Γ is the GB stiffness, and $\xi(x,t)$ denotes the thermal noise. With the EOM and proper setup of initial and boundary conditions, one can numerically simulate the time-evolution of GB from the SDE and finally obtain the XPCS two-time correlation $g_2(t_1,t_2)$, allowing us to study its dependence on physical parameters such as D, Γ and λ_{GB}. More details regarding the model and SDE setup for GB dynamics, as well as the computable observables for XPCS mapping are shown in Methods and Supplementary Information 1.

Results for the continuum model simulation are shown in Fig. 3, where $g_2(t_1,t_2)$ denotes $g_2(\mathbf{q},t_1,t_2)$ evaluated at an in-plane momentum transfer $q_r = 0.045$ Å^{-1}, similar to Fig. 2. To test whether this minimal model can generate the hierarchy of two-time structures seen experimentally, we simulated 2,000 synthetic XPCS maps over a broad kinetic parameter space (Methods). The resulting ensemble spans the entire non-equilibrium measure range from $\mathscr{S}_{\mathrm{noneq}} = 1.8\times10^{-4}$ to $\mathscr{S}_{\mathrm{noneq}} = 0.889$, which also fully covers the experimental range in Fig. 2. The four maps in Fig. 3a,b were selected from successive quartiles of this distribution and show a continuous evolution from nearly stationary dynamics to the block-like and wedge-shaped correlations characteristic of strongly non-equilibrium relaxation, highly similar to the experimental observation in Fig. 2a.

Furthermore, Fig. 3c presents a systematic mapping of the non-equilibrium measure across the parameter space $(D,\Gamma,\lambda_{\mathrm{GB}})$, revealing the interplay between bulk diffusive relaxation, GB curvature-driven dynamics, and effective GB concentration. Physically, the bulk diffusivity D governs the rate of equilibrium decorrelation, effectively driving the system towards time-translation invariance. In contrast, the grain-boundary stiffness Γ controls the strength of curvature-driven restoring forces, thereby controlling the intrinsic relaxation timescale of GB fluctuations. The effective GB concentration λ_{GB} determines the relative weight of non-equilibrium GB motion in the total scattering signal. As a result, non-equilibrium behavior emerges from a competition between fast bulk diffusion and slow, curvature-driven GB dynamics, with larger λ_{GB} amplifying the contribution of the latter. From Fig. 3c, we find that decreasing D or increasing λ_{GB} enhances the persistence of non-stationary features in $g_2(t_1,t_2)$ (right panel), while smaller Γ further promotes non-equilibrium behavior by weakening curvature-driven restoring forces and slowing GB relaxation (left panel). The parameter map also reveals crossover regimes in both (Γ,D) and $(\lambda_{\mathrm{GB}},D)$ spaces, separating regimes dominated by either diffusive equilibration or GB-driven dynamics. Additionally, we highlight in Fig. 3c the region of parameter space (Q1-Q4) used in Fig. 3a, illustrating how the model can be tuned from equilibrium to non-equilibrium regimes and how parameters can be sampled accordingly for batch simulations. More details regarding the parameter map of XPCS GB dynamics from the theoretical model are shown in Supplementary Information 2.2.

Complementary to the continuum modelling, we perform classical MD simulations of a finite $\Sigma5(210)[001]$ Si bicrystal to provide an atomistic consistency check for the separation between bulk diffusion and GB-associated dynamics (Fig. 3d). The periodic axial direction $\hat{z}$ lies parallel to the GB plane. At 900 K, atoms initially proximal to the GB develop substantially larger axial displacements than the remaining atoms classified as bulk, whose mean-square displacement (MSD) $\langle \Delta z^2 \rangle$ remains nearly plateaued over the 10 ns trajectory. At 1100 K, both populations become mobile, while the GB-proximal population retains a larger axial MSD, corresponding to a diffusivity ($1.23\times10^{-7}\mathrm{cm}^2\cdot\mathrm{s}^{-1}$) approximately 2.7× that of the bulk population ($4.56\times10^{-6}\mathrm{cm}^2\cdot\mathrm{s}^{-1}$). These atomistic calculations therefore support the existence of distinct GB-associated and bulk mobility channels underlying the two-component representation in Eq. 5. Because the MD simulations probe elevated-temperature, nanosecond dynamics in a finite bicrystal with free lateral surfaces, we use them only as a qualitative microscopic consistency check rather than to calibrate the effective diffusivity or experimental XPCS timescale. Further details of the MD setup are provided in Methods.

