Digital Discovery

ROYAL SOCIETY OF CHEMISTRY

PAPER

View Article Online
View Journal | View Issue

Cite this: *Digital Discovery*, 2025, **4**, 3363

Material dynamics analysis with deep generative model

Duc-Anh Dao,[a] Minh-Quyet Ha,[a] Tien-Sinh Vu,[a] Shuntaro Takazawa,[b] Nozomu Ishiguro,[b] Yukio Takahashi,[b] Suzuki Masato,[c] Takashi Kakubo,[c] Naoya Amino,[c] Hirosuke Matsui,[d] Mizuki Tada[e] and Hieu-Chi Dam *[ab]

Understanding nanoscale material evolution—including phase transitions, structural deformations, and chemical reactions—under dynamic conditions remains a fundamental challenge in materials science. While advanced imaging techniques enable visualization of transformation processes, they typically capture only discrete temporal observations at specific time intervals. Consequently, intermediate stages and alternative pathways between captured images often remain unresolved, introducing ambiguity in analyzing material dynamics and transformation mechanisms. To address these limitations, we present a two-stage framework using deep generative models to probabilistically reconstruct intermediate transformations. Our framework is based on the hypothesis that generative models trained to reproduce experimental images inherently capture the dynamical processes that generated those observations. By integrating these trained generative models into Monte Carlo simulations, we generate plausible transformation pathways that interpolate unobserved intermediate stages. This approach enables the extraction of meaningful insights and the statistical analysis of material dynamics. This study also evaluates the framework's applicability across three phenomena: tantalum test chart translation, gold nanoparticle diffusion in polyvinyl alcohol solution, and copper sulfidation in heterogeneous rubber/brass composites. The generated transformations closely replicate experimental observations while revealing previously unrecognized dynamic behaviors for future experimental validation. These findings suggest that learned generative models encode physically meaningful continuity, enabling statistical interpolation of unobserved intermediate states and classification of transformation modes under sparse observational constraints.

Received 24th June 2025
Accepted 23rd September 2025

DOI: 10.1039/d5dd00277j

rsc.li/digitaldiscovery

1. Introduction

Understanding and controlling material transformations are essential for tailoring materials to achieve desired structural, mechanical, and functional properties in engineering applications. At the nanoscale, transformations such as phase transitions,[1,2] structural deformations,[3,4] and chemical reactions[5,6] significantly influence material stability, mechanical properties, and chemical reactivity. For instance, molecular chirality—the left- or right-handed absolute configuration of molecules—can lead to markedly different chemical behaviors in materials with identical atomic compositions.[7] Investigating these transformations is essential for designing materials with tailored properties, enabling advancements in catalysis, energy storage, and electronic devices.

Recent advanced imaging techniques, including scanning electron microscopy (SEM),[8] transmission electron microscopy (TEM),[9] and coherent X-ray diffraction imaging (CXDI),[10–12] enable researchers to track structural and compositional changes during material transformations. These methods generate sequential snapshots of material states, often captured in real-time or under *in situ* conditions. Conventional analyses model dynamic behavior from these discrete observations, as illustrated in Fig. 1a. However, these techniques face a fundamental limitation: their discrete temporal sampling inherent in these techniques captures only isolated moments within continuous transformation processes. This constraint is analogous to reconstructing a continuous dynamical system from sparse temporal observations—while each captured state reveals crucial information about the system's evolution, the intermediate dynamics driving transitions between observed states remain inaccessible.

[a]*Japan Advanced Institute of Science and Technology, 1-1 Asahidai, Nomi, Ishikawa 923-1292, Japan. E-mail: dam@jaist.ac.jp*
[b]*International Center for Synchrotron Radiation Innovation Smart, Tohoku University, 2-1-1 Katahira, Aoba, Sendai 980-8577, Japan*
[c]*Materials Performance Laboratory, Research and Advanced Development Division, The Yokohama Rubber Co., Ltd, Oiwake, Hiratsuka, Kanagawa 254-8601, Japan*
[d]*School of Science, Kwansei Gakuin University, 1 Gakuen Kamigahara, Sanda, Hyogo 669-1330, Japan*
[e]*Department of Chemistry, Graduate School of Science, Nagoya University, Furo, Chikusa, Nagoya, Aichi 464-8602, Japan*

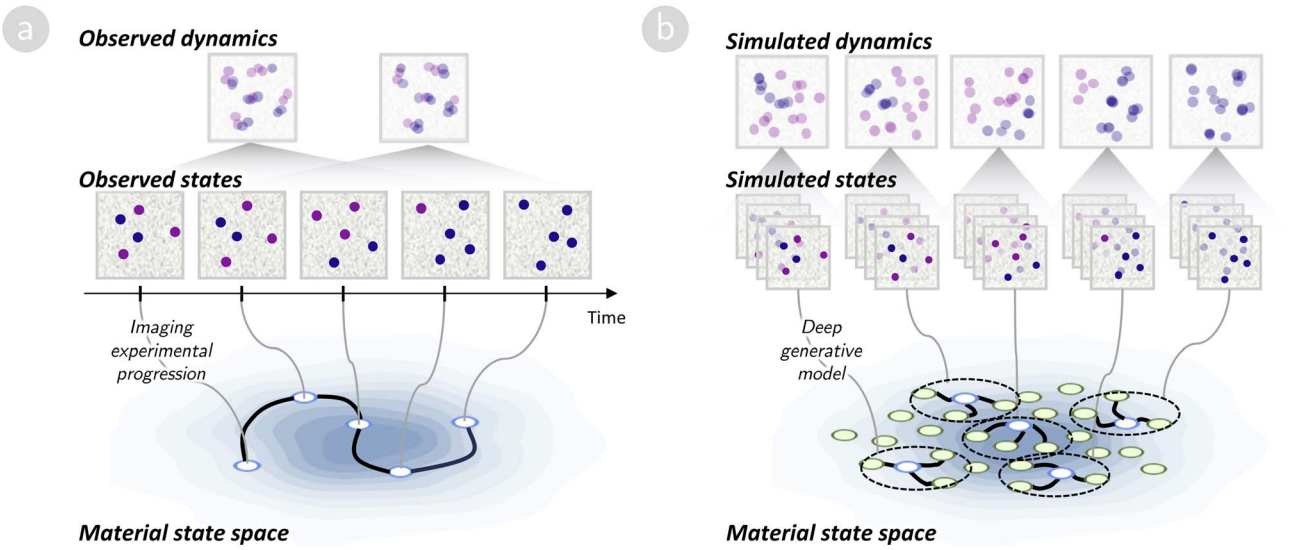

Fig. 1 Conceptual framework for material dynamics analysis. (a) Conventional approach: sequential snapshots of material states captured through microscopic imaging over time, showing discrete observations with unresolved intermediate states. (b) Proposed approach: ensemble of possible material states generated through Monte Carlo sampling in the learned latent space of a deep generative model, enabling exploration of transformation pathways between observed states.

To overcome these observational limitations, methods capable of simulating the transitions between captured sparse snapshots are highly required. Ideally, while denser experimental sampling would yield more complete datasets,[13,14] practical constraints in imaging equipment and measurement time often render this approach infeasible. Therefore, an alternative approach is to augment experimental data with computational methods that simulate plausible transitions between observed states.[15] This strategy forms the conceptual foundation of our approach, addressing data sparsity through statistically grounded analyses that extend beyond directly measurable transitions.

With recent advances, deep generative models offer a promising avenue to address the limitations of observable transitions.[16–19] Models such as generative adversarial networks (GANs),[20] variational auto-encoders (VAEs),[21] and diffusion models[22] are designed to learn mappings from high-dimensional image data to compact latent spaces that preserve essential structural features. Crucially, we hypothesize that by learning to generate experimental images, these models implicitly capture the underlying dynamical processes that produce the observed material states. A key advantage of these latent representations is their local smoothness—small perturbations in latent coordinates typically yield gradual and coherent variations in the generated images, despite the absence of explicit physical constraints during training. This property is particularly valuable for modeling material transformations—such as defect motion or phase-boundary migration—where latent geometry can reflect physical continuity in structural evolution.[23,24] Thus, exploring the latent space offers a potential mechanism to simulate hypothetical transitions that complement sparse experimental observations and statistically enrich the space of observed transformations.[16,18]

The local smoothness of latent spaces suggests a powerful approach to address the challenge of sparse observations. Given that small perturbations in latent coordinates typically yield gradual, coherent changes in generated images,[25,26] we can systematically explore neighborhoods around observed states to generate *material variations*—an ensemble of plausible intermediate states that could occur during transformation. This concept arises from recognizing that nanoscale systems rarely follow deterministic paths. Instead, thermal fluctuations, kinetic barriers, and competing energetic pathways create multiple possible routes between observed states. By sampling these variations through controlled exploration of the latent space, we transform the limitation of discrete observations into an opportunity for richer analysis. Notably, while we cannot directly observe all intermediate states, we can utilize the learned generative model to generate statistically plausible alternatives. This approach parallels established methods in statistical mechanics, where ensemble averages provide more robust insights compared with single trajectories. However, implementing this strategy requires careful consideration of the extent to which the exploration of latent space can be conducted while maintaining physical validity.

Based on these insights, we propose a framework that integrates deep generative modeling with statistical sampling and distributional analysis to characterize material transformations from experimental data with limited transition coverage. Our approach uses a generative model to map compact latent representations to hypothetical material states, enabling efficient navigation on the latent space. To systematically sample local transitions, we implement Monte Carlo (MC) sampling[27] in latent space to explore material variations around a given material state. This probabilistic strategy enables diverse sampling of latent codes under constraints that ensure the

generated transitions remain physically plausible while capturing variation in possible outcomes (Fig. 1b). To further analyze dynamic behaviors, the generated states are grouped into variations that approximate the collective range of transformations possible from the same initial condition, rather than representing single predictions. This approach, based on local sampling and statistical aggregation, enables statistical interpolation of unobserved intermediate states and classification of transformation behaviors, even in the absence of explicit time-series data. Thus, the proposed framework extends generative models' representational capacity, enabling the interpretation of transformation behavior through physically grounded, data-driven reasoning.

2. Methodology

To implement the described concepts, the proposed framework for analyzing material dynamics follows two systematic stages, as shown in Fig. 2: (1) generative model training, where a deep generative model is employed to synthesize microscopy images and construct a structured latent space for material states; and (2) MC simulation, which utilizes the trained model to generate and analyze material variations representing possible transformations from a given material state. Hereinafter, we define a *material variation* as a localized ensemble of synthetic material states derived from perturbations around a latent embedding of a reference configuration. This representation captures the stochastic neighborhood of dynamic transitions without requiring temporally resolved sequences.

