ATOMWORLD-MIRROR: MACRO-STEP WORLD MODELING OF CRITICAL EVOLUTION BACKBONES FOR MATERIALS DYNAMICS

Anonymous authors
Paper under double-blind review

ABSTRACT

Atomistic simulation is a fundamental tool for studying long-term materials evolution, from diffusion and defect dynamics to interfacial reactions and fracture. Yet conventional simulators typically advance at microscopic resolution, spending substantial computation on low-impact local updates before reaching structurally consequential states—an evolutionary-resolution bottleneck that limits long-horizon simulation. We propose AtomWorld-Mirror, a time-aware macro-step world model for the critical evolution backbone of atomic systems. For Step-Wise atomistic simulation, AtomWorld-Mirror distills short micro-event segments into physically reachable transitions between key states, jointly predicting sparse structural edits and accumulated physical time through latent macro-step dynamics. Local reachability, inventory conservation, and continuous-time consistency constrain each transition. By amortizing local atomic physics into a reusable latent macro model and replacing explicit micro-event replay with macro-step inference, this formulation provides a path toward substantially faster prediction of long-term materials evolution while preserving structural validity and time semantics. Across five atomic systems, spanning Cu-rich RPV steel irradiation aging, Cu–Zr metallic glass, and Li_3N-based anti-perovskite solid electrolyte, macro-step inference delivers a speed up of 10^3 to 10^4 times over event-by-event simulation.

1 INTRODUCTION

Conventional atomistic simulation methods usually describe materials evolution through explicit time stepping or micro-event updates, e.g., kinetic Monte Carlo (KMC) and molecular dynamics (MD) (Alder & Wainwright, 1959; Bortz et al., 1975). This Step-Wise formulation treats different state changes uniformly, even though they contribute differently to long-horizon evolution. Some changes are transient fluctuations, whereas others accumulate into persistent structural progress. Figure 1 illustrates this mismatch in Cu-rich reactor pressure vessel (RPV) steel aging trajectories, where Cu-vacancy exchange events associated with key structural evolution occur sparsely and at seed-dependent times within the same fixed replay budget. As a result, standard simulators may spend substantial computation on low-impact intermediate states before reaching structurally consequential configurations (Voter et al., 2002; Soisson et al., 1996; Soisson & Fu, 2007). We characterize this mismatch as an evolutionary-resolution bottleneck. Overcoming it requires compressing low-impact intermediate states while retaining the sparse transitions that carry persistent structural progress. We refer to this retained sequence as the critical evolution backbone.

A rollout over this backbone must determine how far the system should advance while preserving both structural progress and the physical time accumulated in between. The underlying atomistic simulator resolves valid local events and their time costs, but it does not directly indicate which later state should be retained next. That state may lie many microscopic events ahead, and both its structural change and duration depend on the intervening path. Rather than resolving this path event by event, long-horizon modeling can compress it into a sequence of macro steps that predicts the next physically reachable backbone transition together with the time required to reach it.

To provide this capability, we propose AtomWorld-Mirror, a time-aware macro-step world model for atomistic evolution. AtomWorld denotes the underlying atomistic simulator, which serves during

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

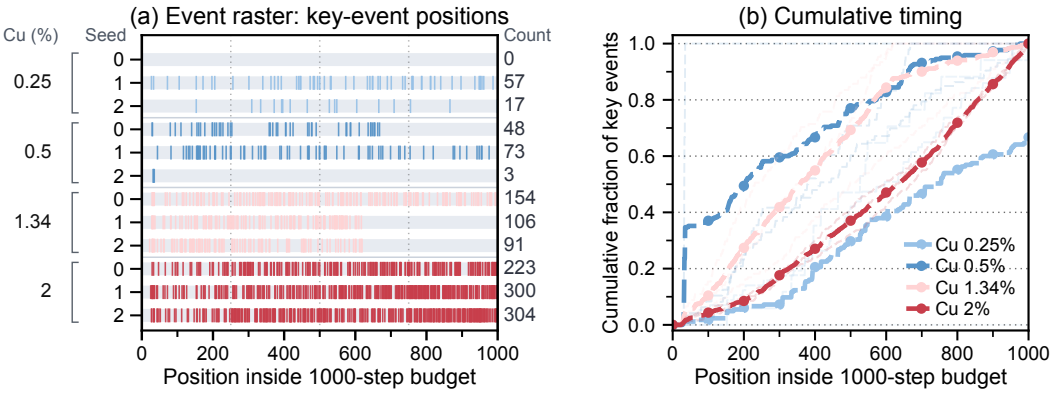


Figure 1: Random timing of key structural evolution within a fixed kinetic simulator replay budget. Each row is one kinetic simulator trajectory at a specified Cu density and random seed; ticks mark Cu-vacancy exchange events within a 1000-micro-event budget, and the cumulative curves show how these events accumulate across the budget.