Bridging theoretical and experimental XPCS with semi-supervised learning

With both continuum simulation and experimental XPCS data available, our goal is to place each experimental observation onto the theoretical parameter manifold and infer the underlying physical parameters $(D, \Gamma, \lambda_{\mathrm{GB}})$ of each sample at a given temperature. A natural starting point is a purely supervised regression model trained on simulated data. Specifically, we train a convolutional neural network (CNN) on the 2,000 simulated XPCS spectra to map $g_2(t_1, t_2)$ images to $(D, \Gamma, \lambda_{\mathrm{GB}})$ using a mean-squared-error (MSE) regression loss $\mathscr{L}_y$. We note that in Fig. 3c, the crossover regime appears smoother and more well-defined in the $(D, \lambda_{\mathrm{GB}})$ space compared to other two-parameter projections. Therefore, in this section, we focus on illustrating the $(D, \lambda_{\mathrm{GB}})$ mapping. As shown in Fig. 4a (left and middle panels), this model achieves excellent performance on the simulated data, with a test-set $R^2 = 0.995$ and $R^2 = 0.893$ for the predicted parameters D and λ_{GB} relative to the ground-truth values, indicating that the mapping from XPCS features to physical parameters is well learned within the theoretical domain. However, this model completely fails when applied directly to experimental data. As is shown in Fig. 4a (right panel), when experimental maps are passed through the simulation-trained predictor, the inferred parameters unanimously collapse to outlier regions (e.g., $\lambda_{\mathrm{GB}} \approx 0$). Such observation from the vanilla model's output is inconsistent with the smooth crossover structure in (D, Γ) identified from theory (Fig. 3). This mismatch directly reflects the simulation-to-experiment domain gap. Further tests on data augmentation indicate that the domain gap in this particular scenario is non-trivial, i.e., it can't be fixed by simple operations like smearing or rescaling on the experimental XPCS data (see Supplementary Information 3.1).

To address this, we introduce a domain adaptation objective $\mathscr{L}_d$ in addition to the regression loss $\mathscr{L}_y$. The domain loss $\mathscr{L}_d$ imposes two constraints: (i) alignment of feature distributions between simulated and experimental data in the learned representation space, implemented via a CORAL loss[33]; and (ii) consistency between predicted parameters and experimentally measured non-equilibrium signatures, enforced during the fine-tuning stage for each sample. More details regarding the formulation of training loss and schedule are described in Methods section. The effect of the domain adaptation is shown in Fig. 4b. We visualize the learned feature space for experimental and simulated data, using the uniform manifold approximation and projection (UMAP)[34] for dimension reduction on the feature vectors extracted from CNN. For the vanilla model trained solely on the continuum model's simulation, the simulated and experimental data occupy distinctively separated regions. However, after training with domain loss $\mathscr{L}_d$ on 2,000 simulated and 63 experimental XPCS spectra, the two domains significantly overlap in the two-dimensional UMAP embedding, indicating successful alignment of representations. Importantly, this alignment does not come at the cost of predictive accuracy on the kinetic parameters in theoretical space. As shown in Fig. 4c (left panel), after training the model with domain loss, the predictor retains high test-set $R^2 = 0.950$ and $R^2 = 0.890$ on simulated data with well-behaved adaptation, demonstrating that the physical mapping learned from theory is preserved. With this adapted ML model, we then infer parameters for experimental XPCS maps (which remain unlabeled), thus completing a semi-supervised learning pipeline. The resulting assignments (Fig. 4c, right panel) now lie along the physically expected crossover regime in the $(D, \lambda_{\mathrm{GB}})$ diagram, with trends that are consistent with the continuum model and directly interpretable in terms of competing diffusion and GB dynamics. In this way, although experimental data are scarce and not directly labeled, they instead act as constraints that select physically meaningful regions of the theoretical manifold, enabling robust and interpretable parameter inference for the non-equilibrium dynamics of GB.

Discussion

Our results establish XPCS as a powerful and direct probe of slow, non-equilibrium GB dynamics over long timescales (minutes to hours), in contrast to prior approaches based on discrete real-space snapshots[35]. By resolving the full two-time correlation function $g_2(t_1, t_2)$, this work moves beyond static or time-averaged descriptions and directly quantifies the breakdown of time-translation invariance through a non-equilibrium metric $\mathscr{S}_{\mathrm{noneq}}$. A key observation is that a minimal continuum model, incorporating the competition between bulk diffusion and curvature-driven GB dynamics, is sufficient to reproduce the essential features of the experimental XPCS maps. Despite its simplicity, this model captures the crossover between equilibrium and non-equilibrium regimes and provides