2.1. Generative model training

In this stage, we develop a deep generative model to construct a continuous latent space that enables sampling of material states while connecting these states to their image representations. This design is based on the hypothesis that, by learning to reproduce the distribution of experimental images, the generative model implicitly encodes the underlying dynamical

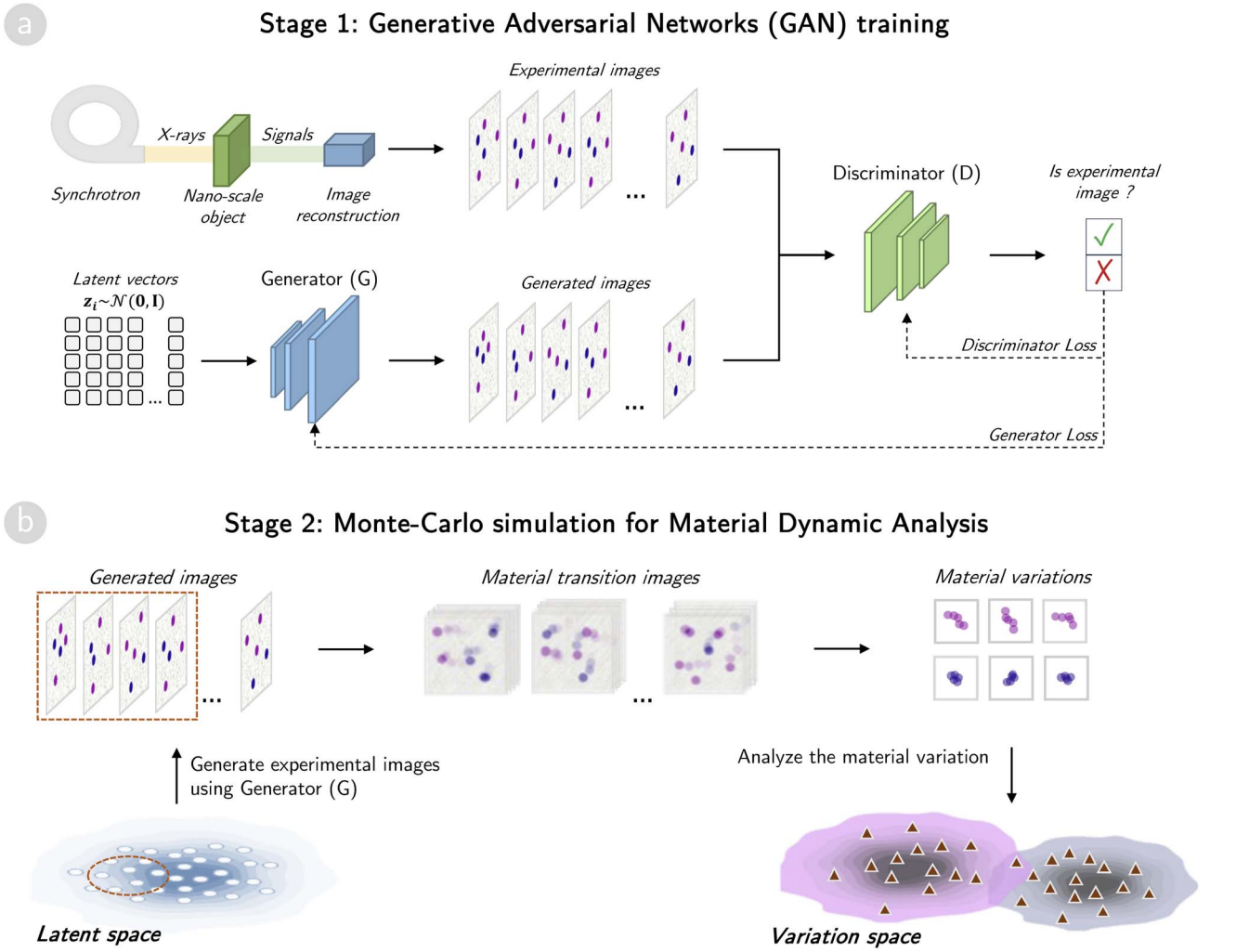

Fig. 2 Overview of the two-stage analytical framework for extracting insights into material dynamics from experimental images. (a) Deep generative model training: a GAN is trained on experimental images to construct latent space of material states. (b) Monte Carlo simulation: the trained model generates material variations through latent space perturbations, which are then analyzed to identify distinct transformation behaviors and dynamic patterns.

processes that generate the observed material configurations. This generative approach provides a structured statistical representation of material configurations, supporting systematic exploration of material variations—structurally similar configurations that may plausibly evolve from a common initial material state over short timescales—in subsequent analyses.

For this construction, we implement GANs,[28] selected for their effectiveness in generating high-quality images from limited training data. Our GAN approach constructs a latent space where the original data instances follow a multivariate normal distribution $\mathcal{N}(0, I)$, enabling systematic sampling of the material state space. The adversarial training process involves a generator network that produces realistic material images, whereas a discriminator network learns to distinguish between generated and real experimental images, progressively improving generation quality. Compared with alternative approaches, GANs offer a practical trade-off between fidelity and efficiency. In contrast, the variational sampling principle in VAEs imposes an additional constraint by enforcing reconstructions through a latent distribution, which tends to blur fine-scale features and restricts analysis of delicate structural or compositional details, particularly at the nano- or microscale. Diffusion models[29] require extensive computational resources and large datasets, making them less practical in this context.

The GAN consists of two competing neural networks: a generator G and a discriminator D, trained on a dataset $\mathcal{D}$ of experimental material images, as shown in Fig. 2a. Both networks use convolutional layers to capture the spatial features present in the material data. The generator G: $\mathbf{Z} \rightarrow \mathbf{X}$ maps a latent vector $z \in \mathbb{R}^d$ from a d-dimensional latent space $\mathbf{Z}$ to a synthetic image $\hat{x}$ in image space $\mathbf{X}$, simulating a possible material configuration. The discriminator D: $\mathbf{X} \rightarrow \mathbb{R}$ assigns higher scores to real images and lower scores to generated ones.

To train the model, we employ the *Wasserstein loss function*,[30] which provides a stochastic objective for comparing real and generated distributions:

$$\min_{\mathrm{G}} \max_{\mathrm{D}} \mathcal{L}(\mathrm{D}, \mathrm{G}) = \mathbb{E}_{x \sim \mathbb{P}_{\mathcal{D}}}[\mathrm{D}(x)] - \mathbb{E}_{z \sim \mathbb{P}_{\mathbf{Z}}}[\mathrm{D}(\mathrm{G}(z))], \quad (1)$$

where $\mathbb{P}_{\mathcal{D}}$ and $\mathbb{P}_{\mathbf{Z}}$ are distributions of observed images x in dataset $\mathcal{D}$ over image space $\mathbf{X}$ and the prior distribution of latent vector z over the latent space $\mathbf{Z}$, respectively. Accordingly, $x \sim \mathbb{P}_{\mathcal{D}}$ denotes a material image sampled from the dataset $\mathcal{D}$ and $z \sim \mathbb{P}_{\mathbf{Z}}$ are latent vectors sampled from the multivariate normal distribution $\mathcal{N}(0, I)$ over the latent space $\mathbf{Z}$.

In material dynamics modeling, the generator is trained to produce images that simulate physically plausible dynamically evolving material states, such as those induced by diffusion, deformation, or sulfidation. These synthetic images represent hypothetical progressions along dynamic trajectories, capturing variations that may be absent from experimental observations. The discriminator acts as a critic, emphasizing realistic image features and enabling the model to internalize key structural and compositional characteristics of material evolution.

To improve training stability and image fidelity, we implement a gradient penalty[31] along with a progressive growing strategy,[32] which begins with low-resolution images and incrementally increases resolution. This multi-scale approach facilitates the learning of both coarse and fine features essential for capturing hierarchical material structures. The configuration of model architecture and training schedule across material systems follows the logic of progressive growing: training begins at low resolution to stabilize global structures, and layers with gradually reduced filter counts are added as resolution increases to capture finer details. Accordingly, higher-resolution datasets require deeper networks and finer resolution schedules, while batch sizes are scaled down to maintain stability, with stricter reductions for volumetric data due to cubic scaling (see Tables S1 and S2). All available experimental data were used for training without partitioning into separate sets, as the objective was not to predict unseen outcomes but to model the full distribution of material states. Model quality was assessed by validating that generated samples reflect the statistical and physical properties of the training data.

2.2. Monte Carlo simulation for material dynamic analysis

In materials science, a material state refers to a specific configuration of atoms, molecules, or microstructures at a given moment. While experimental imaging captures discrete snapshots of these states, it cannot fully resolve the continuous transitions between them. Therefore, we employ MC simulation to examine plausible transformations and material evolution over time.

Our MC approach, illustrated in Fig. 2b, models short-range transformations within the GAN-generated latent space. We adopt a first-order Markov assumption where transformations depend primarily on the current state. This assumption enables us to focus on local changes rather than reconstructing full dynamic trajectories, which require densely sampled, high-resolution data rarely obtainable without specialized instrumentation. By constraining transitions to experimentally plausible magnitudes, we ensure physical relevance while enabling probabilistic analysis of material dynamics without the need to model long-range sequences.

We sample M latent vectors z_i from the normal distribution $\mathcal{N}(0, I)$ over the latent space $\mathbf{Z}$ and input them into the generator G to produce material state images $\hat{x}_i = \mathrm{G}(z_i)$. For each sampled state as a reference, we generate N images of multiple transitioned states by perturbing its latent vector with a stochastic deviation ε_j (where $j = 1, \ldots, N$), as follows:

$$\hat{x}_{i,j} = \mathrm{G}(z_{i,j}) = \mathrm{G}(z_i + \varepsilon_j), \quad (2)$$

where each component of ε_j is drawn from a uniform distribution $\mathcal{U}_{[-\gamma,\gamma]}$. We determine γ by correlating the maximum observed micro-structural displacement over time intervals with the average latent space distances between adjacent time-step reconstructions. The uniform distribution ensures equal sampling probability within a controlled radius around each reference state, providing comprehensive coverage of local transformations.

Latent perturbations can be interpreted as surrogates for stochastic fluctuations in material states under thermal, chemical, or mechanical influences. Although the latent space

is not explicitly linked to thermodynamic or kinetic parameters, its geometry reflects structural similarity, enabling the interpretation of small perturbations as physically feasible local transformations. Therefore, γ serves as a proxy for the expected magnitude of change under short-time dynamics or environmental variations. To select γ, we compare generated transformations to the structural differences observed between experimentally yielded states sampled at the highest available temporal resolution, ensuring that perturbations represent short-range, physically plausible transitions.

