POLITECNICO DI TORINO
Repository ISTITUZIONALE

A rigorous data-driven approach to defect nucleation in metals linking kinetic properties and (dis)order parameters

Original
A rigorous data-driven approach to defect nucleation in metals linking kinetic properties and (dis)order parameters / Perrone, M., Girardier, D.D., Pavan, G.M., Pietrucci, F.. - In: THE JOURNAL OF CHEMICAL PHYSICS. - ISSN 0021-9606. - 165:12(2026), pp. 1-12. [10.1063/5.0342563]

Availability:
This version is available at: 11583/3015907 since: 2026-09-25T07:45:07Z

Publisher:
American Institute of Physics - AIP

Published
DOI:10.1063/5.0342563

Terms of use:

This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

Publisher copyright

(Article begins on next page)

13 November 2026

The Journal of Chemical Physics

RESEARCH ARTICLE | SEPTEMBER 22 2026

A rigorous data-driven approach to defect nucleation in metals linking kinetic properties and (dis)order parameters

Mattia Perrone ; David D. Girardier ; Giovanni M. Pavan ; Fabio Pietrucci

Check for updates

J. Chem. Phys. 165, 124506 (2026)
https://doi.org/10.1063/5.0342563

View Online
Export Citation

Articles You May Be Interested In

Homogeneous nucleation rate of methane hydrate formation under experimental conditions from seeding simulations

J. Chem. Phys. (March 2023)

Generalized Langevin theory for many-body problems in chemical dynamics: The method of partial clamping and formulation of the solute equations of motion in generalized coordinates

J. Chem. Phys. (September 1984)

Rigorous index theory for one-dimensional interacting topological insulators

J. Math. Phys. (April 2023)

25 September 2026 07:31:09

AIP Publishing
The Journal of Chemical Physics
Special Topics Open for Submissions
Learn More

AIP Publishing

The Journal of Chemical Physics ARTICLE pubs.aip.org/aip/jcp

A rigorous data-driven approach to defect nucleation in metals linking kinetic properties and (dis)order parameters

Cite as: J. Chem. Phys. 165, 124506 (2026); doi: 10.1063/5.0342563
Submitted: 7 May 2026 • Accepted: 27 August 2026 •
Published Online: 22 September 2026

View Online Export Citation CrossMark

Mattia Perrone,[1,a),b)] David D. Girardier,[2] Giovanni M. Pavan,[1,b)] and Fabio Pietrucci[2]

AFFILIATIONS

[1] Department of Applied Science and Technology, Politecnico di Torino, Corso Duca degli Abruzzi 24, 10129 Torino, Italy
[2] Institut de Minéralogie, de Physique des Matériaux et de Cosmochimie, UMR 7590 CNRS, Sorbonne Université, Muséum National d'Histoire Naturelle, 75005 Paris, France

[a)] **Present address**: Institut de Minéralogie, de Physique des Matériaux et de Cosmochimie, UMR 7590 CNRS, Sorbonne Université, Muséum National d'Histoire Naturelle, 75005 Paris, France. **Email**: mattia.perrone@sorbonne-universite.fr
[b)] **Authors to whom correspondence should be addressed**: mattia.perrone@polito.it and giovanni.pavan@polito.it

ABSTRACT

Nucleation processes, through which a new structure progressively forms within a pre-existing homogeneous phase, are fundamental in materials science but are also typically non-trivial to elucidate. Cases in which defects (or disorder) nucleate within an initially ordered structure are no exception. A prominent example is the nucleation of dislocations in metals, which critically govern their mechanical, electronic, thermal, and chemical properties. While atomic-level insights can be attained using, e.g., molecular dynamics simulations, systematically characterizing nucleation mechanisms and accurately quantifying kinetic rates remain challenging tasks. In this work, we demonstrate how the choice of the order parameter used to track the transition has a very strong effect on the accuracy of the kinetic rate predicted from the corresponding free-energy barrier and diffusion coefficient, a fact that has often been overlooked in the past. By exploiting this systematic error to our advantage, we demonstrate that it is possible to rigorously characterize the nucleation process using a data-driven scheme based on a variational principle, leading to optimal order parameters and a faithful mechanistic description. We apply this method to characterize, as a representative case study, the nucleation of dislocations in crystalline *fcc* copper by analyzing replica molecular dynamics simulations at the elastic–plastic limit. By means of committor analysis and Langevin modeling, our approach allows us to systematically rank candidate (dis)order parameters, identify the critical nuclei (transition states), and infer the free-energy landscapes. Given its general foundations, this method can be extended to nucleation phenomena in a broad class of materials.

© 2026 Author(s). All article content, except where otherwise noted, is licensed under a Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/). https://doi.org/10.1063/5.0342563

25 September 2026 07:31:09

I. INTRODUCTION

Dislocations are a fundamental family of defects in metals, semiconductors, and a variety of other materials.[1] They correspond to localized regions where atoms deviate from their ideal lattice positions, giving rise to line defects around which atomic planes exhibit discontinuities.[2] Due to the lattice distortions that they induce, dislocations strongly influence many material properties—most notably mechanical characteristics such as, e.g., strength, ductility, and plasticity, but also electronic, thermal, and chemical behavior.[4,73] Understanding dislocation dynamics thus carries significant theoretical and practical importance across materials science and engineering.[4,5]

The emergence of dislocations in response to applied stress is a nucleation phenomenon.[3] Such nucleation can occur homogeneously within a defect-free crystal (as observed, e.g., under nanoindentation or in nanostructured materials[6]) or heterogeneously at structural discontinuities such as surfaces or grain boundaries. Phenomenologically, atoms locally displace from their equilibrium positions, gradually aggregating into a critical nucleus of distorted

© Author(s) 2026

atomic arrangements. This process is orchestrated by a sequence of thermally induced fluctuations that accumulate within the material up to a critical threshold.[7] At this critical configuration, the nucleus has an equal probability (50%) of either reverting to the defect-free state or evolving irreversibly into a stable dislocation. This constitutes the transition state, crossing of which triggers the motion of an entire plane (dislocation propagation).

Such a nucleation process has been primarily studied as a function of four key factors, such as chemical composition (single-element metals and alloys), structural characteristics (bulk material, interfaces, and pre-existing defects), applied stress, and temperature.[6,8,9] In general, atomistic simulations are a particularly valuable tool for studying this phenomenon,[10] as they allow complete control over the aforementioned factors.[6,11] In this context, transition state theory (TST)-based methods have been employed to assess the kinetics and thermodynamics of the nucleation phenomena,[12–16] along with studies aimed at investigating the enthalpic and entropic effects involved in them.[17–19]

Despite being a fundamental problem, the literature lacks studies that address dislocation nucleation in its detailed atomistic mechanism and that identify in a compelling way the transition states and the optimal order parameters to characterize it. In fact, in nucleation, as in thermally activated phenomena in general, the systematic study and testing of order parameters is crucial because they provide both mechanistic insight and a low-dimensional representation of the transition process that in turn affects the estimation of free-energy barriers and kinetic rates.[20–22] The development, comparison, and optimization of candidate order parameters—using advanced techniques based on rigorous statistical mechanics—have thus become a central focus in modern computational physical-chemistry/chemical-physics,[23–26] with significant progress in the last decades, particularly in the study of crystallization processes.[27–32] Nevertheless, despite the interest and considerable efforts in many fields, to the best of our knowledge, relatively little attention has been devoted to the study of dislocation nucleation.

As a notable example, an order parameter to monitor the progress of dislocations was introduced in Ref. 33 in a study on Ni_3Al under dynamic and static loading conditions. Such a parameter was constructed by tracking atoms with relative displacements exceeding a critical threshold compared to their neighbors, allowing the identification of hot spots, regions of localized atomic displacements that act as precursors to homogeneous dislocation nucleation. This is in line with what has been done recently in similar or other contexts with physics-based or abstract descriptors.[34–39] Recently, it has also been shown how all such effects can be related to important fluctuations and how they appear in space and time.[7,40]

The hot spots identified in Ref. 33, characterized by interatomic displacements approaching half the Shockley partial Burgers vector, were shown to evolve into small dislocation loops that subsequently expanded to form slip planes. Subsequently, Ryu *et al.*[18] extended this concept to study dislocation nucleation in copper, examining both homogeneous and heterogeneous nucleation under compression across a wide range of applied stresses and temperatures. In that work, the nucleation kinetic rates were computed using Becker–Döring theory,[41] and the free-energy barriers were determined using umbrella sampling techniques and compared with zero-temperature estimates: the obtained results highlighted the importance of considering entropy contributions, which were shown to enhance nucleation rates by orders of magnitude compared to predictions based solely on, e.g., the activation potential energy.