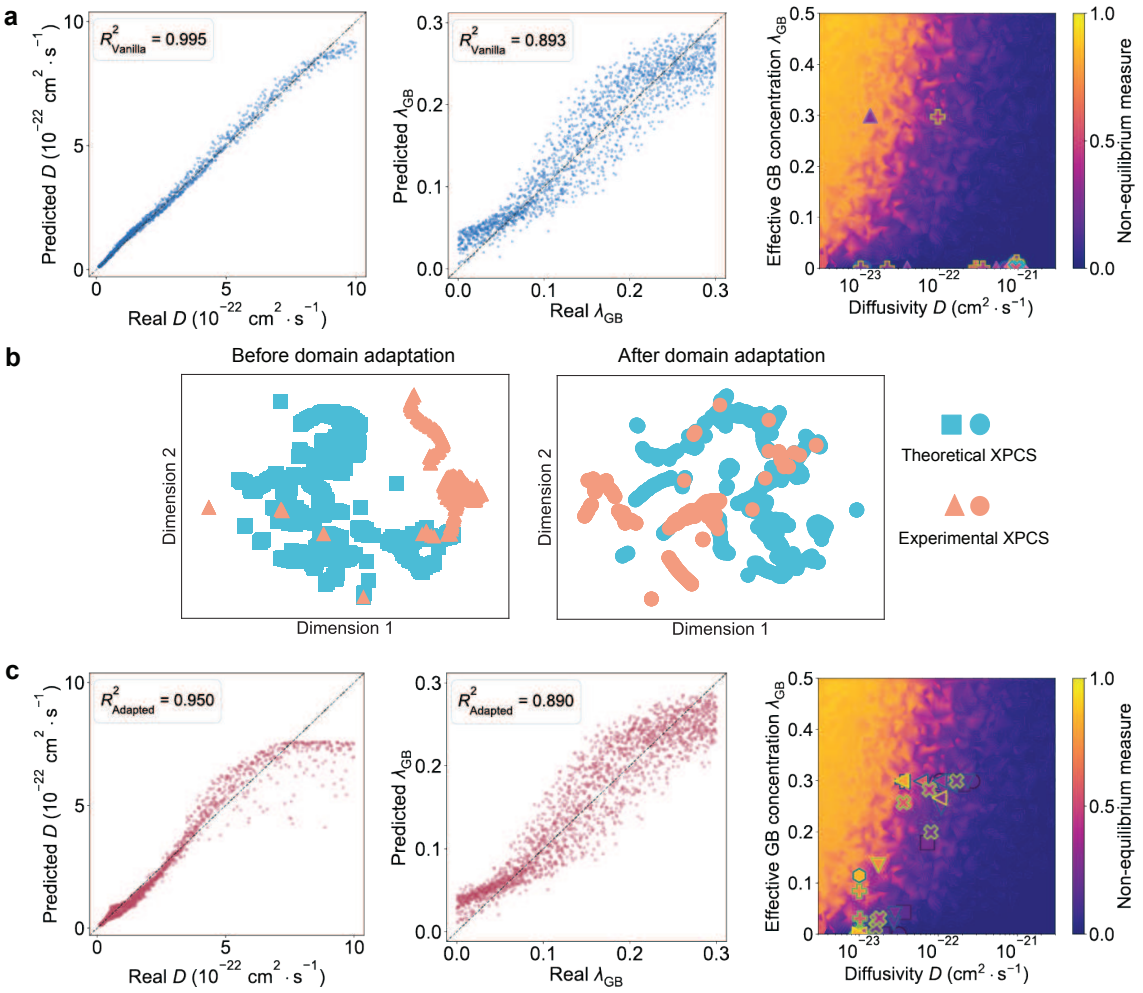


Figure 4. Semi-supervised domain adaptation for mapping experimental XPCS onto the theoretical parameter manifold. a. Prediction results before domain adaptation using the vanilla model trained only on simulated XPCS. Left and middle panel: predicted versus real D and λ_{GB} for simulated test data, with a test-set $R^2 = 0.995$ and 0.893, respectively. Right panel: inferred experimental XPCS locations overlaid on the theoretical (D, λ_{GB}) parameter map, where the background colormap shows the simulated non-equilibrium measure. **b.** Latent feature representations visualized with UMAP before domain adaptation (left) and after contrastive domain adaptation (right). Cyan markers indicate theoretical XPCS samples, and orange markers indicate experimental XPCS samples. **c.** Prediction results after domain adaptation. Left and middle panel: predicted versus real D and λ_{GB} for simulated test data after adaptation, with a test-set $R^2 = 0.950$ and 0.890, respectively. Right panel: inferred experimental XPCS locations on the same theoretical (D, λ_{GB}) map, showing their shift into the physically expected crossover region.