The sampled states $\hat{x}_{i,j}$, representing transitions associated with the reference state $\hat{x}_i$, are grouped into a set denoted as a *material variation* V_i. This set represents possible transformations associated with $\hat{x}_i$ over short timescales, expressed as:

$$V_i = \{\hat{x}_i, \hat{x}_{i,1}, \hat{x}_{i,2}, \ldots, \hat{x}_{i,N}\}. \quad (3)$$

Each material variation V_i is then characterized by the statistical distribution of reference state $\hat{x}_i$ and its N transitioned states $\hat{x}_{i,j}$, reflecting patterns in the transformations occurring within V_i. For each case study, we design features that represent characteristic variations across the sampled state images, guided by domain knowledge of the associated physical processes. These features serve as representations of each material variation V_i.

Subsequently, we apply unsupervised clustering techniques to the space of material variation descriptors to identify representative transformation behaviors across the dataset. This approach enables systematic grouping of variations based on both discrete and continuous transformation patterns, offering insight into dominant dynamic modes. Notably, we do not impose directionality on these transitions. Given that the transformations are derived from differences between generated image samples in the GAN latent space, they are not constrained to a fixed direction of change. Instead, each pair of states can be interpreted as either a forward or backward transition, reflecting the symmetric nature of local variations in this representation. This general formulation is suitable for cases where transition order is uncertain or experimentally unresolved while still preserving local physical plausibility.

3. Results and discussion

3.1. Experimental settings

We evaluate the effectiveness of our proposed GAN-based framework for analyzing material dynamics through a series of carefully structured experiments. These experiments emphasize both quantitative validation against experimental measurements and the extraction of novel insights beyond those obtainable through conventional analytical methods. Building on our previous studies,[1,12,33] we assessed the framework's performance using three primary criteria:

1. Image fidelity: evaluating whether the generated images of material states accurately reflect physically plausible configurations that align with real experimental observations.

2. Latent space smoothness: verifying that small perturbations in latent vectors produce gradual and coherent changes in generated images. This criterion assesses the model's capacity to reflect smooth physical transformations, such as defect movements or boundary migrations.

3. Insight consistency with prior studies: comparing the physical insights obtained from our generative analysis with those from established conventional approaches, emphasizing consistency rather than superiority.

To comprehensively evaluate these criteria, we selected three distinct yet progressively challenging case studies (Table 1), each designed to systematically increase the complexity of both the imaging data and the underlying physical mechanisms. The first case study served as a controlled, proof-of-concept scenario using two-dimensional (2D) phase images obtained through coherent X-ray diffraction imaging (CXDI) of a Ta test chart.[33] These images captured simple, precisely controlled translational and rotational motions without internal structural changes, forming a baseline for validating the fundamental capabilities of the generative model.

Building on this foundational scenario, the second case study also employed CXDI to reconstruct phase images capturing the diffusion of gold nanoparticles (NPs) in a polyvinyl alcohol (PVA) solution.[12] In this study, the dataset featured stochastic and complex NP motions, testing the framework's ability to model and interpret realistic diffusive processes.

The third case study further increased complexity by examining 3D volumetric images obtained using X-ray absorption fine structure computed tomography (XAFS-CT).[1] These 3D images comprise three separate channels corresponding to Cu valence states (Cu^0, Cu^{1+}, Cu^{2+}), offering insights into the chemical degradation (aging) processes within brass–rubber composites over time.

The model architecture is adapted for each case, with the number of convolutional layers and filter sizes tailored to the specific characteristics of each dataset. In the 2D cases (Ta test chart and NP diffusion), we employed 2D convolutional GANs.

Table 1 Summary of experimental datasets used for generative model training and material dynamics analysis. The datasets represent diverse material systems imaged using advanced synchrotron-based techniques (CXDI: Coherent X-ray Diffraction Imaging; XAFS-CT: Hard X-ray Absorption Fine Structure Computed Tomography)

Datasets	No. images (no. channels)	Image size	Pixel resolution	Imaging techniques
Ta test chart[33]	10 530 (1)	178 × 178 pixels	20.72 nm per pixel	CXDI
NP diffusion[12]	2000 (1)	142 × 142 pixels	40.58 nm per pixel	CXDI
Aging brass clumps[1]	4956 (3)	32 × 32 × 32 voxels	0.65 μm per voxel	XAFS-CT

Conversely, for the 3D XAFS-CT case, we used a 3D convolutional GAN to better capture volumetric features. All models were trained on an NVIDIA A100 graphics processing unit (GPU) with 80 GB of memory. A complete description of the training configuration, including stage-wise growth schedules tailored for specific datasets, is provided in Section S1.

Throughout this study, we define a *material variation* as the set of synthetic material states generated by perturbing the latent vector of a single reference configuration. Thus, each variation represents an ensemble of physically admissible, short-range transitions accessible from that configuration and constitutes the fundamental unit of our analysis. For clarity, we adopt context-specific labels—*NP variations* for diffusion-driven displacements of NPs and *clump variations* for sulfidation-induced morphological changes in brass particles—while maintaining a unified underlying concept. In all cases, a *material variation* functions as a structured surrogate for local material dynamics, capturing the spectrum of plausible short-range evolution.

3.2. Proof-of-concept case study of Ta test chart translation

Before applying our framework to complex material systems, we first validate a key assumption that the learned latent space encodes both static material states and meaningful transformation dynamics. To test this hypothesis, we conduct a controlled experiment using a rigid-body system—a Ta test chart subjected to prescribed translation and rotation. This case study offers a ground-truth scenario involving only rigid-body motion without internal deformation. Therefore, any deviations or deformations in the generated images must reflect limitations in the model's learned representation rather than unobserved material behavior.

The dataset is acquired using CXDI to capture the horizontal translation of a Ta test chart moving at 340 nm s^{-1} under triangular aperture illumination.[33] The raw dataset consists of 1755 phase-reconstructed frames, each representing a 7 ms time window of the Ta test chart's continuous motion. To enhance the statistical robustness of our generative model while maintaining physical realism, we augment the dataset through controlled in-plane rotations. Specifically, each original frame is supplemented with five additional images generated by applying random rotations within the range of $[-30°, 30°]$, expanding the total dataset to 10 530 images. For visualization, images were center-cropped to highlight the most coherent region unaffected by peripheral artifacts, as shown in Fig. 3a. However, for model training, full image frames—including boundary artifacts—were retained to maximize the perceivable structural information. The GAN was trained on this dataset to capture the statistical distribution of physically plausible material states and their potential transformations (see Section S1.1 for details).

To evaluate whether the latent space encodes smooth, plausible transitions, we conduct an MC sampling step using the trained GAN. A latent vector z was sampled from the standard multivariate Gaussian prior and perturbed by additive noise drawn from a uniform distribution $\mathcal{U}_{[-1,1]}$ (see more details in Section S3.1). While this range may seem large relative to the unit-variance prior, it corresponds to moderate deviations in high-dimensional latent spaces owing to their geometric properties (see Section S2). This process yields a local variation set V_z, consisting of images that represent material states that can transform between each other. Although these generated images are not temporally ordered, they collectively reflect plausible short-range transitions in the configuration space.

Fig. 3b and c show two representative image sequences, compiled *post hoc* from variations generated in V_z. These sequences are presented to illustrate that latent-space sampling captures the inherent mechanisms underlying the observed rigid-body dynamics: one representative sequence highlights horizontal translation and the other in-plane rotation. To construct these illustrative sequences, we selected subsets of generated images and sorted them according to quantitatively evaluated transformation magnitudes, such as translation offsets or rotation angles, computed *via* a template-matching procedure applied to locally perturbed latent samples (see Section S3.2 for details). Fig. 3b illustrates smooth translational motion along the x-axis, whereas Fig. 3c captures rotational steps of approximately 1–2 degrees. Full sequences are provided in Videos S1 and S2. Compared with the original experiment, the generated translation appears slightly slower, yet the spatial coherence across frames supports the interpretation that the latent space encodes physically meaningful, continuous variations. Intersection-over-union (IOU) distributions between consecutive generated frames and those in the experimental sequence are shown in Fig. 4a. The higher IOU values in the synthetic data suggest that the sampled transitions correspond to smaller structural changes than those observed experimentally, which is consistent with short-range interpolations.

To quantitatively verify the plausibility of the sampled transitions, we analyze their consistency with ground-truth rigid-body dynamics. Given that the transformation in this case study involves only rigid-body motion, which implies that the Ta test chart moves without internal deformation, any geometric inconsistencies in the generated images must originate from limitations in the model's latent representation. To detect such deviations, we track two structural features—slit-like probe lines labeled d_1 and d_2—across all images in the sampled variation set V_z, originating from a reference image I_z, as shown in Fig. 4b. We apply template-matching-based tracking[34] to robustly localize and quantify the dimensions of these features throughout the variation set (more details of the process is shown in Section S3.2). As shown in Fig. 4b, the generated images consistently preserve the contrast between transmitted and absorbed regions, enabling clear delineation of the slits. Notably, the measured slit widths in the generated images are $d_1 = 309 \pm 11$ nm and $d_2 = 340 \pm 12$ nm, closely aligning with the reference widths in I_z and comparable to those in the observed dataset: $d_1 = 311 \pm 13$ nm and $d_2 = 324 \pm 11$ nm. The resulting deviations (26–39 nm) fall within approximately two pixels (1 pixel $\approx$ 20.72 nm), indicating that the generative model preserves key structural features and does not introduce significant deformations during localized transformations.