Building on this scenario, in this work, we aim to provide a more advanced description and deeper insight into the degrees of freedom responsible for the homogeneous nucleation of dislocation defects, identifying the transition states (using committor analysis), reconstructing equilibrium free-energy barriers, and estimating the

25 September 2026 07:31:09

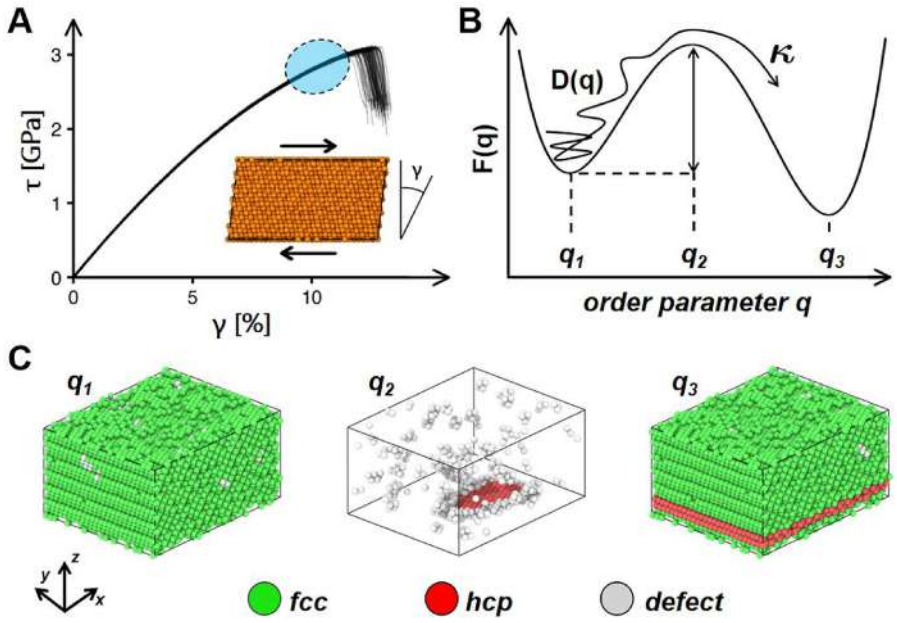

FIG. 1. (a) Stress–strain curve of the Cu crystalline supercell modeled here under simple shear deformation (inset). The light blue shaded area indicates the pre-yield conditions used as the starting point for the equilibrium MD simulations. Close to such an elastic/plastic limit, dislocation nucleation is a rare event. (b) Schematic representation of the free-energy landscape as a function of the order parameter q. The plot illustrates the metastable basins, along with the free-energy barrier ΔF, the diffusion profile $D(q)$, and the kinetic rate k. (c) Snapshots of atomic configurations at three key stages: t_1 corresponds to the initial metastable state, t_2 shows the nucleation of a dislocation, and t_3 depicts the system after the planes have slipped. Atoms are colored by their local structure: green for *fcc*, red for *hcp*, and gray for defect configurations.

© Author(s) 2026

associated kinetic rates based solely on unbiased MD trajectories. We exploit rigorous tools from statistical mechanics and the theory of stochastic processes, specifically maximum-likelihood inference methods applied to Langevin models, complemented by robust diagnostics to validate the optimal models that we identify.[42] We perform equilibrium MD simulations at constant volume under pre-yield conditions [highlighted by the blue region in Fig. 1(a)], ensuring that dislocation nucleation occurs as a thermally activated rare event upon crossing a free-energy barrier [Fig. 1(b)]. We systematically compare different candidate order parameters (also called collective variables) to describe the transformation, critically evaluating their performance in characterizing the process. In fact, an often-neglected, even though important, and easy-to-demonstrate feature of free-energy landscapes is the sizable dependence of the barrier height on the choice of the order parameter:[28,43] as we demonstrate in this work, it is crucial to consider this dependence in the estimation of kinetic rates for the nucleation of dislocations in copper.

As we will demonstrate, this approach provides rich, self-consistent insights into the system: as the quality of the order parameter improves, (i) the transition states are better resolved from the initial local minimum (the order parameter approaches the ideal one, i.e., the committor[43]), (ii) the free-energy landscape aligns closely with the exact profile for the optimal coordinate, and (iii) the kinetic rate systematically decreases until it converges to the exact value (consistent with the variational principle introduced in Ref. 44). We find that dislocations in copper nucleate via a non-trivial mechanism promoted by non-affine displacements within the largest defect cluster of atoms.

The methodology presented here can be applied to scenarios with high free-energy barriers, where brute-force simulations become impractical, by utilizing MD trajectories obtained via transition path sampling.[45] Our results pave the way to automatic identification of optimal order parameters based on machine-learning techniques[46] and to the estimation of accurate nucleation rates for a broad range of different systems, materials, and conditions.

II. METHODS

A. Shearing MD simulations

We started by considering a single *fcc* unit cell of Cu and adjusting its orientation to align the crystallographic directions $[11\bar{2}]$, [111], and $[1\bar{1}0]$ with the x, y, and z Cartesian axes, respectively. Subsequently, we replicated the unit cell 16, 6, and 26 times along each Cartesian direction, obtaining an orthorhombic supercell with 14 976 atoms. The interatomic interactions were modeled using the embedded atom method (EAM),[47] which is reliable in reproducing the bulk mechanical properties of such model systems.[7,48,49] The system was energy minimized and equilibrated in the NpT ensemble for 1 ns under conditions of 300 K and 1 bar. All simulations of this study were conducted using LAMMPS[50] version 23 Jun 2022 with a MD time step of 1 fs. We performed a series of 200 independent replicas with initial velocities randomly drawn from the Maxwell–Boltzmann distribution at 300 K. The Nosé–Hoover schemes[51,52] were used to control temperature and pressure, with coupling constants set to 1 and 2 ps, respectively. An anisotropic barostat was applied, allowing independent fluctuations of the box dimensions along x, y, and z to meet the target ambient pressure of 1 bar, while keeping the simulation box orthorhombic.

Subsequently, we initiated the non-equilibrium simulation phase, during which the simulation box was steadily deformed during the MD run. A simple shear deformation was applied by progressively tilting the angle γ_{xy} between the sides oriented along the x and y axes. The deformation was applied at a constant shear rate of 5×10^9 s^{-1}. Throughout this process, the temperature was controlled at 300 K using the Nosè–Hoover thermostat with the SLLOD integration scheme.[53]

Through this procedure, we are able to generate metastable equilibrium structures of the material at increasing shear deformations γ. These configurations are then used as the starting point for the subsequent analyses, in which the material is simulated under equilibrium conditions at fixed deformation (i.e., at fixed γ).

B. MD simulations with a fixed, sheared cell at pre-nucleation conditions

Subsequently, we conducted simulations in the NVT ensemble at pre-yeld conditions by directly selecting configurations from the out-of-equilibrium shearing simulation described in Sec. II A. In particular, we picked the starting point before dislocation nucleation [marked by a light blue shaded area in Fig. 1(a)]. Under these conditions, the atoms are initially in the ideal lattice positions of the sheared cell. We simulate the system in this metastable state A (local equilibrium) until it spontaneously evolves after some waiting time into a state B, in which one crystalline plane has slipped over another due to the nucleation and propagation of a dislocation. To ensure a reasonable spontaneous rate within the simulations, balancing between sufficient sampling time and proximity to the yield point, we selected two configurations corresponding to imposed shear γ = 10.525% and 10.550%.

C. Putative order parameters to describe and monitor the nucleation process

To monitor the system's evolution over time during the spontaneous nucleation of a dislocation, we project the high-dimensional atomic trajectories in Cartesian space over a one dimensional order parameter (or collective variable), defined as an explicit function of atomic coordinates. In general, an order parameter must be able to discriminate between the system's metastable states, but there is large freedom in the definition, with no standard way to choose an optimal order parameter for a given process, so we considered several candidates. Here below, structural identification of the atoms was performed using the DXA method,[54] which relies on common neighbor analysis.[55] This analysis classifies each atom into one of three possible categories: belonging to a crystalline *fcc* structure, an *hcp* structure, or a non-crystalline (defect) configuration. The set of order parameters considered in this work is as follows:

Potential energy of the system—U: corresponds to the EAM force field used in our simulations.

Number of *fcc* atoms—N_{fcc}: represents the total number of atoms featuring a *fcc* local environment and is used as a global measure of the system's crystallinity.

25 September 2026 07:31:09

© Author(s) 2026

Global structural asymmetry—ΣCSP: is a measure of the total disorder in the system, obtained by summing the centrosymmetry parameter[56] (CSP) for each atom, considering its 12 nearest neighbors. A high ΣCSP value indicates a globally disordered configuration.[7]

Non-*fcc* cluster size—$N_{cluster}$: represents the number of atoms in the largest cluster comprised of defect and/or *hcp* atoms.

Radius of the non-*fcc* cluster—$R_{gyr,\,cluster}$: represents the gyration radius of the largest cluster composed of defect and/or *hcp* atoms.