training as a teacher that provides reference trajectories, local transition rules, and physical-time labels. Mirror distills these microscopic dynamics into a latent macro model that predicts reachable sparse edits and their accumulated physical duration directly from the current configuration. At rollout, it evaluates candidate macro horizons, selects a legal transition, and advances both structure and clock without explicitly replaying the intervening microscopic events (Kingma & Welling, 2014; Hafner et al., 2019; 2020; 2025; Fichthorn & Weinberg, 1991; Chatterjee & Vlachos, 2007). This yields a compressed but physically constrained state-time rollout for long-horizon structural prediction.To provide this capability, we develop AtomWorld-Mirror, a distillation framework that distills conventional atomistic simulation methods into time-aware macro-step world models.

Such compression is useful only if each predicted macro transition remains consistent with the underlying atomistic dynamics, which leads to three physical constraints. (i) *Local Reachability*. AtomWorld-Mirror restricts structural prediction to candidate supports constructed from the local event rules, ensuring that edits remain within sites accessible to the underlying dynamics. (ii) *Inventory Conservation*. The projection-closed sparse-edit formulation preserves material counts when predicted edits are converted into lattice states. (iii) *Continuous-Time Consistency*. Physical duration is modeled as a path-conditioned quantity because related macro endpoints can accumulate different waiting times along different microscopic paths (Gillespie, 1976; 1977; Fichthorn & Weinberg, 1991). Accordingly, teacher path summaries supervise duration modeling during training, while prior path latents provide the corresponding information during rollout.

Together, these components allow Mirror to replace microscopic replay with physically constrained macro evolution without discarding structural or temporal information. Across controlled atomistic simulations, AtomWorld-Mirror captures backbone-level sparse state changes through physically valid edits, aligns path-conditioned time and energy predictions along teacher trajectories, and delivers a speed up of 10^3 to 10^4 times in a timed end-to-end diagnostic against event-by-event simulation.

Our contributions are threefold:

- We introduce a simulator-to-world-model distillation method that converts step-wise atomistic teacher simulators into time-aware macro-step models over critical evolution backbones.
- We propose AtomWorld-Mirror, a time-aware macro-step world model distilled from these teacher simulators to predict critical state-time transitions beyond step-wise microscopic replay.
- Across five atomic systems with KMC and MD teachers, AtomWorld-Mirror delivers physically valid rollouts and 10^3–$10^4\times$ speedup while preserving structural and temporal accuracy.

2 RELATED WORK

Fine-Grained Atomistic Dynamics. Atomistic simulators traverse material evolution through different microscopic representations. Molecular dynamics (MD) integrates atomic coordinates through short time steps (Verlet, 1967; Plimpton, 1995), whereas kinetic Monte Carlo (KMC) samples state-dependent events and advances a residence-time clock (Bortz et al., 1975; Fichthorn & Weinberg, 1991). High-performance implementations such as OpenKMC extend event-driven KMC to very large systems (Li et al., 2019). In RPV steel aging, neutron irradiation drives the formation of Cu-rich precipitates (Soisson et al., 1996; Soisson & Fu, 2007); vacancy-mediated KMC represents this process through explicit local events, transition rates, and physical-time labels.

Accelerated Atomistic Dynamics. Several lines of work reduce the cost of traversing these trajectories. Superbasin KMC aggregates recurrent low-barrier events within local basins and advances their effective dynamics (Fichthorn & Lin, 2013). Rate Scaling KMC reduces the rates of fast events to limit repetitive sampling and accelerate access to slower evolution (Lin et al., 2019). Adaptive KMC discovers transition mechanisms and constructs event tables during simulation, extending the accessible dynamics beyond a predefined event catalogue (Henkelman & Jónsson, 2001; Chatterjee & Vlachos, 2007). Learned potentials, graph and equivariant simulators, and physics-informed models amortize local force and interaction calculations across atoms (Behler & Parrinello, 2007; Bartók et al., 2010; Schütt et al., 2017; Batatia et al., 2022; Zhang et al., 2018; Raissi et al., 2019; Pfaff et al., 2021).

3 PROBLEM SETUP

We formalize this state-time compression problem starting from the microscopic trajectory of an atomistic simulator.

3.1 MICRO-EVENT RESOLUTION AND KEY-STATE EVOLUTION

RPV steel aging makes the resolution mismatch explicit. Let X_t denote the atomic configuration after the t-th microscopic update and T_t its physical clock. A conventional simulator resolves a fine-grained trajectory $\{X_t\}_{t=0}^{N}$, where each transition $X_t \to X_{t+1}$ is a local integration step (Verlet, 1967; Plimpton, 1995; Thompson et al., 2022) or an atom-vacancy event (Bortz et al., 1975; Gillespie, 1977). Each update follows the simulator's local physics, while long-term behavior concentrates in a much smaller set of structurally decisive states. The learning problem is to identify those states and retain the structural progress they carry.