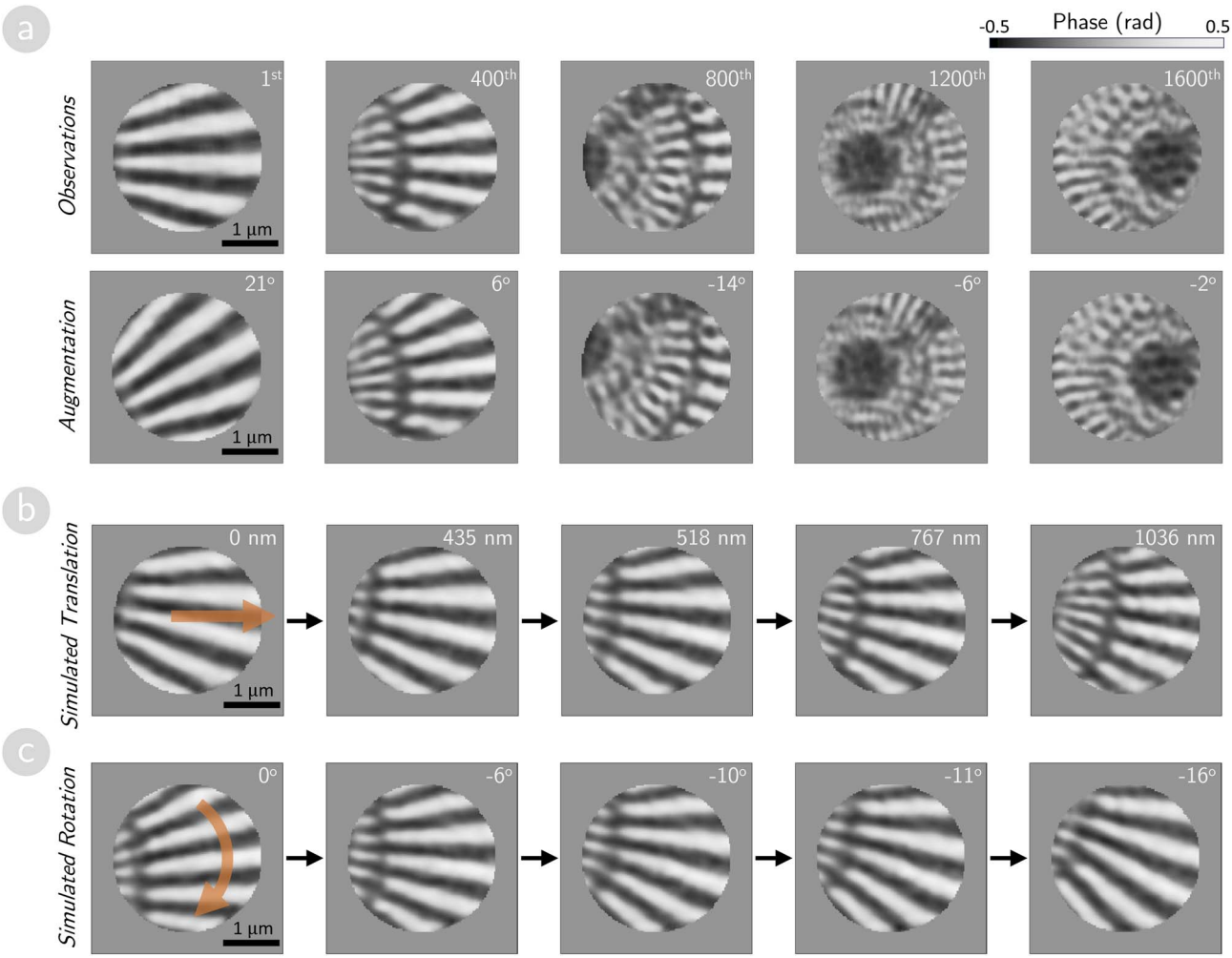

Fig. 3 Sampled phase images depicting rigid-body motion progression of a Ta test chart. (a) Experimental phase images documenting controlled translation and rotation of the Ta test chart. Upper row: Sequential phase images capturing horizontal translation at frames 1st, 400th, 800th, 1200th, and 1600th. Lower row: Corresponding augmented images with applied in-plane rotations, where annotations indicate rotation angles relative to the original experimental frames. (b and c) Representative sequences of GAN-generated images illustrating smooth progressions of translational motion (b) and rotational motion (c) extracted from MC sampling in GAN's latent space. Annotations indicate cumulative displacement distances or rotation angles relative to the initial frame in each sequence, demonstrating the model's ability to generate physically coherent transformation pathways.

Overall, these results support the core hypothesis of our framework—that localized traversal in the GAN latent space, even without explicit temporal supervision, can expose transformation patterns consistent with physical dynamics.

3.3. Investigating dynamics of nano-particles diffusing in aqueous media

Having validated our framework using controlled transformations in the Ta test chart experiment, we subsequently applied it to a more complex, stochastic system: gold NP diffusing in an aqueous PVA solution. This case study assesses the framework's capacity to capture random, thermally driven particle motion, providing insights into diffusion mechanisms and NP–solution interactions. Unlike the deterministic rigid-body motion of the previous study (Section 3.2), the NP–PVA system exhibits stochastic dynamics governed by Brownian motion, introducing significant variability across spatial and temporal scales.

The experimental dataset comprises 2000 sequentially captured, phase-reconstructed images from 2D CXDI, recording real-time NP movement in the PVA solution.[12] Each image has a resolution of 142 × 142 pixels, with a spatial resolution of approximately 40.58 nm per pixel. These images represent snapshots of NP arrangements evolving under thermal agitation, providing an ideal dataset to evaluate the generative model's ability to represent and interpolate diffusive dynamics.

3.3.1. Evaluating fidelity of sampling transitions. We trained a GAN model on this dataset to learn the statistical distribution of plausible NP configurations and their local variations. Detailed model architecture and training procedures are provided in Section S1.4. Fig. 5a (left column) shows representative experimental frames sampled from the 600th,

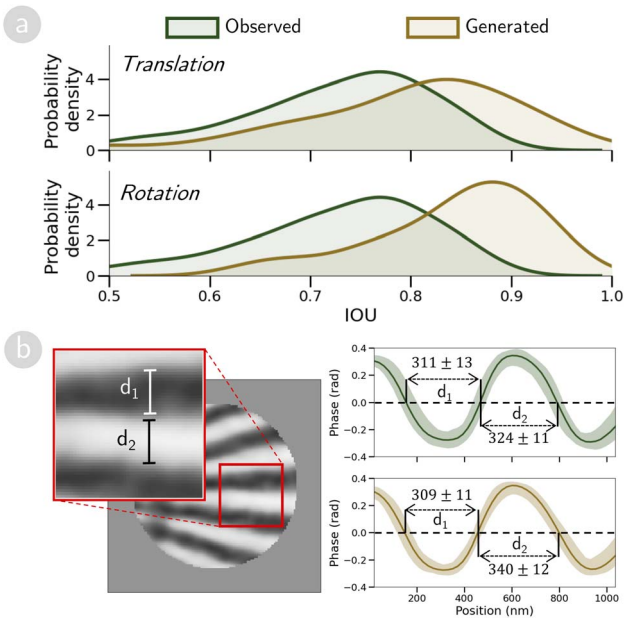

Fig. 4 Proof-of-concept validation of generative modeling on motions of Ta test chart. (a) Intersection-over-union (IOU) distributions between consecutive frames for translation and rotation sequences, comparing observed (green) and generated (yellow) data. (b) Comparison of slit widths extracted from observed and generated CXDI images, where solid lines represent average slit width in the matched region, and the spectrum illustrates width fluctuations.

900th, and 1200th time steps. Corresponding synthetic images, selected from 1000 GAN-generated samples, are shown in Fig. 5a (right column). These samples were generated by decoding latent vectors randomly drawn from a standard normal distribution. The generated samples visually resemble their experimental counterparts, preserving the number, size, and spatial coordination of NPs. This resemblance suggests that the trained model effectively captures both structural fidelity and variation patterns characteristic of the diffusive process. To ensure consistent analysis, we used a validated image-processing method[12] to extract the positions and morphologies of individual NPs over time (see Section S1.2). A comparison of the NP area distributions from observed and generated images (Fig. 5b) yielded a low Wasserstein distance of 0.068, indicating close statistical agreement and supporting the physical plausibility of the generated configurations.

Beyond assessing the fidelity of generated images, we also evaluate the *smoothness* of the GAN's latent space. Specifically, we assess whether small perturbations in latent vectors lead to gradual and coherent changes in generated images. This criterion reflects the model's capacity to represent plausible short-range NP movements. To examine this hypothesis, we generate 100 perturbed variants around each of 1000 latent vectors sampled from a standard normal distribution. Perturbations are then applied by adding noise drawn from a uniform distribution $\mathcal{U}_{[-\gamma,\gamma]}$, with the γ varied from 0.25 to 2.0 in

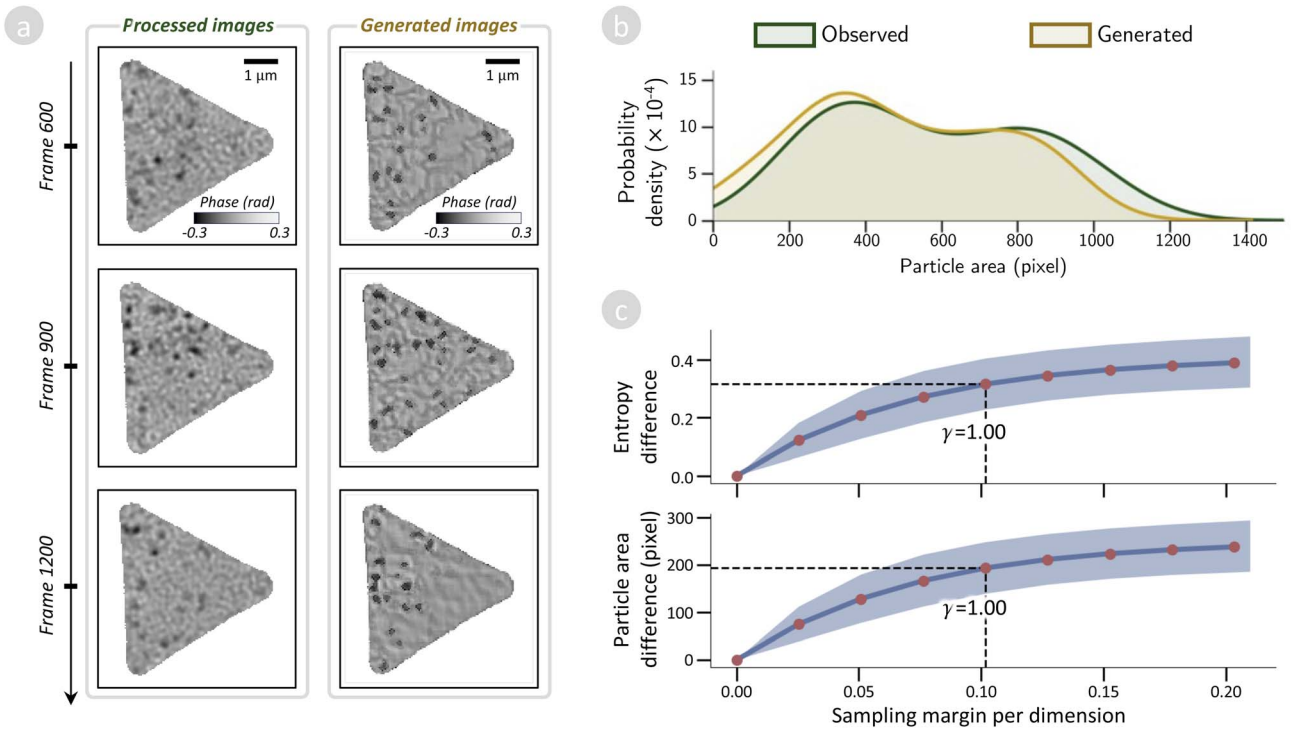

Fig. 5 Evaluation of fidelity for generated NP configurations and diffusion dynamics. (a) Observed and generated phase images as snapshots depicting NP configuration when diffused in a PVA solution. From upper to lower, each row respectively represents observed phase images (for training GAN models) and GAN-generated phase images to represent hypothetical NP configuration under diffusion. (b) Comparison between particle density represented in the observed and generated phase images. (c) Latent space continuity analysis: entropy differences and particle area differences (pixels) as functions of sampling margin per dimension (γ normalized by latent space dimensionality), with $\gamma = 1.00$ matching experimental frame-to-frame variations.

increments of 0.25. This approach enables the analysis of how latent space distance influences structural variations in the generated images. To quantify how these perturbations affect image structure, we measure two metrics between reference and perturbed images: (1) image entropy difference, and (2) particle area difference. Image entropy[35] measures the statistical complexity of grayscale distributions and reflects structural richness in the image. Notably, higher entropy indicates more complex or varied patterns.