Non-affine deformation—$D^2_{min,cluster}$: measures the sum of squared non-affine displacements for atoms within the largest cluster composed of defect and/or *hcp* atoms. For each atom, the local environment is compared with that of the undeformed reference configuration. The quantity D^2_{min}, introduced in Ref. 57, quantifies the extent to which a local atomic neighborhood deviates from an ideal affine (elastic) transformation. Low values correspond to nearly affine (elastic) deformation, whereas high values indicate significant non-affine, plastic rearrangements. Additional computational details are provided in the supplementary material.

Slipped cluster size[18,33]**—N_{slip}**: an atom is labeled as "slipped" if its distance from any of its original nearest neighbors in the reference configuration has changed by more than 0.33 Å. The slipped atoms are grouped into clusters, where two atoms belong to the same cluster if their distance is less than 3.4 Å (the nearest-neighbor distance in perfect *fcc* copper is 2.6 Å). The number of slipped atoms is defined as the number of atoms in the largest cluster divided by two.[18,33] The reference configuration used to calculate the distances corresponds to the undeformed structure, in which the atoms occupy their equilibrium positions at 0 K.

D. Inference of stochastic models

The dynamics of atoms in high-dimensional phase space can be modeled when observed via a low-dimensional space of generalized coordinates (in our case, a single order parameter q) with different forms of stochastic differential equations.[58,59] In this study, we employed the following overdamped Langevin equation:

$$\dot{q} = -\beta D(q)\frac{dF(q)}{dq} + \frac{dD(q)}{dq} + \sqrt{2D(q)}\,\eta(t), \tag{1}$$

where $D(q)$ is the position-dependent diffusion coefficient, $F(q) = -k_B T \log \rho_{eq}(q)$ is the free-energy profile [with $\rho_{eq}(q)$ being the equilibrium probability density], and $\eta(t)$ is a Gaussian white noise with zero mean and $\langle \eta(t_1)\eta(t_2)\rangle = \delta(t_1 - t_2)$.

For a suitably chosen time resolution τ, the projected MD trajectory $q(t)$ can be modeled by the above equation: the theoretical probability of a displacement from q to q' in a time interval τ can be approximated up to order τ^2 with the following propagator:[42,60]

$$p(q', t+\tau \mid q, t) \approx \frac{1}{\sqrt{2\pi\mu(q)}} e^{-(q'-q-\phi(q))^2/2\mu(q)}, \tag{2}$$

having

$$\begin{aligned} \phi &= a\tau + \frac{1}{2}\left(aa' + Da''\right)\tau^2, \\ \mu &= 2D\tau + \left(aD' + 2a'D + DD''\right)\tau^2, \end{aligned} \tag{3}$$

where the prime indicates the derivative with respect to q, and the term $a = -\beta DF' + D'$ is the drift term that appears in Eq. (1). Due to the Markovian character of the Langevin equation, the likelihood of observing an entire trajectory $\{q_k\}_{k=1,\ldots,M}$ is the product of all the individual probabilities between successive points q_k and q_{k+1},

$$\mathscr{L}[F, D] = \prod_{k=1}^{M-1} p(q_{k+1}, t_{k+1} \mid q_k, t_k) \tag{4}$$

and it can be reformulated as

$$-\log \mathscr{L}[F, D] = \sum_{k=1}^{M-1} \left\{ \frac{1}{2}\log(2\pi\mu_k) + \frac{(q_{k+1} - q_k - \phi_k)^2}{2\mu_k} \right\}. \tag{5}$$

The free energy and diffusion functions $F(q)$ and $D(q)$ can be statistically inferred by writing them in a parametric form and by identifying the optimal parameter values via the minimization of the expression in Eq. (5). We use a Monte Carlo algorithm, involving iterative adjustments to free-energy and diffusion profiles with a suitable Metropolis rule (for more details, we refer to Ref. 42 for a comprehensive description and to the GitHub repository https://github.com/physix-repo/optLE for the freely available implementation).

E. Kinetic rates and ranking of order parameters

In this work, the kinetic rate of dislocation nucleation is estimated in two ways: with brute-force MD trajectories and with stochastic models inferred from MD trajectories projected over different order parameters.

In the first approach, equilibrium MD simulations are initialized in the metastable state A under pre-nucleation conditions, and the system's evolution is monitored using an order parameter. A transition is detected when the order parameter hits a suitably defined threshold, and the time of this event is recorded. The average of these times is referred to as the mean first passage time (MFPT), and the rate κ of the process is then defined as the inverse of the MFPT.

In the second approach, the profiles $F(q)$ and $D(q)$ that have been inferred for each order parameter q with the procedure illustrated in Sec. II D are used to estimate the rate using the following analytical solution of the MFPT problem:[58]

$$\kappa^{-1} = \int_{q_0}^{b} dq \frac{e^{\beta F(q)}}{D(q)} \int_{a}^{q} ds e^{-\beta F(s)}, \tag{6}$$

where q_0 corresponds to the position of the minimum of the initial metastable basin, and a and b represent the positions of the reflecting and absorbing boundaries, respectively. The set of order parameters is then ranked based on the rate predicted by the stochastic model: according to the variational principle demonstrated in Ref. 44, the model's rate decreases by improving the quality of the order-parameter definition until reaching the true (brute-force) kinetic rate for the optimal order parameter that corresponds to the committor probability function. See Ref. 43 for technical details and for an implementation of the variational principle into an algorithm for automatic optimization.

25 September 2026 07:31:09

© Author(s) 2026

III. RESULTS AND DISCUSSION

A. Spontaneous nucleation of dislocations at constant applied strain

We begin our analysis by simulating a defect-free copper (Cu) supercell (modeling the bulk of an *fcc* lattice) that is close to the elastic/plastic limit, in which the material is held at a constant shear deformation γ [Fig. 1(a); cf. Sec. II for a detailed description of the system preparation and simulation protocol]. Under such conditions, the system resides in a metastable state where the material retains its *fcc* structure, yet dislocations are prone to nucleate on relatively short timescales. Phenomenologically, atoms perform a large number of thermal fluctuations around their crystalline lattice sites until, eventually, the system undergoes a spontaneous structural transformation to another local free-energy minimum. Such transformation consists of the slipping of one crystalline plane over another, driven by the nucleation and growth of a dislocation, ultimately resulting in two planes of atoms adopting an *hcp* configuration [Figs. 1(b) and 1(c)].

By repeating this computational experiment multiple times (a total of 200 replica MD simulations were run, with the initial atomic velocities randomly drawn from the Maxwell–Boltzmann distribution), the transition occurs as a stochastic process, i.e., at a random time that is found to be exponentially distributed (Poisson statistics). This behavior conforms to the standard phenomenology of thermally activated barrier-crossing rare events (Fig. 1).

In this framework, it is customary to monitor the system's evolution using low-dimensional representations based on collective variables (or descriptors), which are functions of atomic coordinates. When a collective variable can effectively discriminate between the initial (A) and final (B) states, it is customarily termed an order parameter. A summary of the order parameters employed in this study is provided in Sec. II.

Figure 2(a) provides an example of the time evolution of the "global structural asymmetry" ΣCSP order parameter (i.e., the sum at each time step of the CSP_i local centrosymmetry parameters of each atom i in the system), highlighting the two metastable states and the transition. By analyzing all trajectories, the rate of the process is estimated as the inverse of the mean first passage time, i.e., the inverse of the average transition times recorded across all realizations [Fig. 2(b)]. Such a rate, based on transitions that occur spontaneously, is referred to as the "brute-force" rate.

It is worth noting that any reasonable order parameter can be chosen to perform this analysis, while the resulting brute-force rate estimation does not depend on the particular chosen parameter (see Fig. S2 of the supplementary material). The reason is that the transition time has two distinct contributions: the first, which dominates the total, corresponds to the system undergoing many thermal fluctuations in the initial local free-energy minimum (e.g., this phase is on the nanosecond timescale for our system). The second, much faster, involves the system crossing the free-energy barrier connecting the two minima, a process that lasts ~5 ps in our case study.

From the available trajectories that spontaneously nucleate a dislocation, the free-energy landscape as a function of the putative order parameters q can be easily estimated using the relation $F(q) = -k_B T \log \rho_{eq}(q)$, where $\rho_{eq}(q)$ is the equilibrium probability density. It is important to emphasize, however, that such a landscape can be reliably estimated only up to the transition state: until that point, the system is in local equilibrium, allowing us to apply equilibrium statistical mechanics. To this end, it is therefore necessary to localize the transition state region precisely. This can be done, for example, by means of committor analysis.

From this point onward, the results refer to simulations performed at the applied shear deformation $\gamma = 10.55\%$, corresponding to the green curve in Fig. 2(b).

To clearly illustrate how different order parameters characterize the nucleation process, Fig. 3 provides a structured comparison of four representative descriptors. Each column corresponds to a distinct order parameter, showing its temporal evolution during reactive trajectories, the distribution of values within the initial and transition-state regions, the resulting free-energy profile, and the correlation with the committor probability. For visual clarity, the remaining order parameters (N_{fcc}, ΣCSP, and $\mathrm{R}_{\mathrm{gyr,cluster}}$) are shown in the supplementary material (Fig. S3).