We denote a sparse key-state subsequence by

$$\mathcal{B} = (X_{n_0}, X_{n_1}, \ldots, X_{n_M}), \quad 0 = n_0 < n_1 < \cdots < n_M \le N, \quad M \ll N, \tag{1}$$

and call it the critical evolution backbone. The retention criterion is system-dependent: RPV steel retains successive Cu-cluster growth and connectivity milestones, and MD cage-swap systems retain topology-change event signatures. The realized macro horizon is $k_m = n_{m+1} - n_m \in \{1, \ldots, 1024\}$, so each transition ends at the next important state and adapts to the current configuration.

3.2 MACRO-STEP STATE-TIME OBJECTIVE

We train this predictive selection over $\mathcal{K} = \{1, \ldots, 1024\}$. For each state, the next important state determines the realized macro horizon k and the number of teacher micro events compressed into the transition. The corresponding physical duration $\tau_{t:t+k}$ is predicted jointly with the endpoint. A direct full-state objective is

$$p_\theta(X_{t+k}, \tau_{t:t+k} \mid X_t, k), \tag{2}$$

where $\tau_{t:t+k} = T_{t+k} - T_t$ for a realized path. Most lattice sites remain unchanged over a macro step, making full-state prediction inefficient. We express the structural target as sparse edit labels over a physically reachable candidate set $\mathcal{C}_k(X_t)$:

$$\Delta X_{t:t+k} = \{(i, a_i)\}_{i \in \mathcal{C}_k(X_t)}, \tag{3}$$

where a_i is the final edit label or occupancy state at candidate site i. Unchanged candidates retain their current occupancy, so the effective edit remains sparse. The horizon k determines how far the

teacher dynamics can propagate and consequently changes the candidate set. The resulting objective is

$$p_\theta(\Delta X_{t:t+k}, \tau_{t:t+k} \mid X_t, \mathcal{C}_k(X_t), k). \tag{4}$$

This objective couples structural support, endpoint edits, and path duration. A macro transition is defined by their joint prediction.

3.3 KMC SUPERVISION OF CONTINUOUS TIME

In KMC, the current state X_t has possible events $\mathcal{E}(X_t) = \{e_1, \ldots, e_n\}$ with rates $r_i(X_t)$ (Bortz et al., 1975; Gillespie, 1976; Fichthorn & Weinberg, 1991). The total escape rate is

$$R(X_t) = \sum_i r_i(X_t), \tag{5}$$

and event selection and waiting time are

$$p(e_i \mid X_t) = \frac{r_i(X_t)}{R(X_t)}, \qquad \Delta t \sim \mathrm{Exp}(R(X_t)). \tag{6}$$

KMC realizes a continuous-time Markov chain (CTMC). The normalized rates determine which local event occurs, and the total escape rate determines how long the system remains in its current state. Each micro event consequently advances both structure and physical time. Compressing several such events into one macro transition requires an endpoint edit and the time accumulated across the intervening path.

The expected waiting time at X_t is $1/R(X_t)$, and the sampled clock update is $T_{t+1} = T_t + \Delta t_t$. For a teacher segment of k events, define

$$\tau_{\exp}(t,k) = \sum_{j=0}^{k-1} \frac{1}{R(X_{t+j})}, \qquad \tau_{\mathrm{real}}(t,k) = T_{t+k} - T_t = \sum_{j=0}^{k-1} \Delta t_{t+j}. \tag{7}$$

Mirror uses $\tau_{\exp}$ as its primary macro-time target, preserving the expected CTMC clock along the teacher path. The realized duration τ_{real} records one stochastic draw and serves as an uncertainty diagnostic. Paths that reach similar endpoints through different rate environments can accumulate different durations, making macro time path-conditioned. KMC supplies all parts of this supervision directly: local event semantics, reachable trajectories, and continuous-time labels (Chatterjee & Vlachos, 2007).

4 ATOMWORLD-MIRROR

AtomWorld-Mirror implements the macro-transition objective as a teacher-student world model for backbone-aware atomistic evolution. AtomWorld denotes the teacher-exposed evolution world, including configurations, local event support, transition rules, and physical clock. Mirror is the student-side macro-step world model over backbone state-time transitions. The teacher supplies short, locally valid state-time segments, and Mirror predicts reachable sparse edits, accumulated duration, and the next macro latent state for selected horizons. The primary instantiation uses atomistic KMC; other Step-Wise atomistic simulators that provide sparse structural states and physical-time supervision can expose the same teacher interface and be distilled into Mirror.