Interpreting distances in high-dimensional latent spaces is challenging because absolute distances often lose meaning due to the *curse of dimensionality*.[36,37] To address this, we measure the average Euclidean distance between each reference latent vector and its perturbed samples and normalize this average by the number of latent dimensions (see details provided in Section S2). This normalization yields an intuitive measure of locality within the latent space. Consequently, perturbation margins of $\gamma = 0.25$ and $\gamma = 2$ correspond to average per-dimension changes of approximately 0.025 and 0.20, respectively.

As shown in Fig. 5c, both metrics increase smoothly with larger γ, indicating that the latent space encodes a gradient of structural variation. At lower margins (*e.g.*, $\gamma = 0.25$), generated variations show minimal differences from their reference images. Conversely, at higher margins, substantial differences in image entropy and particle area emerge, reflecting broader configurational changes in NP arrangements. Notably, a sampling margin of $\gamma = 1$ (equivalent to a per-dimension sampling margin of approximately 0.1) yields differences in entropy and particle area closely matching those observed between consecutive experimental frames. At this level, the entropy difference corresponds to the upper bound observed between consecutive experimental frames (0.26), and particle area differences remain below 200 pixels. This observation aligns with NP depletion rates of approximately 10 particles per second,[12] indicating that latent vectors sampled within this margin produce configurations reflecting the natural stochastic motion of NPs over a timescale of about one second. Therefore, in subsequent simulations of dynamic behavior, we adopt $\gamma = 1$ to generate physically plausible transitions.

3.3.2. Extraction of dynamic behaviors of NP diffusion. Based on the proximity analysis of GAN's latent space, we conduct MC sampling with a sampling margin $\gamma = 1$ to generate variations. Starting from 200 reference states, we generate perturbed samples around each one, yielding 500 sampled images. We then retain only those with an entropy difference ≤ 0.2, a threshold corresponding to the upper bound 0.26 of the experimentally observed transition between consecutive phase images (see more details in Section S4.1). The retained samples are then pixel-wise averaged to produce images summarizing local NP diffusion over intervals of approximately 1 s. Each of these averaged images encodes the diffusion traces of multiple NPs, including multiple NP variations associated with the same reference state. From each image, we extract individual NP variations—each representing the diffusion trace of a single NP—characterizing them using two descriptors: anisotropy and

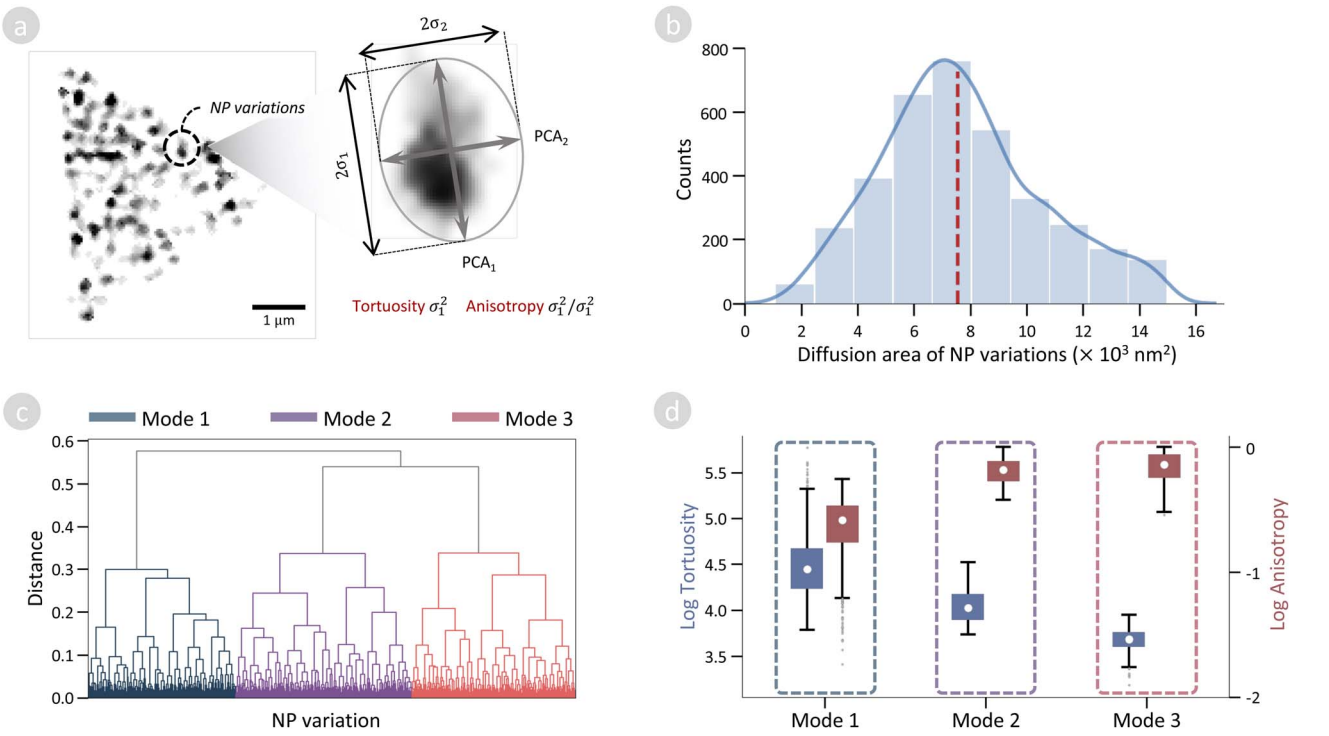

Fig. 6 Analysis of NP diffusion modes in PVA solution. (a) Schematic showing how motion descriptors (anisotropy and tortuosity) are extracted from a single NP variation. (b) Distribution of diffusion areas measured from sampled NP variations, with mean value indicated by red line. (c) Hierarchical clustering dendrogram revealing three distinct diffusion modes. (d) Box plots showing anisotropy and tortuosity distributions for each identified mode, with outliers shown as gray dots.

tortuosity, as illustrated in Fig. 6a (see more details of this process in Section S4.2). Anisotropy quantifies the directional preference of each diffusing motion, while tortuosity captures the degree of path irregularity.[12,38,39] Notably, these descriptors are widely used in both biological and materials diffusion studies for distinguishing transport modes across heterogeneous media. Fig. 6b shows the distribution of diffusion areas derived from the extracted NP diffusion traces. The mean sampled diffusion area over approximately 1 s intervals is 7712 $\pm$ 2862 nm^2, which closely aligns with the diffusion rate of 7550 $\pm$ 850 nm^2 per second, measured by X-ray Photon Correlation Spectroscopy (XPCS) in a previous study.[12]

To quantitatively compare these NP variations, we construct a distance matrix using a metric scaled to account for the range and distribution of the descriptors (see Section S4.3 for details). We then apply hierarchical agglomerative clustering (HAC) using the average linkage method[40,41] to identify structurally coherent groups based on geometric separation in descriptor space. This approach enables the classification of distinct diffusion modes by grouping trajectories with similar path characteristics. As shown in the dendrogram in Fig. 6c, three primary clusters emerge, suggesting distinct classes of NP motion within the PVA solution.

Fig. 6d shows the distribution of anisotropy and tortuosity values across these clusters. The first cluster exhibits directionally biased diffusion with lower tortuosity, indicating persistent movement potentially influenced by channel-like structures or local field gradients within the polymer.[42] The second cluster exhibits random, isotropic diffusion characterized by high tortuosity and low directional persistence—features typical of classic Brownian motion and consistent with unstructured diffusion in homogeneous fluids.[43] The third cluster consists of highly confined diffusion characterized by limited displacement and no dominant direction, suggesting transient immobilization or trapping—a phenomenon known to occur in polymer matrices and crowded soft-matter environments.[44,45]

3.4. Investigating dynamics of sulfidation in aging process of rubber/brass composites

Having demonstrated our framework's effectiveness in analyzing systems with well-defined or partially understood dynamics—from rigid-body motion in a Ta test chart to stochastic diffusion of NPs—we now apply it to a more complex and less-characterized phenomenon: the sulfidation dynamics of copper within aging rubber/brass composites. This process is critical for the durability and performance of steel-cord-reinforced rubber tires, where brass-coated wires embedded in a rubber matrix significantly enhance interfacial adhesion. Despite its industrial importance, the physicochemical mechanisms driving Cu sulfidation during aging are not well quantified, with most existing studies offering primarily empirical descriptions.

To examine these sulfidation dynamics, we analyzed a 3D imaging dataset acquired using XAFS-CT[1] at the SPring-8 synchrotron facility.[46] This advanced imaging technique enables precise 3D mapping of elemental compositions and oxidation states, allowing spatially resolution of metallic Cu, monovalent Cu_2S, and CuS distributions within rubber/brass composites. The dataset was generated through controlled aging experiments conducted in a previous study:[1] approximately 2000 brass particles (Cu/Zn ratio = 75/25) were embedded within a rubber matrix, heated at 443 K for 10 min to form adhesive layers, and subsequently aged under controlled conditions. The aging process and corresponding XAFS-CT imaging were repeated over multiple intervals, producing imaging datasets at aging times of 0, 3, 14, and 28 days, denoted as Aging-0, Aging-3, Aging-14, and Aging-28, respectively.