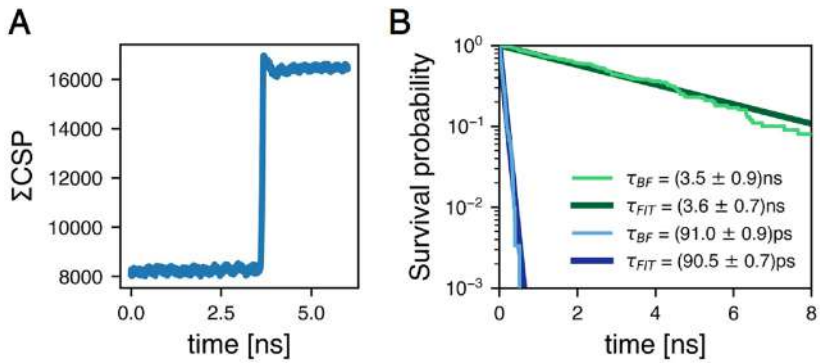

FIG. 2. (a) Example of time evolution of an order parameter (ΣCSP), illustrating the separation between state A (initial crystalline metastable state) and state B (defect crystal post-transition state). (b) Distribution of the transition times under two different applied shear strain conditions ($\gamma = 10.525\%$, green curves, and $\gamma = 10.550\%$, blue curves). The survival probability, shown on a semilogarithmic scale, represents the fraction of systems that have not yet undergone the transition at a given time. The mean transition time is estimated both by direct brute-force sampling (τ_{BF}) and by fitting an exponential distribution (τ_{fit}), with means and standard errors shown in the legend.

25 September 2026 07:31:09

© Author(s) 2026

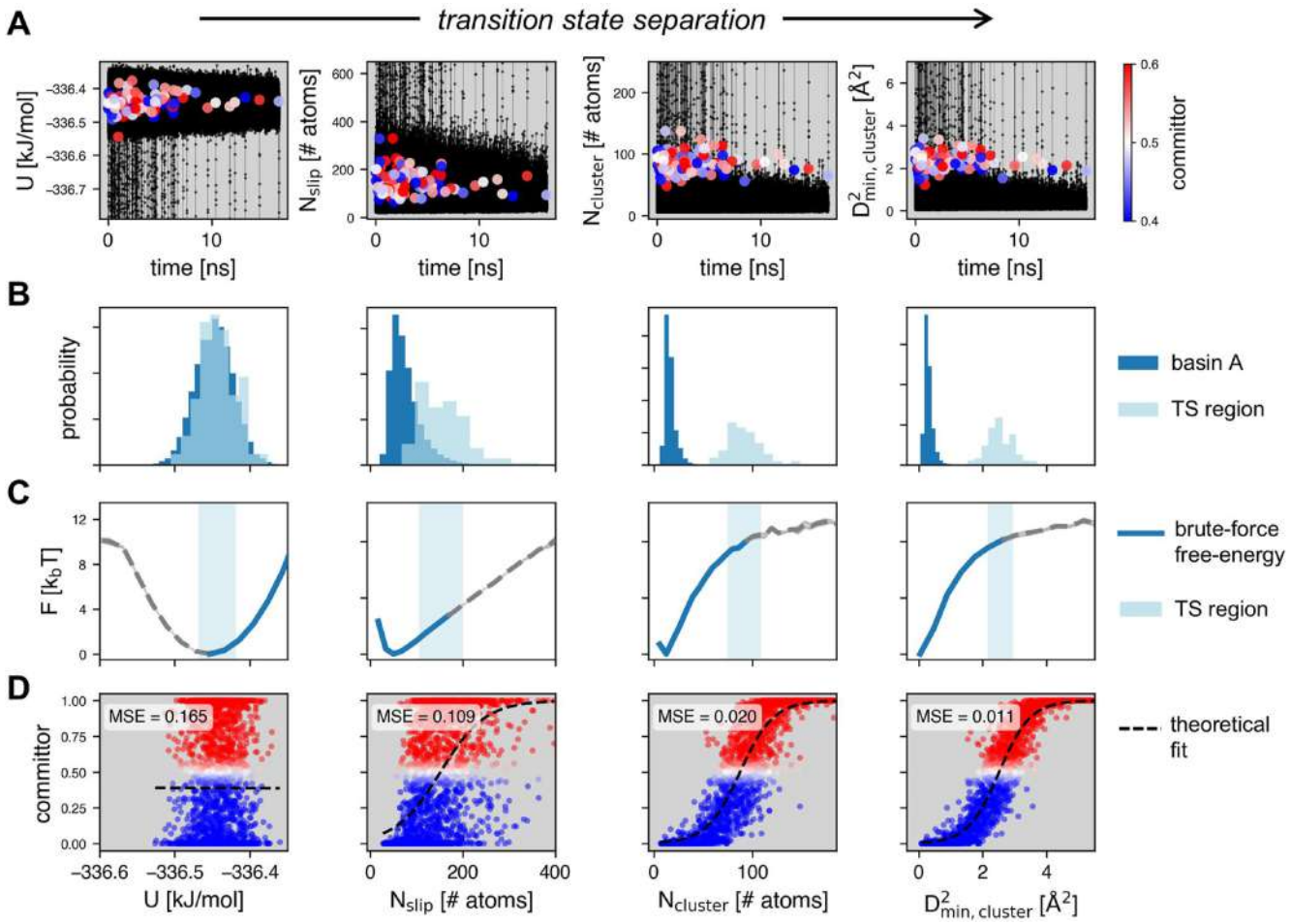

FIG. 3. (a) Reactive trajectories illustrating the evolution of the order parameters over time. Scatter points highlight the transition states, with colors representing the committor probability. (b) Probability histograms for the initial basin (blue) and the transition-state region (light blue) with committor probability between 0.4 and 0.6. (c) Free-energy barrier estimates as a function of the order parameters. The solid blue curve represents the equilibrium free-energy profile, while the dashed gray curve extends the profile in the out-of-equilibrium region beyond the transition. The blue shaded region highlights the transition-state region. (d) Committor probability as a function of the order parameters. The dashed line shows an analytical fit obtained using a logistic function, with the corresponding mean squared error.

Figure 3 demonstrates that, even though the different order parameters are all capable of distinguishing the perfect crystal from the system including a full-grown dislocation, they display significant differences in their ability to resolve the transition states from the initial local minimum (note that, by definition, the committor probability as a function of the atomic configuration would be an optimal order parameter from this viewpoint[28]).

Given that the free-energy landscape changes, in principle, for different order parameters (see the discussion in Ref. 43, with the accompanying change of the diffusion landscape), a poor resolution leads, in practice, to an important shortcoming: i.e., a free-energy barrier that is incompatible (too low) with the observed slow kinetics of the nucleation process. This phenomenology is essentially due to an overlap of the probability distributions of reactants and products, which inevitably leads to an apparent increase of the probability of the transition region.

Therefore, inspecting the barrier height offers a first, straightforward criterion to attempt a ranking of the quality of different order parameters. Clearly, according to this criterion, the barrier of ~10 k_BT in the free-energy landscape of the $N_{cluster}$ and $D^2_{min,cluster}$ order parameters indicates that the latter are likely higher quality descriptors compared to, e.g., the potential energy or the N_{slip} order parameters, which display much lower barriers. Overall, it is possible to observe a clear trend: a greater separation of the transition state from the initial basin *A* leads to a higher estimated free-energy barrier and, at the same time, to a better correlation with the committor [Fig. 3(d)].

It is important to observe how the order parameters N_{slip} and $N_{cluster}$ are conceptually similar, as they both rely on identifying atoms in a non-*fcc* environment and then determining the largest aggregate of such atoms. Yet, it is evident that even a slightly different definition of "non-*fcc* atoms" results in significantly different