Mirror learns macro state-time transitions between consecutive states of the critical evolution backbone $\mathcal{B}$ of Eq. 1, distilled from KMC-generated segments. For each candidate macro horizon, it takes the current lattice configuration X_t, its reachable candidate set $\mathcal{C}_k(X_t)$, and state-derived context as input. It predicts a sparse lattice edit, the next latent state, the accumulated physical duration, and reward/energy-related quantities. The planner scores the resulting transitions, retains a legal candidate, and projects its edit before execution. The projected state becomes the input to the next macro step. One macro step maps

$$\big(X_t,\ \mathcal{C}_k(X_t),\ k\big) \ \longmapsto\ \big(\Delta\hat{X}_{t:t+k},\ \hat{z}_{t+k},\ \hat{\tau}_{t:t+k}\big), \tag{8}$$

and rollout applies $\hat{X}_{t+k} = X_t \oplus \Delta\hat{X}_{t:t+k}$ and repeats. Each training example contains an endpoint edit, a microscopic path summary, and two duration targets derived from a teacher segment.

The teacher KMC simulator supplies configurations, local reachable supports, microscopic path summaries, and continuous-time labels. Mirror maps the current lattice state and active candidate patch into latent representations, predicts horizon-conditioned macro dynamics, decodes a sparse edit within the candidate support, and predicts accumulated duration. Projection precedes the latent, reward, and duration losses, giving training and rollout the same projection-closed forward path. Figure 5 shows the pipeline, Table 2 in Appendix B consolidates component-level dimensions, inputs, and outputs, and Appendix A lists the data splits and filtering rules.

4.1 TEACHER SUPERVISION

Given an initial configuration X_t, the teacher advances the underlying microscopic dynamics over a short segment of horizon $k \in \mathcal{K}$, producing the local trajectory $\{X_{t+i}\}_{i=0}^{k}$ (Fichthorn & Weinberg, 1991; Chatterjee & Vlachos, 2007). Each supervised macro sample contains

$$(X_t, k, X_{t+k}, \Delta X_{t:t+k}, \tau_{\text{exp}}, \tau_{\text{real}}, s_{\text{path}}, \mathcal{C}_k). \tag{9}$$

Here $\mathcal{K}$ is the controlled Multi-K horizon set, $\Delta X_{t:t+k}$ is the start-to-end sparse edit, s_{path} summarizes the microscopic path, and $\mathcal{C}_k$ is the reachable candidate set. Rollout receives the current state X_t, candidate support $\mathcal{C}_k(X_t)$, horizon k, and state-derived context. The endpoint X_{t+k}, sparse-edit target, path summary, and duration labels supervise training. The realized duration is

$$\tau_{\text{real}} = \sum_{j=0}^{k-1} \Delta t_{t+j}, \tag{10}$$

while the primary time target is the path-conditioned accumulated expected time

$$\tau_{\text{exp}} = \sum_{j=0}^{k-1} \frac{1}{R(X_{t+j})}. \tag{11}$$

We use τ_{exp} as the main duration target because it reflects the expected CTMC clock. τ_{real} remains an auxiliary stochastic target and time-uncertainty diagnostic.

The three physical commitments stated in Section 1 turn these supervised segments into Mirror rollouts: local reachability restricts edits to the teacher-derived candidate support, inventory conservation constrains the projection that converts latent predictions into lattice edits, and continuous-time consistency makes duration path-conditioned (Gillespie, 1976; 1977; Fichthorn & Weinberg, 1991). These commitments enter Mirror through latent dynamics, sparse-edit projection, and duration losses.

4.2 MIRROR MACRO-STEP WORLD MODEL

Mirror is a macro-step world model developed from the DreamerV4 latent dynamics architecture (Hafner et al., 2025) and adapted to teacher-supervised atomistic state-time transitions. Its encoder combines a latent-variable representation with graph message passing and equivariant site features (Kingma & Welling, 2014; Gilmer et al., 2017; Satorras et al., 2021),

$$z_t = \text{Enc}_\theta(X_t, g_t), \tag{12}$$

where g_t may contain local composition, defect statistics, rate summaries, vacancy information, or other global context. The active-patch encoder also produces site embeddings for $\mathcal{C}_k(X_t)$, coupling global latent dynamics with the local edit space. The same X_t can lead to different reachable endpoints, and each k defines a distinct admissible path family. AtomWorld-Mirror represents this variation with a horizon-conditioned path latent h_t.

During training, the posterior path latent conditions on the endpoint and teacher path summary,

$$q_\phi(h_t \mid z_t, z_{t+k}, s_{\text{path}}, k), \tag{13}$$

which identifies the supervised transition. During inference, the model samples or selects from a prior

$$p_\theta(h_t \mid z_t, g_t, k). \tag{14}$$

The macro dynamics then predicts a horizon-conditioned next latent,

$$\hat{z}_{t+k} = z_t + F_\theta(z_t, h_t, k). \tag{15}$$

A KL term transfers posterior information into the rollout prior. At rollout, every input is computed from the current state, state-derived context, and macro horizon.