From this time-resolved dataset, we systematically tracked the chemical evolution of individual brass particles throughout the entire aging period. Initially, 6590 particles were extracted from the experimental data (details of the data collection process are shown in Section S1.3). However, to ensure consistent volumetric representation within our generative modeling framework, we retained only particles whose dimensions fit entirely within a 32 $\times$ 32 $\times$ 32 voxel grid (voxel size = 0.65 μm per voxel). Particles exceeding this spatial constraint were excluded from subsequent modeling. This preprocessing step resulted in a final dataset comprising 4956 particles, each tracked across the aging intervals (0, 3, 14, and 28 days).

3.4.1. Evaluating fidelity of sampling transitions. We trained a GAN to model the statistical distribution of particle transformations during aging. Complete model and training details are provided in Section S1.4. Fig. 7a shows the progressive chemical evolution of a representative particle observed experimentally. Corresponding synthetic images, selected from 1000 GAN-generated samples, are shown in Fig. 7b. These synthetic images closely align with experimental observations in particle shape and distribution of Cu valence states (see a comparison with VAE-generated images in Section S1.5), indicating that the GAN effectively captures structural and chemical changes throughout aging. Quantitatively, intensity distributions for each Cu state (Cu, Cu_2S, and CuS) show minimal differences between observed and synthetic data (Wasserstein distances: Cu, 0.010; Cu_2S, 0.017; CuS, 0.006), confirming the physical fidelity of the generated particle states (Fig. 7c).

Beyond assessing the fidelity of generated images, we evaluated the *smoothness* of the GAN's latent space, specifically, whether small perturbations in latent vectors lead to gradual and coherent changes in generated images. We applied the same local perturbation protocol described in Section 3.3.1, generating 100 perturbed variants around each of 1000 latent vectors previously sampled from the standard normal distribution. Perturbations were applied by adding noise drawn from a uniform distribution $\mathcal{U}_{[-\gamma,\gamma]}$, with the γ varied from 0.25 to 2.0 in increments of 0.25. To quantify the deviation of perturbed samples from their respective references, we measured two metrics: (1) the complement of the 3D IOU image (IOU_{3D}), which captures shape-based dissimilarity (see Section S5.1), and (2) total compositional density difference, which quantifies voxel-wise changes in elemental distribution. As shown in Fig. 7d, both metrics gradually increase with increasing γ,

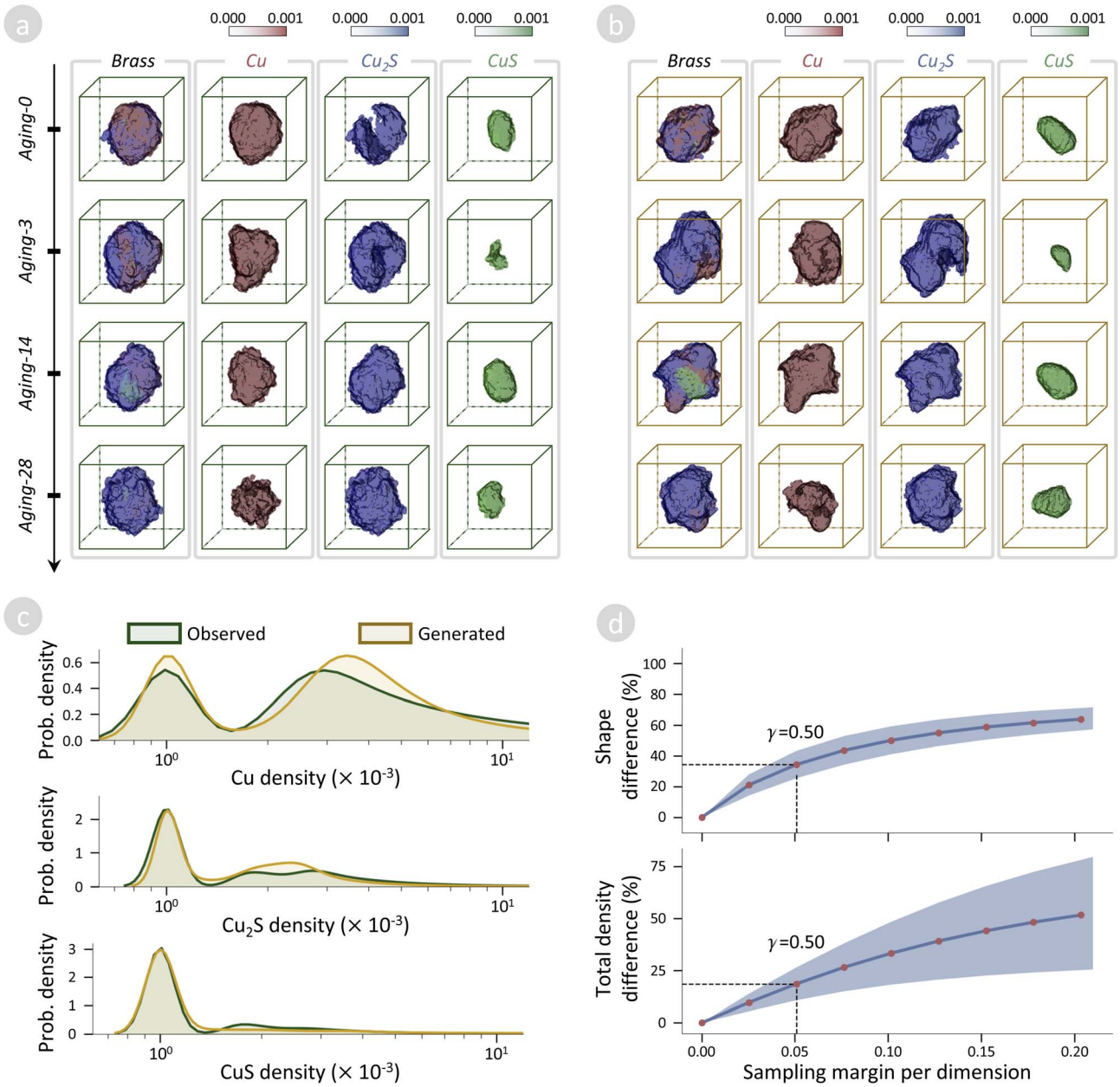

Fig. 7 Evaluation of fidelity for generated brass clumps and transitions. (a) Experimental 3D XAFS-CT images showing brass clump evolution during aging (0–28 days), with separated visualizations of brass matrix and copper species (Cu, Cu_2S, CuS). (b) GAN-generated images representing plausible sulfidation states at corresponding aging stages. (c) Density distributions of copper species comparing observed (solid lines) and generated (shaded areas) images, demonstrating close statistical agreement. (d) Latent space continuity analysis: shape differences (%) and total density differences (%) as functions of sampling margin per dimension (γ normalized by latent space dimensionality for intuitive evaluation of perturbation magnitude), with $\gamma = 0.5$ selected for subsequent dynamic analysis.

indicating that the latent space encodes a continuous and interpretable spectrum of compositional transformations.

Notably, a sampling margin of $\gamma = 0.5$ (equivalent to a per-dimension sampling margin of approximately 0.05) yields an $IOU_{3D} \approx 0.7$ and total compositional mass deviations under 20%, closely matching observed differences between adjacent sulfidation stages.[1] This suggests that latent vectors sampled within this margin produce brass clumps consistent with natural aging processes observed within the experimental time window. Therefore, in subsequent simulations of aging behavior, we adopt $\gamma = 0.5$ to generate physically plausible transitions.

3.4.2. Extraction of dynamic behaviors of sulfidation. Based on the latent space proximity analysis, we selected a sampling margin of $\gamma = 0.5$ (corresponding to an average distance of 0.05 per dimension) for MC sampling. From 3000 reference latent vectors, 400 perturbed samples were generated per reference. Accordingly, to enforce physical validity, we applied two constraints: (1) compositional mass deviation $\leq$20%, and (2) $IOU_{3D} \geq 0.5$ to retain morphological identity.

Invalid samples, likely representing cross-clump or unrealistic changes, were discarded (see Section S5.1).

Each sampled set defines a *clump variation*, representing the ensemble of admissible transformations around a given reference clump. From each set, we derived a single data instance in the form of a descriptor vector summarizing sulfidation-induced redistribution behaviors. Specifically, we computed the interquartile range (IQR) of voxel intensities for Cu, Cu_2S, and CuS across all valid samples in the set. Subsequently, we calculated the variance of these IQR values as a descriptor of sulfidation-induced redistribution extent. Thus, each clump variation was represented by a 3D vector [$Var(IQR_{Cu})$, $Var(IQR_{Cu_2S})$, $Var(IQR_{CuS})$], capturing compositional responses within its local transformation space (see Section S5.2). We applied Multidimensional scaling (MDS) to these descriptors, generating a 2D embedding space that visualizes the similarity among clump variations based on their transformation behaviors (Fig. 8a). Details of the embedding space construction are