a low-dimensional, physically interpretable parameterization in terms of the bulk diffusivity D, the GB stiffness Γ and the effective GB concentration λ_{GB}. Here, D sets the bulk relaxation rate, while Γ and λ_{GB} describe the effective GB relaxation strength and GB-associated dynamic weight, respectively, with λ_{GB} expected to increase with GB density and decrease with grain size. This indicates that the dominant physics of GB relaxation shown in the experimental data can be effectively recovered without resorting to time-consuming atomic-scale simulation. The primary challenge, however, lies in *quantitatively* connecting the continuum theoretical model to experimental observations. As demonstrated by the failure of the vanilla ML model in Fig. 4a, models trained solely on simulated data do not generalize to experiment due to the substantial domain gap arising from noise, instrumental effects, and unmodeled microstructural complexity. Our semi-supervised domain adaptation framework addresses this challenge by utilizing both labeled simulated data and unlabeled experimental data, where experimental measurements do not act as supervision in the conventional sense, but instead restrict the model to physically plausible regions of the theoretical parameter manifold. This enables robust and interpretable parameter inference while preserving the predictive accuracy learned from XPCS simulation.

Recent advances in AI for physics and materials engineering have increasingly emphasized the need for a closed loop between AI, computation, and experiment[36–42]. In this context, the domain gap between theory and experiment has become a central challenge across all subfields. Several works have attempted to bridge this gap using deep learning. For instance, domain-adversarial learning has been used to extract superconducting phase information from angle-resolved photoemission spectroscopy (ARPES) spectra by aligning simulated and experimental feature representations[43], while adversarial distribution alignment frameworks train generative models on simulations and then match them to distributions of experimental observables, such as nuclear magnetic resonance (NMR) spectroscopy, via a min-max objective with Kullback-Leibler (KL) regularization[44]. However, these approaches rely on min-max optimization, which are prone to instability, sensitive to objective design, and challenging to train in regimes with limited experimental data. In this work, we find that a simple CORAL loss based feature alignment, combined with a physically motivated non-equilibrium consistency constraint, provides a stable and effective solution to the XPCS problem. By avoiding adversarial min-max optimization and relying on feature-space convex alignment losses, the framework achieves robust training, even in data-scarce experimental regimes. Our results propose an alternative route to bridge the domain gap between theory and experiment using deep learning. More details on the comparison between our approach and other methods, including min-max adversarial learning and contrastive learning[45], are shown in Supplementary Information 3.2.

Despite these advances, several limitations of the present framework should be noted. First, the continuum model provides a coarse-grained description of GB and does not explicitly capture atomistic mechanisms such as interactions between multiple grains and dynamic heterogeneity[46–48]. Second, the domain adaptation relies on a scalar non-equilibrium metric, which may not fully capture the richness of the underlying dynamics encoded in the two-time correlation function. Third, the approach assumes that experimental observations lie within the span of the simulated parameter space, which may not hold in more complex materials systems or under extreme conditions (see minor outlier examples in Supplementary Information 3.4). Addressing these limitations will require both improved physical modeling and more expressive data representations. For example, on the physics side, incorporating additional degrees of freedom into the continuum model, such as structural disorder, anisotropy, or multi-boundary interactions, could further improve the fidelity of the simulation; on the experimental side, combining XPCS with complementary modalities may help break degeneracies in interpreting non-equilibrium dynamics.

Looking forward, our results demonstrate that combining XPCS, minimal physical modeling, and domain-adaptive ML enables quantitative and interpretable access to non-equilibrium defect dynamics. More broadly, the integration of experiment, theory, and machine learning suggests a pathway toward a closed-loop framework, in which experimental data iteratively refine theoretical models, and theory guides the interpretation and acquisition of new measurements. Our approach charts a concrete roadmap to bridging the gap between theory and experiment in complex physical systems by data-driven domain adaptation, even when they are separated by indirect measurement processes and modeling approximations.