25 September 2026 07:31:09

© Author(s) 2026

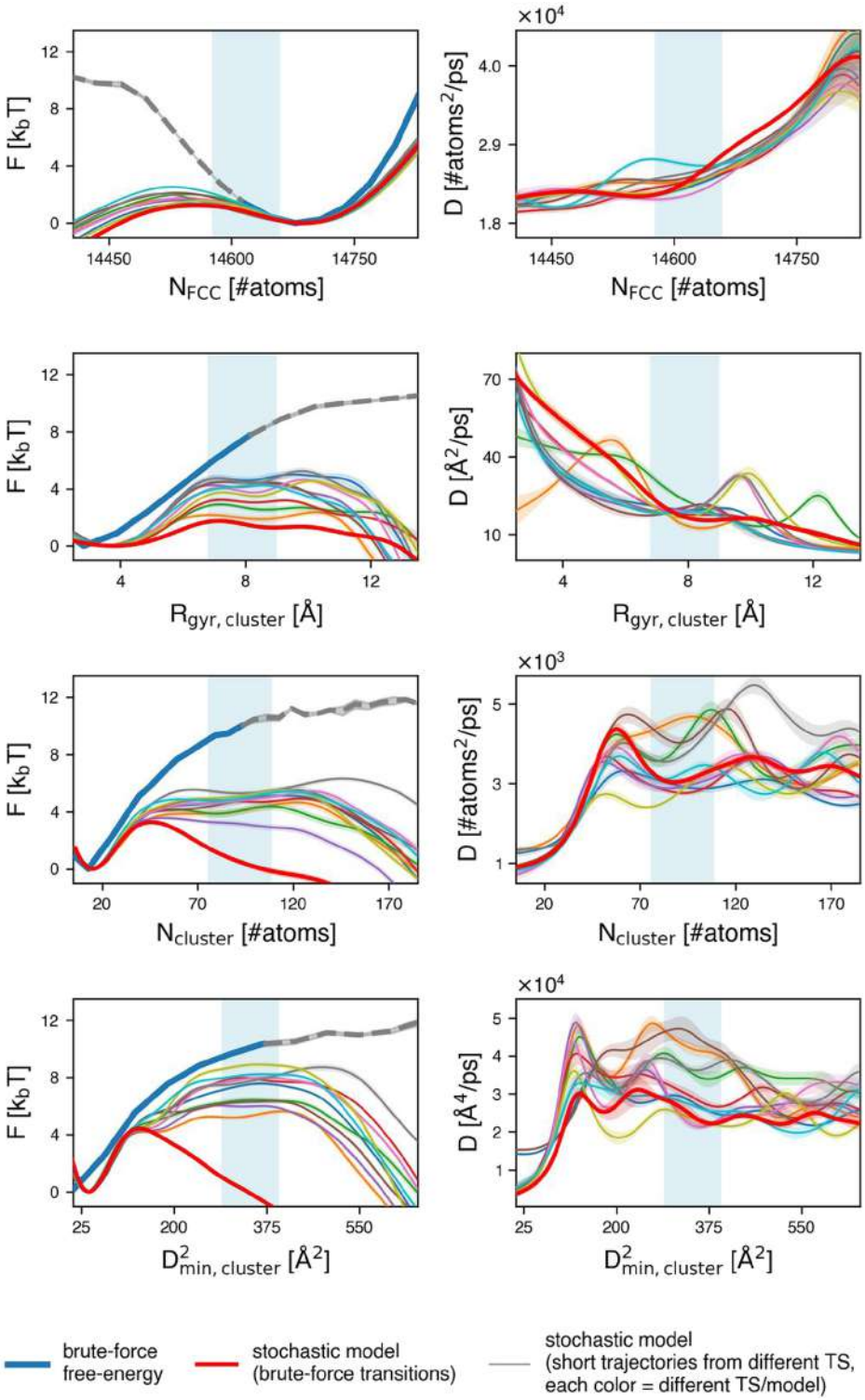

FIG. 4. Estimated profiles of free-energy (F) and diffusion-coefficient (D) obtained using a Langevin model for various order parameters. The thick red line represents the inferred profiles from spontaneous brute-force transitions, while the thin colored lines correspond to models constructed from trajectories initialized at ten randomly selected transition states. The solid blue line shows the brute-force free-energy estimate, for comparison. The shaded light blue region highlights the transition state region.

25 September 2026 07:31:09

© Author(s) 2026

qualities of these descriptors. This demonstrates that, despite the relatively simple visual interpretation of the nucleation process, human intuition is not reliable for identifying the best suited descriptors, and it is crucial to conduct a systematic assessment of order parameters, leading toward advanced optimization methods to refine their definition.[43,61–63]

In addition, we emphasize that while this study is based on order parameters defined on the instantaneous atomic coordinates, the system's evolution can also be studied using descriptors that incorporate the temporal evolution of atomic environments between consecutive time steps (namely, descriptors that are functions of Δt). Among these is LENS,[34] a descriptor of the shuffling of neighbors in the local environment surrounding each atom in the system, which has been recently employed, e.g., to successfully track dislocation dynamics over longer timescales.[7,40] Figures S1 and S2 of the supplementary material illustrate how LENS could potentially be used as an order parameter [despite being a time-dependent descriptor, $f(\Delta t)$].

B. Ranking order parameters via stochastic models

Another powerful approach that allows extracting relevant properties of the nucleation process (including, e.g., the free-energy barrier and the kinetic rate) and ranking the order parameters consists in parameterizing stochastic models starting from data extracted from the MD simulations. For each order parameter q, it is possible to choose a time resolution τ, such that the discretized one-dimensional trajectory $\{q_k \equiv q(k\tau)\}_{k=1,\ldots,n}$ (obtained by projecting the high-dimensional MD trajectory) is well approximated by a Markovian model in the form of an overdamped Langevin equation: to this aim, τ must be longer than the decay time of the auto-correlation function of the order-parameter velocity, so that $\langle \dot{q}_k \dot{q}_{k+1} \rangle \approx 0$. In other words, $\dot{q}(t)$ relaxes to the equilibrium Maxwell–Boltzmann distribution within τ. Likelihood maximization is then applied to the observed trajectory in order to parameterize free-energy and diffusion landscapes, respectively, $F(q)$ and $D(q)$, producing an optimal Langevin model of the nucleation dynamics (see Sec. II as well as Refs. 42 and 59 for more details).

Figure 4 compares the brute-force free-energy landscapes obtained from the set of spontaneous reactive trajectories (blue line, i.e., the region of local equilibrium, up to the mean transition-state point) with the landscapes of the Langevin models parameterized on the same trajectories (thick red line) as a function of different order parameters q. The two free-energy profiles differ because the blue curve represents the equilibrium distribution sampled from MD, whereas the red one is reconstructed from a Langevin model fitted to reproduce the observed dynamics $q(t)$. We underline that, as expected, the nucleation barrier depends on the choice of q: since $F(q)$ has a one-to-one correspondence with the marginal equilibrium probability $\rho_{eq}(q)$, a simple geometric argument proves that a good order parameter, effectively separating the metastable states and the transition states in different q−regions, has a higher barrier than a bad order parameter, which overlaps the three states in the same region.[43] Note that we decided not to infer stochastic models for Σ*CSP* and *U*, because their velocity decorrelation time is comparable to the characteristic timescale for the system to cross the barrier (Fig. S2 of the supplementary material), leading to a scarce number of samples on the transition path.[59]

After inferring $F(q)$ and $D(q)$, the nucleation rate can be straightforwardly obtained for every model, using Eq. (6). Figure 5 presents the brute-force rate (horizontal red line, computed in Sec. III A) alongside the rate inferred from the Langevin models (cross marks) for each q. Once again, different choices of the order parameter lead to different results. The variational principle formally demonstrated in Ref. 44 (and first applied to MD simulations in Ref. 43) states that the rate predicted by a stochastic model built on a given q decreases as the quality of q improves, with the minimum rate ideally coinciding with the exact brute-force rate in the case of the committor as the order parameter.

The nucleation rates predicted by Langevin models in Fig. 5 display significant differences and, according to the variational

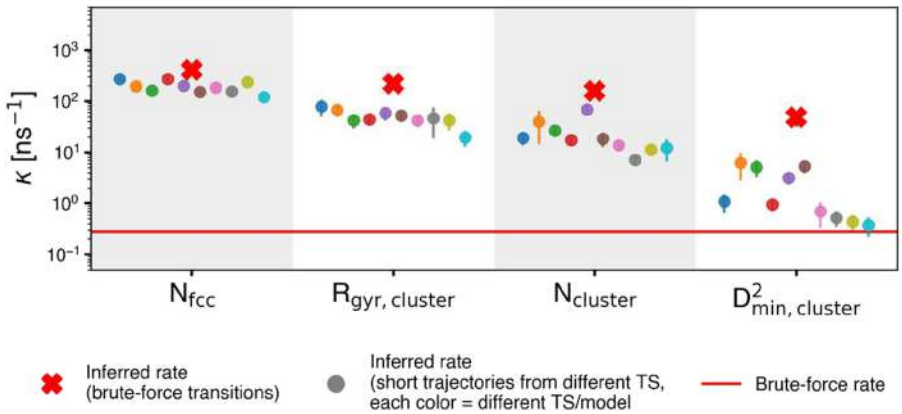

FIG. 5. Comparison of kinetic rates as a function of the order parameters, using different approaches. The solid red line represents the rate directly observed from brute-force simulations initialized in the initial basin A. The red crosses and the colored dots correspond to estimations obtained from stochastic models. In the case of red crosses, the models were built using all spontaneous transitions. In the case of colored dots, the models were inferred from short trajectories initialized from ten different transition-state configurations (TSs), each color referring to a distinct TS-based model. Improved order parameters yield more accurate rate estimations, with the inferred rates systematically converging toward the brute-force value.