4.3 REACHABILITY-CONSTRAINED SPARSE EDITS

The edit decoder operates on $\mathcal{C}_k(X_t)$, the reachable candidate support generated by the teacher dynamics. Candidate-restricted event supports are standard in spatial and adaptive KMC (Chatterjee & Vlachos, 2007; Henkelman & Jónsson, 2001), with related barrier and path-sampling context in NEB and time-reversal path-sampling work (Henkelman et al., 2000; Liu, 2023). For candidate site i, the decoder takes site embedding u_i, patch context, predicted macro latent $\hat{z}_{t+k}$, path latent h_t, and horizon k, and predicts change logit ℓ_i and final type distribution π_i,

$$(\ell_i, \pi_i) = D_\theta(u_i, c_{k,i}, \hat{z}_{t+k}, h_t, k), \tag{16}$$

Let $m_i = \mathbf{1}\{x_{t+k,i} \neq x_{t,i}\}$ be the teacher change mask and a_i be the target final type. A compact sparse-edit loss is

$$\mathcal{L}_{\text{edit}} = -\sum_{i \in \mathcal{C}_k(X_t)} \left[m_i \log \sigma(\ell_i) + (1 - m_i) \log(1 - \sigma(\ell_i)) + \beta m_i \log \pi_i(a_i) \right], \tag{17}$$

where $c_{k,i}$ is the active-patch context, the first two terms learn sparse support, the last learns changed-site type, and $\beta = 1$ in the reported experiments. This candidate-restricted output space focuses prediction on sites that the local dynamics can affect and supplies the support for physically reachable vacancy-mediated edits (Soisson et al., 1996; Soisson & Fu, 2007).

Inventory conservation further requires the candidate patch to preserve material counts:

$$\sum_{i \in \mathcal{C}_k(X_t)} \mathbf{1}\{\hat{x}_{t+k,i} = c\} = \sum_{i \in \mathcal{C}_k(X_t)} \mathbf{1}\{x_{t,i} = c\}, \qquad c \in \{\text{Fe}, \text{Cu}, \text{Vac}\}, \tag{18}$$

Outside the patch, sites are copied from the current configuration. Projection is implemented as paired vacancy–atom transport, selecting vacancy-to-atom and atom-to-vacancy pairs under the budget

$$d_{\text{vac}} \leq k, \qquad d_{\text{Cu}} \leq k, \qquad d_{\text{vac}} + d_{\text{Cu}} \leq 2k, \qquad |\Delta \hat{X}_{t:t+k}| \leq 2k, \tag{19}$$

where distances use the periodic BCC hop metric. Projection is applied before latent, reward, and duration losses are evaluated. The projected state is re-encoded and passed to the next macro step, aligning optimization with the state consumed during rollout.

4.4 CONTINUOUS-TIME DURATION MODELING AND LOSS

Duration is a core macro-transition output because stochastic simulation couples event selection and waiting time through the same transition rates (Gillespie, 1976; 1977; Fichthorn & Weinberg, 1991; Lin et al., 2019). We separate path-conditioned expected CTMC duration from realized sampled duration: τ_{exp} is the main macro-time target, while τ_{real} is auxiliary stochastic supervision.

The duration branch predicts expected macro time in log space,

$$b_k(X_t) = \log k - \log R(X_t), \qquad \mu_{\text{exp}} = b_k(X_t) + r_\theta^\tau(z_t, \hat{z}_{t+k}, h_t, g_t, \rho_t, k), \tag{20}$$

where $b_k(X_t)$ is the log start-state CTMC estimate $k/R(X_t)$, ρ_t is edit/projection context, and r_θ^τ is a residual. The residual corrects the start-state estimate using rate changes accumulated along the teacher path. The expected-time head captures both the initial escape-rate scale and path-dependent dynamics. It also predicts $\log \sigma_{\text{exp}}$, and the main duration loss is the implementation's log-space Gaussian negative log-likelihood:

$$\mathcal{L}_\tau = \log \sigma_{\text{exp}} + \frac{1}{2} \left(\frac{\log(\tau_{\text{exp}} + \epsilon) - \mu_{\text{exp}}}{\sigma_{\text{exp}}} \right)^2. \tag{21}$$

Realized duration remains stochastic even under identical rates, so it is modeled as a conditional lognormal auxiliary target:

$$\mu_{\text{real}} = b_k(X_t) + r_\theta^{\text{real}}(z_t, \hat{z}_{t+k}, h_t, g_t, \rho_t, k), \qquad \tau_{\text{real}} \sim \text{LogNormal}(\mu_{\text{real}}, \sigma_{\text{real}}). \tag{22}$$