Methods

Continuum dynamics simulation

Since we assume that all atoms not affected by the GB are at equilibrium, their dynamics can be described by thermal Browninan motion, i.e.

$$\left\langle [\mathbf{r}_i(0) - \mathbf{r}_i(t)]^2 \right\rangle_{\text{eq}} = 6Dt \tag{8}$$

which results in the equilibrium contribution in Eq. 6. As for the motion of atoms on GB[32], the velocity can be expressed as $\mathbf{v}(\mathbf{r},t) = M\mathbf{F}(\mathbf{r},t)$ in the continuum limit. Furthermore, the force $\mathbf{F}(\mathbf{r},t)$ consists of both a curvature restoring force $\Gamma\sum_i h_{ii}$ under the small slope approximation, where h_{ii} is the local i-th principal curvature of GB and only $h_{xx} \approx \frac{\partial^2 z}{\partial x^2}$ is nonzero, and thermal noise $\xi(\mathbf{r},t)$ satisfying $\langle \xi(\mathbf{r},t)\xi(\mathbf{r}',t')\rangle = \frac{2k_BT}{M}\delta(\mathbf{r}-\mathbf{r}')\delta(t-t')$. Putting these together we arrive at Eq. 7, which can be readily discretized and solved numerically given appropriate initial and boundary conditions. For the numerical simulation, we set

$$z(x=0) = z(x=L) = 0; \quad z(t=0) = A\frac{x(L-x)}{L^2}. \tag{9}$$

where L is the system size and $A \neq 0$ marks that the GB is initially not under equilibrium shape, i.e. it will start oscillating. With these conditions and EOM, we can simulate the time-dependent evolution of GB, compute the GB contribution to the intermediate scattering function $\langle e^{i\mathbf{q}\cdot(\mathbf{r}_i(t)-\mathbf{r}_j(t'))}\rangle_{\text{GB}}$, and finally produce the simulated XPCS signal in Eq. 4 combined with Eq. 6. For the simulation ensemble used in Fig. 3 and for supervised model training, we generated 2,000 parameter sets with Γ sampled linearly from 1×10^{18} s$\cdot$cm^{-2} to 6×10^{18} s$\cdot$cm^{-2}, D sampled logarithmically from 3×10^{-24} cm$^2\cdot$s^{-1} to 3×10^{-21} cm$^2\cdot$s^{-1}, λ_{GB} sampled linearly from 0 to 0.5, and T sampled linearly from 300 to 500 K.

Molecular dynamics simulation

MD simulations are performed using LAMMPS[49] with the environment-dependent interatomic potential (EDIP) for Si[50]. We construct a $\Sigma5(210)[001]$ symmetric-tilt bicrystal with a 36.87° misorientation. The bicrystal is modeled as a cylinder with radius 30 Å and axial length 50 Å, surrounded by a 15 Å lateral vacuum region. Free boundary conditions are applied in the lateral (x,y) directions, while the axial z direction is periodic. The GB lies in the yz plane, such that z provides an unbounded direction parallel to the interface. Atoms that are unphysically close upon joining the two lattices are removed using a 2.0 Å overlap criterion. Three trajectories with independent initial velocity seeds are analyzed at each temperature. Each structure is energy-minimized at zero axial pressure, equilibrated for 10 ps in the NVT ensemble, and subsequently simulated for 10 ns at 900 and 1100 K under axial NPT conditions with $P_{zz} = 0$. A timestep of 1 fs is used throughout the MD simulation. Atoms are classified using the relaxed initial configuration. The GB-proximal group consists of atoms satisfying $|x| \leq 7.5$ Å and $r < 26$ Å, where r is the radial distance from the cylinder axis; the latter criterion excludes a 4 Å lateral-surface shell. The bulk reference comprises all remaining atoms with $r < 26$ Å. Axial mean-squared displacements are calculated after removing the center-of-mass motion of each group.

Semi-supervised learning with domain adaptation

We train the parameter inference model with three sub-components: vanilla supervised training on simulations, global CORAL-surrogate adaptation to experimental data, and sample-wise fine-tuning. The network contains two parts, as shown in Fig. 1: a convolutional feature extractor F_θ with parameters θ, mapping the XPCS spectra $g_2(t_1,t_2)$ $(0 \leq t_1,t_2 \leq 2{,}500$ s) to a feature vector $h \in \mathbb{R}^{128}$; a multilayer-perceptron (MLP) predictor P_ϕ with parameters ϕ, mapping the feature vector to predicted parameters. Given an input two-time correlation map x, the model predicts physical parameters

$$\hat{y} = P_\phi(F_\theta(x)), \quad y = (D,\Gamma,\lambda_{\text{GB}}). \tag{10}$$

The vanilla model is trained only on simulated data, for which the parameter labels are known. Its supervised regression loss is

$$\mathscr{L}_y = \frac{1}{N_s}\sum_{i=1}^{N_s}\left|P_\phi(F_\theta(x_i^s)) - y_i^s\right|^2, \tag{11}$$

where $N_s = 2{,}000$ is the total number of simulated data, x_i^s, y_i^s denote the simulated g_2 maps and their corresponding parameters. In practice, all parameters are normalized to $[0,1]$ either linearly (for Γ and λ_{GB}) or logarithmically (for D).