25 September 2026 07:31:09

© Author(s) 2026

principle, when ranked in decreasing order, indicate a progressive improvement of the order parameter. It is possible to observe that the best estimate among the predicted rates from models, represented by the crosses in Fig. 5, is provided by the non-affine displacements of atoms in the largest defect cluster ($D^2_{\mathrm{min},cluster}$), which overestimates the true rate by a factor of ~150. In terms of free-energy profiles, $D^2_{\mathrm{min},cluster}$ consistently offers the closest approximation to the brute-force barrier, with a $5k_BT$ underestimation. It is also worth noting the trend that variables that simultaneously better identify the transition state and better approximate the committor provide more accurate barriers and rates. Nevertheless, the discrepancy with direct estimates remains non-negligible.

This discrepancy could have, in principle, two origins. The first one could be the sub-optimal character of the order parameter, according to the variational principle. The second one could be the possibility that the set of MD trajectories employed to train the model includes significantly different transition paths (for instance, because of different atomic structures of the critical nuclei), and that such pathway heterogeneity renders inappropriate the use of a single stochastic model based on a unique variable to faithfully describe the process (like a "blanket pulled in many directions"). Such a hypothesis is supported by the sizable dispersion of the transition states, shown in Fig. 3(a), over an interval of values for each order parameter.

For this reason, we randomly selected ten atomic configurations from the transition-state ensemble (namely, structures with a committor value of ~0.5, visually rendered in Fig. 6). We sampled the transition process by starting 200 MD simulations starting from each of these configurations, with initial atomic velocities randomly drawn from the Maxwell–Boltzmann distribution. This approach allowed us, for each order parameter q, to estimate 10 different stochastic models with the corresponding $F(q)$ and $D(q)$ landscapes (shown with thin lines in Fig. 4) and kinetic rates (reported in Fig. 5 with dots).

Remarkably, building a set of separate models to account for the heterogeneous transition mechanism systematically improves the kinetic estimates for every order parameter. The best estimate is, again, provided by $D^2_{\mathrm{min},cluster}$, which in this case, on average, overestimates the true rate by less than one order of magnitude (i.e., very accurate, considering that good precision in estimating the kinetics of rare events is ~1 order of magnitude). Furthermore, separating the pathways leads to free-energy profiles with higher barriers, approaching the exact brute-force ones. The spread observed in the results originated from different transition states, which supports our earlier hypothesis: when performing spontaneous (brute-force) nucleation transitions, the material indeed follows pathways that are similar yet structurally distinct, since random thermal fluctuations lead to a specific critical nucleus in each transition.

Such heterogeneity motivates a closer examination of the critical nuclei themselves. As shown in Fig. 6, defect clusters have an elongated morphology aligned with the shear direction. The cluster size falls within the range of 80–120 atoms, with peripheral atoms forming branching structures extending from the main body. The majority of atoms within the cluster are in defect configurations (neither *fcc* nor *hcp*), while *hcp*-stacked atoms, when present, are predominantly located at the cluster's core. Non-affine atomic displacements are not uniformly distributed across the cluster: some atoms exhibit minimal displacement, while others undergo substantial local rearrangements. Crucially, all critical nuclei display localized hot spots of high non-affine displacement, frequently correlated with *hcp* atoms, indicating that the crucial stages of atomic slip are dominated by non-affine contributions. These hot spots serve as starting points for the collective slip of an entire atomic plane. This sequence of events is visualized in the movie of the supplementary material, where the initial non-affine core (highlighted in yellow) progressively expands, causing collective propagation of the dislocation across the system. Interestingly, these results place the phenomenology of deformation in crystalline solids in close parallel with that of amorphous materials, where non-affine dynamics also play a pivotal role in, e.g., governing plasticity in glasses,[64,65] identifying topological defects,[66–68] and influencing the rheological properties of viscoelastic materials.[69,70] Moreover, a direct

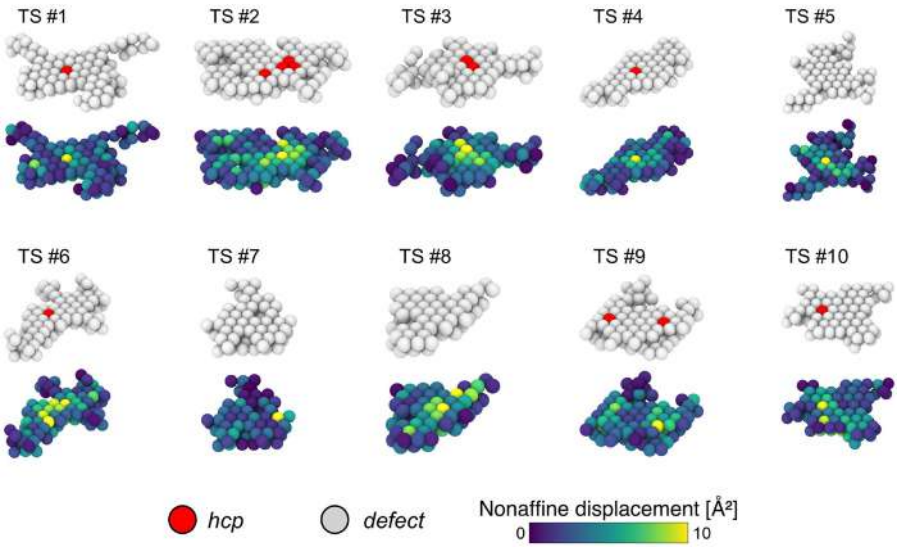

FIG. 6. Three-dimensional representations of the ten randomly selected transition states (TS) used to investigate dislocation nucleation through single transition pathways. For each TS, the upper panels show the largest defect cluster, with atoms colored according to their local structure (defect or *hcp*). The lower panels display the same clusters colored by non-affine displacement (D^2_{min}), highlighting localized irreversible rearrangements. Clusters exhibit hot spots of high non-affine dynamics, which act as initiation sites for the collective slip of atomic planes (see also the movie of the supplementary material).

25 September 2026 07:31:09

© Author(s) 2026

quantitative link exists between the local breaking of inversion symmetry around an atom and the emergence of non-affine displacements in defective crystals, in defect-free non-centrosymmetric crystals (e.g., quartz), and in glasses.[71] Such non-affine fluctuations reduce the local shear modulus (i.e., the local shear rigidity),[72] thereby facilitating dislocation nucleation and plastic rearrangements.

In conclusion, all these observations underline the importance of identifying optimal order parameters to assess the mechanistic details and to reliably infer the thermodynamics and kinetics of the process, as well as the key advantages of the stochastic models framework, providing multi-faceted insights into the transition process.

IV. CONCLUSIONS

In this work, we investigated the microscopic mechanism of spontaneous dislocation nucleation in a perfect copper crystal under constant applied stress. Our results demonstrate that the identification of transition state configurations as well as of high-quality order parameters is paramount to achieve a quantitative understanding of the free-energy barrier and kinetic rate of the process.

Starting from hundreds of all-atom molecular dynamics trajectories and exploiting a combination of rigorous tools based on statistical mechanics, we critically assess the low performance of several widespread and intuitive order parameters while proving the superiority of others. The systematic ranking of different generalized coordinates is achieved in two different ways: based on the distributions of committor probabilities and based on the rates predicted by data-driven stochastic models of the projected dynamics, optimized by likelihood maximization. The two approaches are shown to agree: lower overlap of crystalline configurations with transition states and lower kinetic rate of the optimal stochastic model concur with each other in predicting a higher quality of the order parameter.

We underline how the possibility to precisely identify the transition state, granted by our method, allows us to obtain detailed insight into the transition process, such as determining the structure and size of critical nuclei. In a complementary fashion, optimal order parameters are useful to pinpoint the subtle structural peculiarities of the critical nuclei. We find that dislocations nucleate in a non-trivial way, and tracking their evolution requires monitoring both the nucleus composed of defect atoms and the contributions from non-affine displacements. In particular, our findings highlight that the nucleation process is initiated at localized hot spots of high non-affine displacement that drive the collective slip of entire atomic planes. It is likely that such hot spots of non-affine activity are not merely correlated with the nucleation events but indicate regions of locally softened shear response, where the crystal could first lose mechanical stability.

We remark that the theoretical approach proposed in this work is not restricted to the availability of brute-force simulations of spontaneous nucleation, but it can be applied to, e.g., transition path sampling simulations that are computationally efficient for high-barrier processes,[43,45] in this case leading to the quantitative prediction of kinetic rates solely from short unbiased trajectories.

Clearly, the proposed systematic ranking of order parameters lends itself to future machine learning algorithms for automatic optimization. Furthermore, the approach is not specific to copper or to metallic systems, but it introduces a more rigorous and systematic framework for investigating the formation of defects in generic materials, as well as, e.g., phase nucleation and transitions in a variety of complex molecular systems.

SUPPLEMENTARY MATERIAL

The supplementary material encompasses details on the following aspects: non-affine displacements (definition and tools used for computation); committor analysis (computational protocol and histogram test for the order parameters); consistency checks on the stochastic models (velocity autocorrelations of the order parameters, comparison of stochastic models built with different time resolutions, analysis of the gaussian character and of the temporal correlations of the noise generated by the models).