Default training uses prior-side duration losses, matching the quantities available during rollout. We use smooth-L_1 regression for $\mathcal{L}_{\text{reward}}$ and $\mathcal{L}_{\text{latent}}$. The projection consistency loss $\mathcal{L}_{\text{proj}}$ aligns

the re-encoded projected state with the teacher endpoint latent, and $\mathcal{L}_{\text{real}}$ is the lognormal negative log-likelihood of τ_{real}. The implementation objective is

$$\begin{aligned}\mathcal{L} = \lambda_{\text{edit}}\mathcal{L}_{\text{edit}} + \lambda_{\text{reward}}\mathcal{L}_{\text{reward}} + \lambda_{\text{latent}}\mathcal{L}_{\text{latent}} + \lambda_{\text{proj}}\mathcal{L}_{\text{proj}} \\ + \lambda_{\text{KL}}D_{\text{KL}}(q_\phi \parallel p_\theta) + \lambda_\tau\mathcal{L}_\tau + \lambda_{\text{real}}\mathcal{L}_{\text{real}}.\end{aligned} \tag{23}$$

The macro horizon is selected independently at each step from $\mathcal{K} = \{1, \ldots, 1024\}$. The model encodes X_t, draws h_t from the prior, and decodes a projected sparse edit for each candidate k. Candidates violating reachability or inventory are discarded, and the surviving legal candidates are ranked by predicted structural progress per unit physical time:

$$k^\star = \arg\max_{k \in \mathcal{K}_{\text{legal}}} \frac{\hat{r}_\theta(z_t, \hat{z}_{t+k}, k)}{\hat{\tau}_{t:t+k} + \epsilon}, \tag{24}$$

The selected edit is applied as $\hat{X}_{t+k} = X_t \oplus \Delta\hat{X}_{t:t+k}$, the duration $\hat{\tau}_{t:t+k}$ advances the clock, and the step repeats to form a backbone trajectory.

The backbone $\mathcal{B}$ of Eq. 1 remains stable across realized macro horizons and trajectory seeds. We measure sensitivity over $k \in \{1, 2, 4, \ldots, 1024\}$ micro steps on 16 held-out trajectories and compute the persistence ratio, defined as the fraction of backbone events surviving across horizons. Persistence stays above 0.98 throughout. Across trajectory seeds, the largest-Cu-cluster identity reaches 0.9066, showing that the persistent structural hierarchy remains stable under microscopic stochasticity.

5 Experimental design

5.1 Experimental scope

Across five atomic systems, Mirror maintains zero reachability violations and zero inventory violations, and delivers an end-to-end Speed Up of 10^3 to 10^4 times. A reachability violation is an edit outside the candidate set $\mathcal{C}_k(X_t)$, and an inventory violation is a projected patch that changes any species count; both are reported as the fraction of macro steps affected. The evaluation begins with binary RPV steel aging and extends to 3-element and 10-element RPV steels under the same KMC teacher. Cu–Zr metallic glass and Li_3N-based anti-perovskite solid electrolyte use MD as the teacher, with cage-swap topology events defining their macro-step boundaries. The primary experiments instantiate this interface with KMC, and Appendix D reports all per-system results, including the MD-teacher systems that reuse the same Mirror architecture and training loop with tuple construction adapted to topology-change events.

The RPV steel systems center the study on alloy aging, where Cu-rich precipitate formation under irradiation controls the degradation of nuclear-reactor pressure vessels (Soisson et al., 1996; Soisson & Fu, 2007; Li et al., 2019). Cu–Zr metallic glass extends the evaluation to an amorphous alloy with cage-swap dynamics and structural relaxation (Cheng & Ma, 2011). The Li_3N-based anti-perovskite solid electrolyte extends it to ionic migration in a solid-electrolyte composition for lithium batteries (Cordier et al., 1989; Dembitskiy et al., 2025; Zhao & Daemen, 2012). Together, these systems test state-time macro modeling across composition, structural order, and simulator type, while the measured Speed Up supports longer aging, diffusion, and parameter-sweep studies.

AtomWorld-Mirror uses a state-time macro transition as its prediction unit. It identifies structurally decisive waypoints along the critical evolution backbone and distills segments from a traditional atomistic simulator teacher into sparse reachable edits, next latent states, and accumulated physical duration. Each transition retains the local support and inventory conditions required for a valid rollout, while macro-step inference bypasses explicit replay of intervening micro-events. We apply this teacher interface to event-driven KMC, MD-generated topology-change segments in Cu–Zr metallic glass (Cheng & Ma, 2011), and Li_3N-based anti-perovskite solid electrolyte (Cordier et al., 1989; Dembitskiy et al., 2025; Zhao & Daemen, 2012), as well as to other Step-Wise atomistic simulators.

5.2 Event-driven instantiation

We instantiate AtomWorld as an atom-vacancy KMC process to evaluate projected edit validity, expected-duration alignment, closed-loop rollout consistency, and timed macro-inference efficiency.