To adapt the simulation-trained model to the experiment domain, we use the unlabeled experimental maps through feature alignment and non-equilibrium consistency. The total domain loss $\mathscr{L}_d$ is comprised of two parts: the first part, feature alignment loss, is quantified using the CORAL loss[33]

$$\mathscr{L}_{\mathrm{CORAL}} = \|\mu_s - \mu_e\|^2 + \frac{1}{4d^2}\|C_s - C_e\|_F^2, \tag{12}$$

where $\mu_{s,e} = \mathbb{E}_{i\in\mathscr{D}_{\mathrm{batch}}}[F_\theta(x_i^{s,e})]$ and $C_{s,e} = \mathbb{E}_{i\in\mathscr{D}_{\mathrm{batch}}}[F_\theta(x_i^{s,e})F_\theta(x_i^{s,e})^\top]$ are the minibatch feature means and covariance matrices for the simulated and experimental domains, $d = 128$ is the feature dimension, and $\|\cdot\|_F$ is the Frobenius norm. This term aligns the first- and second-order feature statistics directly in the learned representation. The second part, non-equilibrium measure constraint, further uses the measured non-equilibrium measure $\mathscr{S}_{\mathrm{noneq}}$ as a weak experimental constraint. A separate surrogate network Q_ω is first trained on simulated data to approximate the forward map from normalized parameters to the non-equilibrium measure,

$$Q_\omega(y_i^s) \approx \mathscr{S}_{\mathrm{noneq}}(x_i^s). \tag{13}$$

After this pretraining, Q_ω is frozen. During domain adaptation, the parameters predicted for experimental maps are required to reproduce the measured experimental non-equilibrium measures:

$$\mathscr{L}_{\mathrm{sur}} = \frac{1}{N_e}\sum_{j=1}^{N_e}\left|Q_\omega(P_\phi(F_\theta(x_j^e))) - \mathscr{S}_{\mathrm{noneq}}(x_j^e)\right|^2. \tag{14}$$

Overall, the domain loss is given by $\mathscr{L}_d = w_{\mathrm{CORAL}}\mathscr{L}_{\mathrm{CORAL}} + w_{\mathrm{sur}}\mathscr{L}_{\mathrm{sur}}$. where we set the hyperparameters $w_{\mathrm{CORAL}} = 1.0$ and $w_{\mathrm{sur}} = 1.2$ in our training (more details of hyperparameters are shown in Supplementary Information 3.3). The full objective for semi-supervised learning with domain adaptation is

$$\mathscr{L} = \mathscr{L}_y + \mathscr{L}_d. \tag{15}$$

The model is optimized using Adam with learning rate $\eta = 3\times10^{-4}$, while the non-equilibrium surrogate is pretrained using Adam with learning rate 10^{-3}.

Finally, we fine-tune the globally adapted model separately for each material-dose group. For this step, the adapted model is copied, the surrogate Q_ω is kept frozen, and only the selected trainable part of the predictor is updated using $\mathscr{L}_{\mathrm{sur}}$ on the experimental shots from that group. This material-wise fine-tuning accounts for residual sample-specific offsets while keeping the inferred parameters tied to the simulation-derived non-equilibrium manifold.

XPCS measurements

X-ray photon correlation spectroscopy (XPCS) measurements were performed at the Coherent Hard X-ray (CHX, 11-ID) beamline at NSLS-II, Brookhaven National Laboratory. A partially coherent beam with energy 9.65 keV ($\lambda = 0.1285$ nm) was selected using a Si(111) double-crystal monochromator. Scattering patterns were collected using a Dectris Eiger X 4M detector positioned 10.08 m downstream. Measurements were carried out in a grazing-incidence reflection geometry with an incidence angle of 0.25°.

Code availability

The code used for this work is publicly available at https://github.com/vipandyc/XPCS_DomainML.

Acknowledgments

The authors thank DJ Srolovitz for the helpful discussions. MC acknowledges support from U.S. Department of Energy (DOE), Office of Science (SC), Basic Energy Sciences Award No. DE-SC0020148. BY thanks support from National Science Foundation (NSF) ITE-2345084. MC, DP, YC, JJT, and ML thank the support from DOE Genesis Mission Multimodal AI for 2D Quantum Magnets (MAIQMag) project. This research used resources at the CHX beamline of the National Synchrotron Light Source II, a U.S. Department of Energy (DOE) Office of Science User Facility operated for the DOE Office of Science by Brookhaven National Laboratory under Contract No. DE-SC0012704. ML acknowledges the support from the Future Energy Systems Center (FUEC) through the MIT Energy Initiative and support from R. Wachnik.