ACKNOWLEDGMENTS

The authors acknowledge Line Mouaffac and Jérémy Diharce for discussions and help with methodological aspects. F.P. acknowledges Luca Barbisan, Anna Marzegalli, and Francesco Montalenti for discussions and for sharing their expertise. The authors are grateful for the Reviewer's feedback, which highlighted the connection between the emergence of local non-affine displacements and the softening of the mechanical response of materials. G.M.P. acknowledges the funding received from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (Grant Agreement No. 818776—DYNAPOL). The authors acknowledge ISCRA for awarding this project access to the LEONARDO supercomputer, owned by the EuroHPC Joint Undertaking and hosted by CINECA (Italy).

AUTHOR DECLARATIONS

Conflict of Interest

The authors have no conflicts to disclose.

Author Contributions

Mattia Perrone: Conceptualization (equal); Formal analysis (equal); Investigation (equal); Writing – original draft (equal); Writing – review & editing (equal). **David D. Girardier**: Formal analysis (equal); Investigation (equal); Writing – original draft (equal); Writing – review & editing (equal). **Giovanni M. Pavan**: Conceptualization (equal); Formal analysis (equal); Investigation (equal); Writing – original draft (equal); Writing – review & editing (equal). **Fabio Pietrucci**: Conceptualization (lead); Formal analysis (equal); Investigation (equal); Writing – original draft (equal); Writing – review & editing (equal).

DATA AVAILABILITY

Complete data and materials pertaining to the atomistic simulations and data analyses conducted herein (input files, model files, analysis tools, etc.) are available at:

25 September 2026 07:31:09

© Author(s) 2026

https://doi.org/10.5281/zenodo.22278801. Other information needed is available from the corresponding authors upon reasonable request.

REFERENCES

[1]N. W. Ashcroft, N. D. Mermin, and S. Rodriguez, "Solid state physics," Am. J. Phys. **46**, 116–117 (1978).
[2]P. M. Anderson, J. P. Hirth, and J. Lothe, *Theory of Dislocations* (Cambridge University Press, 2017).
[3]D. Hull and D. J. Bacon, *Introduction to Dislocations* (Elsevier, 2011), Vol. 37.
[4]V. Bulatov and W. Cai, *Computer Simulations of Dislocations* (Oxford University Press, 2006).
[5]M. A. Meyers and K. K. Chawla, *Mechanical Behavior of Materials* (Cambridge University Press, 2008).
[6]L. Y. Chen, M.-r. He, J. Shin, G. Richter, and D. S. Gianola, "Measuring surface dislocation nucleation in defect-scarce nanostructures," Nat. Mater. **14**, 707–713 (2015).
[7]M. Perrone, M. Cioni, M. Delle Piane, and G. M. Pavan, "Unsupervised tracking of local and collective defects dynamics in metals under deformation," J. Chem. Phys. **162**, 214507 (2025).
[8]R. E. Miller and D. Rodney, "On the nonlocal nature of dislocation nucleation during nanoindentation," J. Mech. Phys. Solids **56**, 1203–1223 (2008).
[9]A. Singer, M. Zhang, S. Hy, D. Cela, C. Fang, T. A. Wynn, B. Qiu, Y. Xia, Z. Liu, A. Ulvestad *et al.*, "Nucleation of dislocations and their dynamics in layered oxide cathode materials during battery charging," Nat. Energy **3**, 641–647 (2018).
[10]N. Bertin, R. B. Sills, and W. Cai, "Frontiers in the simulation of dislocations," Annu. Rev. Mater. Res. **50**, 437–464 (2020).
[11]J. Shin, L. Y. Chen, U. T. Sanli, G. Richter, S. Labat, M.-I. Richard, T. Cornelius, O. Thomas, and D. S. Gianola, "Controlling dislocation nucleation-mediated plasticity in nanostructures via surface modification," Acta Mater. **166**, 572–586 (2019).
[12]L. D. Nguyen, K. L. Baker, and D. H. Warner, "Atomistic predictions of dislocation nucleation with transition state theory," Phys. Rev. B **84**, 024118 (2011).
[13]D. H. Warner and W. A. Curtin, "Origins and implications of temperature-dependent activation energy barriers for dislocation nucleation in face-centered cubic metals," Acta Mater. **57**, 4267–4277 (2009).
[14]A. Ghafarollahi and W. A. Curtin, "Theory of double-kink nucleation in dilute bcc alloys," Acta Mater. **196**, 635–650 (2020).
[15]D. Rodney, "Activation enthalpy for kink-pair nucleation on dislocations: Comparison between static and dynamic atomic-scale simulations," Phys. Rev. B **76**, 144108 (2007).
[16]Y. Zhang, K. Ding, S. Stangebye, D. Chen, J. Kacher, O. Pierron, and T. Zhu, "Atomistic modeling of surface and grain boundary dislocation nucleation in fcc metals," Acta Mater. **237**, 118155 (2022).
[17]A. Yelon, B. Movaghar, and R. S. Crandall, "Multi-excitation entropy: Its role in thermodynamics and kinetics," Rep. Prog. Phys. **69**, 1145–1194 (2006).
[18]S. Ryu, K. Kang, and W. Cai, "Predicting the dislocation nucleation rate as a function of temperature and stress," J. Mater. Res. **26**, 2335–2354 (2011).
[19]S. Bagchi and D. Perez, "Anomalous entropy-driven kinetics of dislocation nucleation," Nat. Commun. **16**, 912 (2025).
[20]J. Lam and F. Pietrucci, "Critical comparison of general-purpose collective variables for crystal nucleation," Phys. Rev. E **107**, L012601 (2023).
[21]F. M. Dietrich, X. R. Advincula, G. Gobbo, M. A. Bellucci, and M. Salvalaglio, "Machine learning nucleation collective variables with graph neural networks," J. Chem. Theory Comput. **20**, 1600–1611 (2023).
[22]E. R. Beyerle, Z. Zou, and P. Tiwary, "Recent advances in describing and driving crystal nucleation using machine learning and artificial intelligence," Curr. Opin. Solid State Mater. Sci. **27**, 101093 (2023).
[23]S. W. Hall, P. Minh, and S. Sarupria, "LeaPP: Learning pathways to polymorphs through machine learning analysis of atomic trajectories," J. Chem. Theory Comput. **21**, 4121–4133 (2025).
[24]S. Hussain and A. Haji-Akbari, "Role of nanoscale interfacial proximity in contact freezing in water," J. Am. Chem. Soc. **143**, 2272–2284 (2021).
[25]Y. Bi, B. Cao, and T. Li, "Enhanced heterogeneous ice nucleation by special surface geometry," Nat. Commun. **8**, 15372 (2017).
[26]A. Haji-Akbari and P. G. Debenedetti, "Direct calculation of ice homogeneous nucleation rate for a molecular model of water," Proc. Natl. Acad. Sci. U. S. A. **112**, 10582–10588 (2015).
[27]G. C. Sosso, J. Chen, S. J. Cox, M. Fitzner, P. Pedevilla, A. Zen, and A. Michaelides, "Crystal nucleation in liquids: Open questions and future challenges in molecular dynamics simulations," Chem. Rev. **116**, 7078–7116 (2016).
[28]S. Jungblut and C. Dellago, "Pathways to self-organization: Crystallization via nucleation and growth," Eur. Phys. J. E **39**, 77 (2016).
[29]L. Lupi, A. Hudait, B. Peters, M. Grünwald, R. Gotchy Mullen, A. H. Nguyen, and V. Molinero, "Role of stacking disorder in ice nucleation," Nature **551**, 218–222 (2017).
[30]Arjun, T. A. Berendsen, and P. G. Bolhuis, "Unbiased atomistic insight in the competing nucleation mechanisms of methane hydrates," Proc. Natl. Acad. Sci. U. S. A. **116**, 19305–19310 (2019).
[31]Y. Liang, G. Díaz Leines, R. Drautz, and J. Rogal, "Identification of a multi-dimensional reaction coordinate for crystal nucleation in Ni_3Al," J. Chem. Phys. **152**, 224504 (2020).
[32]G. T. Beckham and B. Peters, "Optimizing nucleus size metrics for liquid–solid nucleation from transition paths of near-nanosecond duration," J. Phys. Chem. Lett. **2**, 1133–1138 (2011).
[33]A. H. W. Ngan, L. Zuo, and P. C. Wo, "Size dependence and stochastic nature of yield strength of micron-sized crystals: A case study on Ni_3Al," Proc. R. Soc. A **462**, 1661–1681 (2006).
[34]M. Crippa, A. Cardellini, C. Caruso, and G. M. Pavan, "Detecting dynamic domains and local fluctuations in complex molecular systems via timelapse neighbors shuffling," Proc. Natl. Acad. Sci. U. S. A. **120**, e2300565120 (2023).
[35]M. Becchi, F. Fantolino, and G. M. Pavan, "Layer-by-layer unsupervised clustering of statistically relevant fluctuations in noisy time-series data of complex dynamical systems," Proc. Natl. Acad. Sci. U. S. A. **121**, e2403771121 (2024).
[36]C. Caruso, A. Cardellini, M. Crippa, D. Rapetti, and G. M. Pavan, "TimeSOAP: Tracking high-dimensional fluctuations in complex molecular systems via time variations of SOAP spectra," J. Chem. Phys. **158**, 214302 (2023).
[37]D. Rapetti, M. Delle Piane, M. Cioni, D. Polino, R. Ferrando, and G. M. Pavan, "Machine learning of atomic dynamics and statistical surface identities in gold nanoparticles," Commun. Chem. **6**, 143 (2023).
[38]M. Cioni, D. Polino, D. Rapetti, L. Pesce, M. Delle Piane, and G. M. Pavan, "Innate dynamics and identity crisis of a metal surface unveiled by machine learning of atomic environments," J. Chem. Phys. **158**, 124701 (2023).
[39]M. Cioni, M. Delle Piane, D. Polino, D. Rapetti, M. Crippa, E. A. Irmak, S. Van Aert, S. Bals, and G. M. Pavan, "Sampling real-time atomic dynamics in metal nanoparticles by combining experiments, simulations, and machine learning," Adv. Sci. **11**, 2307261 (2024).
[40]C. Caruso, M. Crippa, A. Cardellini, M. Cioni, M. Perrone, M. Delle Piane, and G. M. Pavan, "Classification and spatiotemporal correlation of dominant fluctuations in complex dynamical systems," PNAS Nexus **4**, pgaf038 (2025).
[41]R. Becker and W. Döring, "The kinetic treatment of nuclear formation in superheated vapors," Ann. Phys. **24**, 719–752 (1935).
[42]K. Palacio-Rodriguez and F. Pietrucci, "Free energy landscapes, diffusion coefficients, and kinetic rates from transition paths," J. Chem. Theory Comput. **18**, 4639–4648 (2022).
[43]L. Mouaffac, K. Palacio-Rodriguez, and F. Pietrucci, "Optimal reaction coordinates and kinetic rates from the projected dynamics of transition paths," J. Chem. Theory Comput. **19**, 5701–5711 (2023).
[44]W. Zhang, C. Hartmann, and C. Schütte, "Effective dynamics along given reaction coordinates, and reaction rate theory," Faraday Discuss. **195**, 365–394 (2016).
[45]C. Dellago and P. G. Bolhuis, "Transition path sampling and other advanced simulation techniques for rare events," in *Advanced Computer Simulation Approaches for Soft Matter Sciences III* (Springer, 2009), pp. 167–233.
[46]P. Gkeka, G. Stoltz, A. Barati Farimani, Z. Belkacemi, M. Ceriotti, J. D. Chodera, A. R. Dinner, A. L. Ferguson, J.-B. Maillet, H. Minoux *et al.*, "Machine learning force fields and coarse-grained variables in molecular dynamics: Application