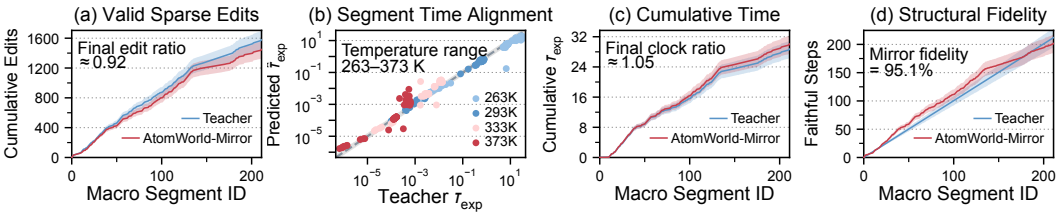


Figure 2: Controlled macro-step validation against the teacher simulator. The panels show (a) cumulative structural edits, (b) single-segment expected-time alignment across temperatures, (c) long-trajectory cumulative expected time, and (d) cumulative correctly typed edits for Teacher and Mirror.

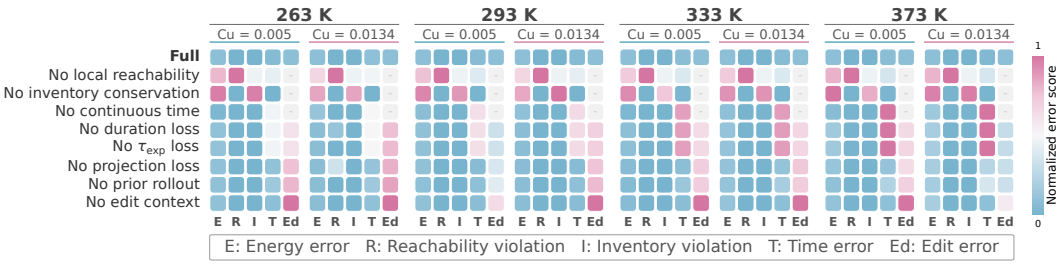


Figure 3: Cu-density and temperature ablation matrices for Multi-K teacher-probe rollout. Darker cells denote larger values of the corresponding failure metric.

Unless otherwise stated, the controlled physical setup uses a 40^3 BCC lattice, Cu density 0.0134, vacancy density 0.0002, a 2NN event neighborhood, 16 defect-graph shells, and reward scale 10.0. The quantitative matrix uses adaptive $k \in \{1, \ldots, 1024\}$ and horizon-selected macro diagnostics. Figure 2 summarizes the controlled macro-step validation against the KMC teacher/reference.

Given these paired macro segments, the first evaluation separates raw prediction quality from projection-closed physical validity. Structural quality is summarized by the edit F1 score, the harmonic mean of precision and recall between predicted and teacher change masks over $\mathcal{C}_k(X_t)$, and by type accuracy, the fraction of teacher-changed sites whose predicted species matches a_i. The expected-time head is evaluated against the path-conditioned CTMC target τ_{exp} using log-MAE, log-RMSE, log-correlation, and scale ratio:

$$\text{logMAE}_\tau = \frac{1}{N}\sum_{n=1}^{N}\left|\log \hat{\tau}_{\text{exp}}^{(n)} - \log \tau_{\text{exp}}^{(n)}\right|, \qquad \text{scale}_\tau = \frac{1}{N}\sum_{n=1}^{N}\frac{\hat{\tau}_{\text{exp}}^{(n)}}{\tau_{\text{exp}}^{(n)}}. \tag{25}$$

Scale ratio tests systematic time bias, log-correlation tests relative variation across teacher segments, and log-RMSE weights the largest per-segment errors. The stochastic τ_{real} head is evaluated separately as an auxiliary lognormal using NLL, probability-integral-transform diagnostics, and 68%/95% interval coverage, the fraction of segments whose realized duration falls inside the predicted interval. Appendix A reports training parameters, compute resources, and runtime settings.

5.3 Closed-loop validity and ablations

The model maintains an autonomous state $\hat{X}_m$ by applying the previous projected sparse edit. The evaluated model trajectory is the closed-loop sequence $\hat{X}_0 \xrightarrow{k_0} \hat{X}_1 \xrightarrow{k_1} \cdots \xrightarrow{k_{M-1}} \hat{X}_M$, while each transition is scored against a KMC probe started from the current model state. This protocol evaluates whether the self-generated macro trajectory remains locally reachable, inventory-preserving, temporally calibrated, and physically plausible under local probes.