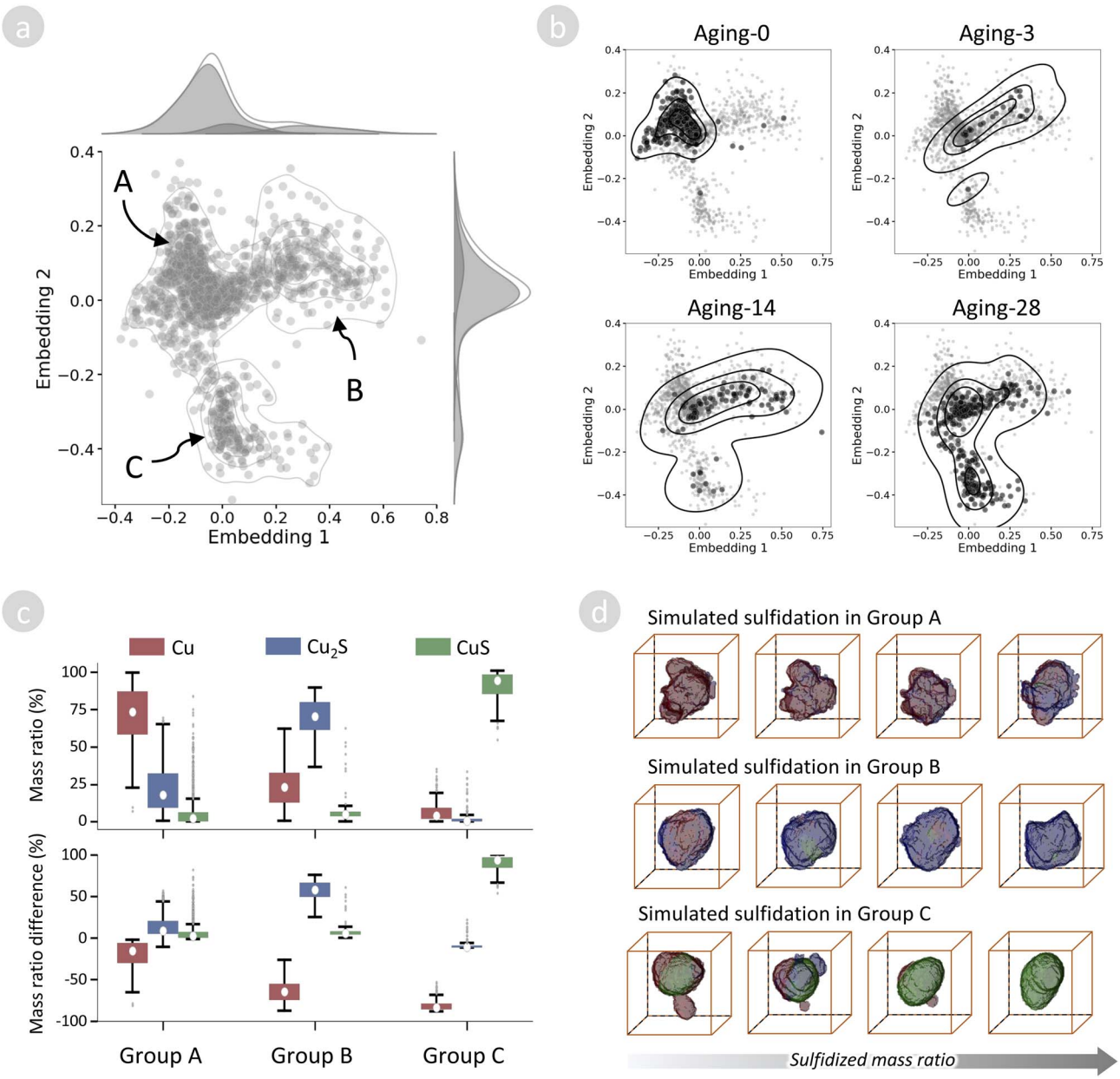

Fig. 8 Analysis of sulfidation dynamics in aging rubber/brass composites. (a) Two-dimensional embedding space obtained *via* multidimensional scaling (MDS) of clump transformation descriptors, revealing three distinct sulfidation modes, denoted as Groups A, B, and C. (b) Temporal mapping showing progression of sulfidation modes across aging stages (0–28 days), with contour lines indicating density distributions. For each stage, darker points denote variations associated with that stage, while lighter ones indicate variations involved with other stages. (c) Box plots summarizing the differences in compositional mass ratios (Cu, Cu_2S, CuS) across the identified sulfidation modes (the gray dots indicate outliers of the respective attributes). (d) Representative 3D visualizations of sulfidation progression for each mode, showing transformation from left to right with increasing sulfidized mass ratio.

shown in Section S5.3. Notably, the embedded clump variations self-organized into three distinct groups—denoted as A, B, and C—each exhibiting characteristic transformation patterns.

To interpret the underlying dynamic behaviors, clump variations were mapped to experimentally temporal observations (Fig. 8b). Clump variations in Group A predominantly corresponded to the initial aging stage (Aging-0). Conversely, clump variations in Group B were primarily observed after 3 to 14 days, while those in Group C appeared mostly in later stages (14 to 28 days). Previous experimental studies indicate that early aging (0–14 days) primarily involves sulfidation and conversion of metallic Cu to monovalent Cu_2S. By contrast, the later aging stage (14–28 days) is characterized by increased formation of bivalent CuS. This temporal alignment suggests that clump variations in Groups A and B are primarily associated with the sulfidation and consumption of Cu to form the monovalent Cu_2S phase, while clump variations in Group C are more closely associated with the formation of the bivalent CuS phase.

Fig. 8c provides further support for the inferred sulfidation phases. Clump variations in Group A exhibited high metallic Cu content, indicating minimal sulfidation, whereas Groups B and C showed pronounced Cu_2S and CuS enrichment, respectively. Further simulation of transitions within these groups revealed that clump variations in Groups A and B exhibited distinct degrees of monovalent sulfidation. Notably, the transformation amplitude in Group A was significantly lower compared with Group B. Conversely, Group C exhibited significant bivalent sulfidation, confirming distinct transformation dynamics at later stages.

The representative simulation of these clump variations, shown in Fig. 8d, revealed distinct patterns: Cu_2S preferentially accumulated near clump surfaces during the early and intermediate stages, whereas CuS formation prominently occurred in interior regions during late stages. These spatial distributions closely match experimental observations, where early sulfidation (Cu_2S formation) dominates interfacial regions during initial vulcanization, whereas deeper CuS formation is associated with progressive mechanical degradation observed in aging or cracked regions.[1,47,48]

4. Discussion

This study introduced a generative framework that infers material dynamics from sparse experimental observations by utilizing the latent space of a trained GAN to perform statistical interpolation of unobserved intermediate states and to classify transformation modes. Unlike prior approaches that interpolate between states or simulate single transitions, our method samples local ensembles—termed material variations—to capture the spectrum of plausible short-range transformations. These ensembles enable statistical analysis of transformation behaviors without requiring time-resolved data. This study's core novelty lies in shifting from deterministic path prediction to distributional analysis of latent-space perturbations. Each variation functions as a surrogate for stochastic dynamics, allowing the extraction of interpretable descriptors—such as anisotropy, tortuosity, and compositional variance—and uncovering hidden transformation modes. Notably, our results validate the central hypothesis that physically plausible transformations generated through latent space exploration demonstrate the generative models' ability to capture essential aspects of the dynamical processes underlying the training images. Across three distinct case studies, these variations revealed mechanistic behaviors consistent with experimental observations, despite the absence of explicit physical constraints or temporal supervision. This outcome highlights the emergent structure of the latent space, where small perturbations yield coherent and physically plausible transformations. Consequently, the framework connects deep generative modeling with physical reasoning, enabling data-driven discovery despite limited observational coverage.

Despite its effectiveness, the framework presents two main limitations. First, like any data-driven model, the proposed framework depends on the coverage and diversity of the training data; rare or under-sampled transformation patterns may not be reliably inferred. Second, the approach models transitions as isolated, memoryless perturbations, following a first-order Markovian assumption. Although this simplification facilitates analysis, it restricts the ability to capture trajectory-dependent behaviors such as hysteresis, irreversibility, or long-range causality. Addressing these limitations motivates future directions toward trajectory-level modeling. Incorporating temporal priors through latent ordinary differential equations (ODEs), time-conditioned generative models, or reinforcement learning-based latent navigation could enable the generation of physically consistent transformation sequences. Such advances would transition the framework from sampling local alternatives to simulating complete dynamic pathways governed by kinetic or thermodynamic constraints.

Nonetheless, even without explicit temporal structure, the current formulation provides insights into short-range transformation mechanisms. It also supports interpretable, probabilistic reasoning under sparse observational conditions, reveals emergent structural variation, and facilitates comparative analysis across heterogeneous datasets. Accordingly, the framework functions both as an independent tool for latent dynamics interpretation and as a foundational step toward generative simulation of material evolution. As machine learning continues to advance materials discovery, our study offers a principled and extensible approach to unveiling hidden dynamics from high-dimensional experimental observations.

Author contributions

D.-A. D.: conceptualization, methodology, software, validation, investigation, writing – original draft, writing – review & editing, visualization. M.-Q. H.: conceptualization, validation, investigation, writing – original draft, writing – review & editing, visualization. T.-S. V.: software, validation, investigation, data curation. S. T.: validation, data curation. N. I.: validation, data curation. Y. T.: validation, investigation, data curation. S. M.: validation, data curation. T. K.: validation, data curation. N. A.: validation, data curation. H. M.: validation, investigation, data curation. M. T.: validation, investigation, data curation. H.-C.

D.: conceptualization, methodology, validation, investigation, writing – original draft, writing – review & editing, visualization, supervision, project administration, funding acquisition.

Conflicts of interest

The authors report there are no competing interests to declare.

Data availability

The datasets supporting the findings of this study are available as follows: (1) Ta test chart translation dataset: available at the Zenodo repository[49] from our previous study. (2) Gold nanoparticle diffusion in polyvinyl alcohol solution dataset: available at the Zenodo repository[49] from our previous study. (3) Copper sulfidation in heterogeneous rubber/brass composites: a representative sample dataset that enables reproduction of the key findings is available at the Zenodo repository.[50] Due to proprietary restrictions from the authors' institution, the complete dataset cannot be made publicly available. However, the full data can be obtained from the corresponding authors upon reasonable request.

Code availability: the Python implementation and custom code used to train the GAN model are publicly available on GitHub,[51] and an archived version is stored in Zenodo[52] to ensure long-term accessibility and reproducibility.

Supplementary information: comprehensive documentation of the test datasets, image preparation, model architecture, training regime, and analysis methods. It also presents additional results and images that support the main text and enable reproducibility. See DOI: **https://doi.org/10.1039/d5dd00277j**.

Acknowledgements

This work was supported by the JSPS KAKENHI under Grants 20K05301, JP19H05815, and JP23H05403; JST-CREST Program (Innovative Measurement and Analysis) under Grant JPMJCR2235.

References

1 H. Matsui, Y. Muramoto, R. Niwa, T. Kakubo, N. Amino, T. Uruga, M.-Q. Ha, D.-T. Dinh, H.-C. Dam and M. Tada, Machine learning-derived reaction statistics for 3d spectroimaging of copper sulfidation in heterogeneous rubber/brass composites, *Commun. Mater.*, 2023, **4**, 88.

2 S. Wang, R. Du, Y. Guo, S. Sun and X. Ke, Structural phase transitions, mechanical and electronic properties of $ZrSe_2$ under high pressures via the first-principles calculations, *Comput. Mater. Sci.*, 2023, **226**, 112214.

3 T. Ishida, F. Cleri, K. Kakushima, M. Mita, T. Sato, M. Miyata, N. Itamura, J. Endo, H. Toshiyoshi, N. Sasaki, *et al.*, Exceptional plasticity of silicon nanobridges, *Nanotechnology*, 2011, **22**, 355704.

4 K. Ishizuka, M. Tomitori, T. Arai and Y. Oshima, Mechanical analysis of gold nanocontacts during stretching using an in-situ transmission electron microscope equipped with a force sensor, *Appl. Phys. Express*, 2020, **13**, 025001.