References

1. Rupert, T., Gianola, D., Gan, Y. & Hemker, K. Experimental observations of stress-driven grain boundary migration. *Science* **326**, 1686–1690 (2009).
2. Wang, L. *et al.* Tracking the sliding of grain boundaries at the atomic scale. *Science* **375**, 1261–1265 (2022).
3. Fang, Z. *et al.* Atomic-scale observation of dynamic grain boundary structural transformation during shear-mediated migration. *Sci. Adv.* **8**, eabn3785 (2022).
4. Zhang, B., Tang, Y., Mei, Q., Li, X. & Lu, K. Inhibiting creep in nanograined alloys with stable grain boundary networks. *Science* **378**, 659–663 (2022).
5. Barr, C. M. *et al.* Irradiation-induced grain boundary facet motion: In situ observations and atomic-scale mechanisms. *Sci. Adv.* **8**, eabn0900 (2022).
6. Bernier, J. V., Suter, R. M., Rollett, A. D. & Almer, J. D. High-energy x-ray diffraction microscopy in materials science. *Annu. Rev. Mater. Res.* **50**, 395–436 (2020).
7. Zewail, A. H. Four-dimensional electron microscopy. *science* **328**, 187–193 (2010).
8. Tian, Y. *et al.* Grain rotation mechanisms in nanocrystalline materials: Multiscale observations in pt thin films. *Science* **386**, 49–54 (2024).
9. Robinson, I., Clark, J. & Harder, R. Materials science in the time domain using bragg coherent diffraction imaging. *J. Opt.* **18**, 054007 (2016).
10. Yau, A., Cha, W., Kanan, M. W., Stephenson, G. B. & Ulvestad, A. Bragg coherent diffractive imaging of single-grain defect dynamics in polycrystalline films. *Science* **356**, 739–742 (2017).
11. García-Moreno, F. *et al.* Tomoscopy: Time-resolved tomography for dynamic processes in materials. *Adv. Mater.* **33**, 2104659 (2021).
12. McKenna, I., Poulsen, S., Lauridsen, E. M., Ludwig, W. & Voorhees, P. W. Grain growth in four dimensions: A comparison between simulation and experiment. *Acta materialia* **78**, 125–134 (2014).
13. Zhang, J. *et al.* Grain boundary mobilities in polycrystals. *Acta Materialia* **191**, 211–220 (2020).
14. Gottstein, G., Ma, Y. & Shvindlerman, L. Triple junction motion and grain microstructure evolution. *Acta Materialia* **53**, 1535–1544 (2005).
15. Upmanyu, M., Srolovitz, D., Shvindlerman, L. & Gottstein, G. Misorientation dependence of intrinsic grain boundary mobility: simulation and experiment. *Acta materialia* **47**, 3901–3914 (1999).
16. Taylor, J. E., Cahn, J. W. & Handwerker, C. A. Overview no. 98 i—geometric models of crystal growth. *Acta Metall. et Materialia* **40**, 1443–1474 (1992).

17. Mishin, Y. M. Atomistic theory and computer simulation of grain boundary diffusion. In *Defect and Diffusion Forum*, vol. 143, 1357–1372 (Trans Tech Publ, 1997).

18. Kobayashi, R., Warren, J. A. & Carter, W. C. A continuum model of grain boundaries. *Phys. D: Nonlinear Phenom.* **140**, 141–150 (2000).

19. Trautt, Z. & Mishin, Y. Grain boundary migration and grain rotation studied by molecular dynamics. *Acta Materialia* **60**, 2407–2424 (2012).

20. Leheny, R. L. Xpcs: Nanoscale motion and rheology. *Curr. opinion colloid & interface science* **17**, 3–12 (2012).

21. Bikondoa, O. X-ray photon correlation spectroscopy for the characterization of soft and hard condensed matter. *X-ray Neutron Tech. for Nanomater. Charact.* 95–156 (2016).

22. Lu, X., Mochrie, S., Narayanan, S., Sandy, A. & Sprung, M. Temperature-dependent structural arrest of silica colloids in a water–lutidine binary mixture. *Soft Matter* **6**, 6160–6177 (2010).

23. Zinn, T., Sharpnack, L. & Narayanan, T. Phoretic dynamics of colloids in a phase separating critical liquid mixture. *Phys. Rev. Res.* **2**, 033177 (2020).

24. Porter, Z. *et al.* Understanding the superconductivity and charge density wave interaction through quasi-static lattice fluctuations. *Proc. Natl. Acad. Sci.* **121**, e2412182121 (2024).