25 September 2026 07:31:09

© Author(s) 2026

to materials and biological systems," J. Chem. Theory Comput. **16**, 4757–4775 (2020).
[47]Y. Mishin, M. J. Mehl, D. A. Papaconstantopoulos, A. F. Voter, and J. D. Kress, "Structural stability and lattice defects in copper: Ab initio, tight-binding, and embedded-atom calculations," Phys. Rev. B **63**, 224106 (2001).
[48]M. I. Mendelev, M. J. Kramer, C. A. Becker, and M. Asta, "Analysis of semi-empirical interatomic potentials appropriate for simulation of crystalline and liquid al and cu," Philos. Mag. **88**, 1723–1750 (2008).
[49]S. M. Rassoulinejad-Mousavi, Y. Mao, and Y. Zhang, "Evaluation of copper, aluminum, and nickel interatomic potentials on predicting the elastic properties," J. Appl. Phys. **119**, 244304 (2016).
[50]A. P. Thompson, H. M. Aktulga, R. Berger, D. S. Bolintineanu, W. M. Brown, P. S. Crozier, P. J. in't Veld, A. Kohlmeyer, S. G. Moore, T. D. Nguyen, R. Shan, M. J. Stevens, J. Tranchida, C. Trott, and S. J. Plimpton, "LAMMPS—A flexible simulation tool for particle-based materials modeling at the atomic, meso, and continuum scales," Comput. Phys. Commun. **271**, 108171 (2022).
[51]S. Nosé, "A unified formulation of the constant temperature molecular dynamics methods," J. Chem. Phys. **81**, 511–519 (1984).
[52]W. G. Hoover and B. L. Holian, "Kinetic moments method for the canonical ensemble distribution," Phys. Lett. A **211**, 253–257 (1996).
[53]B. D. Todd and P. J. Daivis, *Nonequilibrium Molecular Dynamics: Theory, Algorithms and Applications* (Cambridge University Press, 2017).
[54]A. Stukowski, V. V. Bulatov, and A. Arsenlis, "Automated identification and indexing of dislocations in crystal interfaces," Modell. Simul. Mater. Sci. Eng. **20**, 085007 (2012).
[55]J. D. Honeycutt and H. C. Andersen, "Molecular dynamics study of melting and freezing of small Lennard-Jones clusters," J. Phys. Chem. **91**, 4950–4963 (1987).
[56]C. L. Kelchner, S. J. Plimpton, and J. C. Hamilton, "Dislocation nucleation and defect structure during surface indentation," Phys. Rev. B **58**, 11085–11088 (1998).
[57]M. L. Falk and J. S. Langer, "Dynamics of viscoplastic deformation in amorphous solids," Phys. Rev. E **57**, 7192–7205 (1998).
[58]R. Zwanzig, *Nonequilibrium Statistical Mechanics* (Oxford University Press, 2001).
[59]D. D. Girardier, H. Vroylandt, S. Bonella, and F. Pietrucci, "Inferring free-energy barriers and kinetic rates from molecular dynamics via underdamped Langevin models," J. Chem. Phys. **159**, 164111 (2023).
[60]A. N. Drozdov, "High-accuracy discrete path integral solutions for stochastic processes with noninvertible diffusion matrices," Phys. Rev. E **55**, 2496–2508 (1997).
[61]D. Doria, S. Martino, M. Becchi, and G. M. Pavan, "Data-driven assessment of optimal spatiotemporal resolutions for information extraction in noisy time series data," J. Chem. Phys. **162**, 234110 (2025).
[62]S. Martino, D. Doria, C. Lionello, M. Becchi, and G. M. Pavan, "A data driven approach to classify descriptors based on their efficiency in translating noisy trajectories into physically-relevant information," Mach. Learn.: Sci. Technol. **6**, 035039 (2025).
[63]T. S. Domingues, S. Hussain, and A. Haji-Akbari, "Divergence among local structure, dynamics, and nucleation outcome in heterogeneous nucleation of close-packed crystals," J. Phys. Chem. Lett. **15**, 1279–1287 (2024).
[64]H. Wang, W. Dmowski, Y. Tong, Z. Wang, Y. Yokoyama, J. Ketkaew, J. Schroers, and T. Egami, "Nonaffine strains control ductility of metallic glasses," Phys. Rev. Lett. **128**, 155501 (2022).
[65]J. Dong, H. Peng, H. Wang, Y. Tong, Y. Wang, W. Dmowski, T. Egami, B. Sun, W. Wang, and H. Bai, "Non-affine atomic rearrangement of glasses through stress-induced structural anisotropy," Nat. Phys. **19**, 1896–1903 (2023).
[66]A. Zaccone, *Theory of Disordered Solids* (Springer, Cham, 2023).
[67]M. Baggioli, I. Kriuchevskyi, T. W. Sirk, and A. Zaccone, "Plasticity in amorphous solids is mediated by topological defects in the displacement field," Phys. Rev. Lett. **127**, 015501 (2021).
[68]P. Desmarchelier, S. Fajardo, and M. L. Falk, "Topological characterization of rearrangements in amorphous solids," Phys. Rev. E **109**, L053002 (2024).
[69]J. Clara-Rahola, T. A. Brzinski, D. Semwogerere, K. Feitosa, J. C. Crocker, J. Sato, V. Breedveld, and E. R. Weeks, "Affine and nonaffine motions in sheared polydisperse emulsions," Phys. Rev. E **91**, 010301 (2015).
[70]A. Zaccone, "General theory of the viscosity of liquids and solids from nonaffine particle motions," Phys. Rev. E **108**, 044101 (2023).
[71]R. Milkus and A. Zaccone, "Local inversion-symmetry breaking controls the boson peak in glasses and crystals," Phys. Rev. B **93**, 094204 (2016).
[72]A. Zaccone and E. Scossa-Romano, "Approximate analytical description of the nonaffine response of amorphous solids," Phys. Rev. B **83**, 184205 (2011).
[73]J. P. Hirth, J. Lothe, and T. Mura, *Theory of Dislocations* (ASME International, 1983).

25 September 2026 07:31:09

© Author(s) 2026