5 H. Matsui, N. Ishiguro, Y. Tan, N. Maejima, Y. Muramoto, T. Uruga, K. Higashi, D.-N. Nguyen, H.-C. Dam, G. Samjeské, *et al.*, Variation of local structure and reactivity of Pt/C catalyst for accelerated degradation test of polymer electrolyte fuel cell visualized by operando 3D CT-XAFS imaging, *ChemNanoMat*, 2022, **8**, e202200008.

6 Y. Wang, H. Zhao, C. Liu, Y. Ootani, N. Ozawa and M. Kubo, Mechanisms of chemical-reaction-induced tensile deformation of an Fe/Ni/Cr alloy revealed by reactive atomistic simulations, *RSC Adv.*, 2023, **13**, 6630–6636.

7 R. Noyori, Chiral metal complexes as discriminating molecular catalysts, *Science*, 1990, **248**, 1194–1199.

8 A. Mohammed and A. Abdullah, Scanning electron microscopy (SEM): A review, in *Proceedings of the 2018 International Conference on Hydraulics and Pneumatics—HERVEX*, Băile Govora, Romania, 2018, vol. 2018, pp. 7–9.

9 D. B. Williams, C. B. Carter, D. B. Williams and C. B. Carter, *The transmission electron microscope*, Springer, 1996.

10 A. Barty, S. Boutet, M. J. Bogan, S. Hau-Riege, S. Marchesini, K. Sokolowski-Tinten, N. Stojanovic, R. Tobey, H. Ehrke, A. Cavalleri, *et al.*, Ultrafast single-shot diffraction imaging of nanoscale dynamics, *Nat. Photonics*, 2008, **2**, 415–419.

11 J. Kang, S. Takazawa, N. Ishiguro and Y. Takahashi, Single-frame coherent diffraction imaging of extended objects using triangular aperture, *Opt. Express*, 2021, **29**, 1441–1453.

12 S. Takazawa, D.-A. Dao, M. Abe, H. Uematsu, N. Ishiguro, T. Hoshino, H. C. Dam and Y. Takahashi, Coupling X-ray photon correlation spectroscopy and dynamic coherent X-ray diffraction imaging: Particle motion analysis from nano-to-micrometer scale, *Phys. Rev. Res.*, 2023, **5**, L042019.

13 P. Lyngby and K. S. Thygesen, Data-driven discovery of 2D materials by deep generative models, *npj Comput. Mater.*, 2022, **8**, 232.

14 A. Merchant, S. Batzner, S. S. Schoenholz, M. Aykol, G. Cheon and E. D. Cubuk, Scaling deep learning for materials discovery, *Nature*, 2023, **624**, 80–85.

15 T. Lookman, P. V. Balachandran, D. Xue and R. Yuan, Active learning in materials science with emphasis on adaptive sampling using uncertainties for targeted design, *npj Comput. Mater.*, 2019, **5**, 21.

16 S. Otten, S. Caron, W. de Swart, M. van Beekveld, L. Hendriks, C. van Leeuwen, D. Podareanu, R. Ruiz de Austri and R. Verheyen, Event generation and statistical sampling for physics with deep generative models and a density information buffer, *Nat. Commun.*, 2021, **12**, 2985.

17 A. Khan, C.-H. Lee, P. Y. Huang and B. K. Clark, Leveraging generative adversarial networks to create realistic scanning transmission electron microscopy images, *npj Comput. Mater.*, 2023, **9**, 85.

18 X. Lyu and X. Ren, Microstructure reconstruction of 2D/3D random materials via diffusion-based deep generative models, *Sci. Rep.*, 2024, **14**, 5041.

19 B. Murgas, J. Stickel and S. Ghosh, Generative adversarial network (GAN) enabled statistically equivalent virtual

microstructures (SEVM) for modeling cold spray formed bimodal polycrystals, *npj Comput. Mater.*, 2024, **10**, 32.
20 I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville and Y. Bengio, Generative adversarial networks, *Commun. ACM*, 2020, **63**, 139–144.
21 D. P. Kingma and M. Welling, An introduction to variational autoencoders, *Found. Trends Mach. Learn.*, 2019, **12**, 307–392.
22 J. Ho, A. Jain and P. Abbeel, Denoising diffusion probabilistic models, *Adv. Neural Inf. Process. Syst.*, 2020, **33**, 6840–6851.
23 C. X. Hernández, H. K. Wayment-Steele, M. M. Sultan, B. E. Husic and V. S. Pande, Variational encoding of complex dynamics, *Phys. Rev. E*, 2018, **97**, 062412.
24 S. V. Kalinin, O. Dyck, S. Jesse and M. Ziatdinov, Exploring order parameters and dynamic processes in disordered systems via variational autoencoders, *Sci. Adv.*, 2021, **7**, eabd5084.
25 P. Bojanowski, A. Joulin, D. Lopez-Paz and A. Szlam, Optimizing the latent space of generative networks, *arXiv*, 2017, preprint, arXiv:1707.05776, DOI: **10.48550/arXiv.1707.05776**.
26 T. Karras, S. Laine and T. Aila, A style-based generator architecture for generative adversarial networks, in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2019, pp. 4401–4410.
27 N. Metropolis and S. Ulam, The Monte Carlo method, *J. Am. Stat. Assoc.*, 1949, **44**, 335–341.
28 I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville and Y. Bengio, Generative adversarial networks, *Commun. ACM*, 2020, **63**, 139–144.
29 P. Dhariwal and A. Nichol, Diffusion models beat gans on image synthesis, *Adv. Neural Inf. Process. Syst.*, 2021, **34**, 8780–8794.
30 M. Arjovsky, S. Chintala and L. Bottou, Wasserstein generative adversarial networks, in *International conference on machine learning*, PMLR, 2017, pp. 214–223.
31 I. Gulrajani, F. Ahmed, M. Arjovsky, V. Dumoulin and A. C. Courville, Improved training of wasserstein gans, *Adv. Neural Inf. Process. Syst.*, 2017, **30**, 5769–5779.
32 T. Karras, T. Aila, S. Laine and J. Lehtinen, Progressive growing of gans for improved quality, stability, and variation, *arXiv*, 2017, preprint, arXiv:1710.10196, DOI: **10.48550/arXiv.1710.10196**.
33 T.-S. Vu, M.-Q. Ha, A. M. Bachtiar, D.-A. Dao, T. Tran, H. Kino, S. Takazawa, N. Ishiguro, Y. Sasaki, M. Abe, H. Uematsu, N. Okawa, K. Ozaki, K. Kobayashi, Y. Honjo, H. Nishino, Y. Joti, T. Hatsui, Y. Takahashi and H.-C. Dam, PID3Net: a deep learning approach for single-shot coherent X-ray diffraction imaging of dynamic phenomena, *npj Comput. Mater.*, 2025, **11**, 66.
34 R. Brunelli, *Template matching techniques in computer vision: theory and practice*, John Wiley & Sons, 2009.
35 R. K. Pathria, *Statistical mechanics*, Elsevier, 2016.
36 C. C. Aggarwal, A. Hinneburg and D. A. Keim, On the surprising behavior of distance metrics in high dimensional space, in *International conference on database theory*, Springer, 2001, pp. 420–434.
37 K. Beyer, J. Goldstein, R. Ramakrishnan and U. Shaft, When is "nearest neighbor" meaningful?, in *Database Theory—ICDT'99: 7th International Conference Jerusalem, Israel, January 10–12, 1999, Proceedings 7*, Springer, 1999, pp. 217–235.
38 J. Lerner, P. A. Gomez-Garcia, R. L. McCarthy, Z. Liu, M. Lakadamyali and K. S. Zaret, Two-parameter mobility assessments discriminate diverse regulatory factor behaviors in chromatin, *Mol. Cell*, 2020, **79**, 677–688.
39 J. E. Maris, F. T. Rabouw, B. M. Weckhuysen and F. Meirer, Classification-based motion analysis of single-molecule trajectories using DiffusionLab, *Sci. Rep.*, 2022, **12**, 9595.
40 D. Müllner, Modern hierarchical, agglomerative clustering algorithms, *arXiv*, 2011, preprint, arXiv:1109.2378, DOI: **10.48550/arXiv.1109.2378**.
41 K. P. Murphy, *Probabilistic machine learning: an introduction*, MIT Press, 2022.
42 S. Kang, J.-H. Kim, M. Lee, J. W. Yu, J. Kim, D. Kang, H. Baek, Y. Bae, B. H. Kim, S. Kang, *et al.*, Real-space imaging of nanoparticle transport and interaction dynamics by graphene liquid cell tem, *Sci. Adv.*, 2021, **7**, eabi5419.
43 E. A. Codling, M. J. Plank and S. Benhamou, Random walk models in biology, *J. R. Soc. Interface*, 2008, **5**, 813–834.
44 M. J. Saxton, Anomalous diffusion due to obstacles: a Monte Carlo study, *Biophys. J.*, 1994, **66**, 394–401.
45 F. Höfling and T. Franosch, Anomalous transport in the crowded world of biological cells, *Rep. Prog. Phys.*, 2013, **76**, 046602.
46 T. Uruga, M. Tada, O. Sekizawa, Y. Takagi, T. Yokoyama and Y. Iwasawa, SPring-8 BL36XU: Synchrotron radiation X-ray-based multi-analytical beamline for polymer electrolyte fuel cells under operating conditions, *Chem. Rec.*, 2019, **19**, 1444–1456.
47 L. Liu, R. Chen, W. Liu, Y. Zhang, X. Shi and Q. Pan, Fabrication of superhydrophobic copper sulfide film for corrosion protection of copper, *Surf. Coat. Technol.*, 2015, **272**, 221–228.
48 K. Ozawa and K. Mase, Evidence for chemical bond formation at rubber–brass interface: Photoelectron spectroscopy study of bonding interaction between copper sulfide and model molecules of natural rubber, *Surf. Sci.*, 2016, **654**, 14–19.
49 T.-S. Vu, H.-C. Dam and Y. Takahashi, *Diffraction intensity of Ta test chart and gold nanoparticles from dynamic single-shot CXDI experiments*, 2024, DOI: **10.5281/zenodo.12144981**.
50 D. A. Dao and H.-C. Dam, *Material dynamics analysis with deep generative model*, 2025, DOI: **10.5281/zenodo.15730987**.
51 D. A. Dao, *Material dynamics analysis with deep generative model*, GitHub, 2025, **https://github.com/ducanh2526/Material-Dynamics-Analysis-with-Deep-Generative-Model**.
52 D. A. Dao, *Material dynamics analysis with deep generative model*, 2025, DOI: **10.5281/zenodo.17156489**.