25. Shen, L. *et al.* Interplay between relaxational atomic fluctuations and charge density waves in la 2- x sr x cuo 4. *Phys. Rev. B* **108**, L201111 (2023).

26. Tumbleson, Z. *et al.* Thermodynamic phase transitions of nematic order in magnetic helices. *Sci. Adv.* **11**, eadt5680 (2025).

27. Horwath, J. P. *et al.* Ai-nerd: Elucidation of relaxation dynamics beyond equilibrium through ai-informed x-ray photon correlation spectroscopy. *Nat. Commun.* **15**, 5945 (2024).

28. Bux, S. K. *et al.* Nanostructured bulk silicon as an effective thermoelectric material. *Adv. Funct. Mater.* **19**, 2445–2452 (2009).

29. Siegert, A. *On the fluctuations in signals returned by many independently moving scatterers* (Radiation Laboratory, Massachusetts Institute of Technology, 1943).

30. Berne, B. J. & Pecora, R. *Dynamic light scattering: with applications to chemistry, biology, and physics* (Courier Corporation, 2000).

31. Allen, M. P. & Tildesley, D. J. *Computer simulation of liquids* (Oxford university press, 2017).

32. Chen, D. & Kulkarni, Y. Atomistic modeling of grain boundary motion as a random walk. *Phys. Rev. Mater.* **2**, 093605 (2018).

33. Sun, B. & Saenko, K. Deep coral: Correlation alignment for deep domain adaptation. In *European conference on computer vision*, 443–450 (Springer, 2016).

34. McInnes, L., Healy, J. & Melville, J. Umap: Uniform manifold approximation and projection for dimension reduction. *arXiv preprint arXiv:1802.03426* (2018).

35. Qiu, C. *et al.* Grain boundaries are brownian ratchets. *Science* **385**, 980–985 (2024).

36. Carrasquilla, J. Machine learning for quantum matter. *Adv. Physics: X* **5**, 1797528 (2020).

37. Chen, Z. *et al.* Machine learning on neutron and x-ray scattering and spectroscopies. *Chem. Phys. Rev.* **2** (2021).

38. Cheng, M., Okabe, R., Chotrattanapituk, A. & Li, M. Machine learning detection of majorana zero modes from zero-bias peak measurements. *Matter* **7**, 2507–2520 (2024).

39. Han, X.-Q. *et al.* Ai-driven inverse design of materials: Past, present, and future. *Chin. Phys. Lett.* **42**, 027403 (2025).

40. Fu, C.-L. *et al.* Ai-driven defect engineering for advanced thermoelectric materials. *Adv. Mater.* **37**, 2505642 (2025).

41. Cheng, M. *et al.* Artificial intelligence-driven approaches for materials design and discovery. *Nat. Mater.* 1–17 (2026).

42. Cheng, M. *et al.* A foundation model for non-destructive defect identification from vibrational spectra. *Matter* 102728 (2026).

43. Chen, X. *et al.* Detecting thermodynamic phase transition via explainable machine learning of photoemission spectroscopy. *Newton* **1** (2025).

44. Nelson, K., Kreiman, T., Levine, S. & Krishnapriyan, A. S. Bridging the simulation-to-experiment gap with generative models using adversarial distribution alignment. *arXiv preprint arXiv:2604.01169* (2026).

45. Oord, A. v. d., Li, Y. & Vinyals, O. Representation learning with contrastive predictive coding. *arXiv preprint arXiv:1807.03748* (2018).

46. Nye, A., Leff, A., Barr, C. M. & Taheri, M. Direct observation of recrystallization mechanisms during annealing of cu in low and high strain conditions. *Scripta Materialia* **146**, 308–311 (2018).

47. Zhang, L., Han, J., Srolovitz, D. J. & Xiang, Y. Equation of motion for grain boundaries in polycrystals. *npj Comput. Mater.* **7**, 64 (2021).

48. Rohrer, G. S. *et al.* Grain boundary migration in polycrystals. *Annu. Rev. Mater. Res.* **53**, 347–369 (2023).

49. Thompson, A. P. *et al.* Lammps-a flexible simulation tool for particle-based materials modeling at the atomic, meso, and continuum scales. *Comput. physics communications* **271**, 108171 (2022).

50. Justo, J. F., Bazant, M. Z., Kaxiras, E., Bulatov, V. V. & Yip, S. Interatomic potential for silicon defects and disordered phases. *Phys. review B* **58**, 2539 (1998).

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- XPCSGrainBoundarySIclean.pdf