Figure 3 reports Multi-K teacher-probe ablation matrices across Cu concentrations 0.005 and 0.0134 and temperatures 263, 293, 333, and 373K. Shared ablation rows measure candidate-support validity,

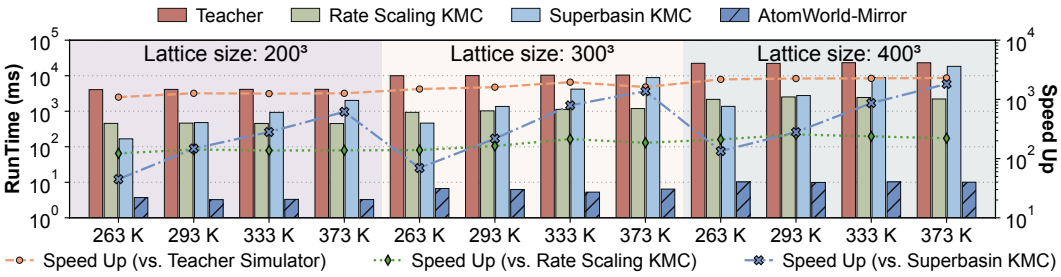


Figure 4: End-to-end timing diagnostic across lattice sizes and temperatures.

inventory preservation, path-conditioned duration, and the components that maintain editable states during rollout. These matrices use two non-learned references. The KMC teacher supplies the rate-governed state-time target, and the copy-state baseline keeps the lattice fixed while using the CTMC start-state duration from $R(X_t)$. Removing local reachability measures candidate-support dependence, removing inventory conservation measures material-count preservation, and replacing learned duration with a start-state CTMC baseline measures the contribution of path-conditioned information.

5.4 Long-horizon diagnostics and efficiency

Replacing explicit micro-event replay with macro-step inference delivers a Speed Up of 10^3 to 10^4 times across lattice size, temperature, and material system (Fig. 4, Appendix D, Appendix I). The timing includes candidate construction, projection, benchmark orchestration, and batched neural inference for one adaptive macro transition. The diagnostic measures replay reduction under reachable, inventory-preserving, and temporally calibrated edits, with the reference cost set by classical and KMC simulators (Plimpton, 1995; Thompson et al., 2022; Li et al., 2019).

Macro-step cost tracks the candidate patch defined by the event neighborhood. Inference cost stays near-constant in system size, and a model trained on 40^3 runs on larger configurations without retraining. Across Multi-K steps with k ranging from 1 to 1024, one macro step replaces 480 teacher micro events on average and advances the clock by 12.98 seconds of expected time; Appendix H reports this timing.

Capturing Cu-rich precipitate evolution in RPV steel requires following vacancy-mediated kinetics across large spatial domains and long physical timescales. We evaluate a 50-year autonomous closed-loop rollout of Mirror on the full 5.4×10^{10}-atom RPV steel aging system across four temperatures in Appendix I.

The teacher-forced diagnostic measures state-conditioned time and reward alignment along a contiguous KMC trajectory $\{X_{t_m}\}_{m=0}^{M}$, where $X_{t_m} \xrightarrow[\text{teacher}]{k_m} X_{t_{m+1}}$. The readout accumulates predicted and reference expected times and the final clock-scale ratio:

$$\hat{T}_M^{\text{exp}} = \sum_{m=0}^{M-1} \hat{\tau}_{\text{exp},m}, \qquad T_M^{\text{exp}} = \sum_{m=0}^{M-1} \tau_{\text{exp},m}, \qquad \rho_T = \frac{\hat{T}_M^{\text{exp}}}{T_M^{\text{exp}} + \epsilon}. \tag{26}$$

6 Conclusion

We introduced a simulator-to-world-model distillation method that converts step-wise atomistic teachers into time-aware macro-step models over critical evolution backbones. AtomWorld-Mirror instantiates this framework by predicting reachable sparse edits and path-conditioned physical time while preserving material inventory. Across five atomic systems with KMC and MD teachers, it achieves physically valid rollouts and 10^3–$10^4\times$ speedups over event-by-event simulation.

AI USE STATEMENT

In this work, we used generative AI (large language model) tools solely as a writing aid: to help translate and polish the phrasing of the manuscript (word choice, grammar, and clarity edits) during preparation of this paper. We did not use generative AI tools to design the method, derive the formulation, run experiments, generate results, or produce the reported findings; the core methodology, experimental design, implementation, and analysis are the authors' own work. All AI-assisted text was reviewed and edited by the authors. We take responsibility for the final content of this work, including text, claims, and artifacts.

ETHICS STATEMENT

This work uses simulated atomistic data from kinetic Monte Carlo teachers for RPV steel alloys and molecular dynamics teachers for Cu–Zr metallic glass and Li_3N-based anti-perovskite solid electrolyte; it does not involve human subjects, personal data, crowdsourcing, or deployed decision systems, and we are not aware of any direct negative societal impact of the method as presented. The released code and modeling assets build on cited open simulation, modeling, and formatting resources, whose original authors are credited via citation in the main text and appendix.

REPRODUCIBILITY STATEMENT

Sections 4 and 5 and Appendix A describe the model components, teacher supervision, loss terms, data splits, evaluation metrics, closed-loop teacher-probe protocol, teacher-forced diagnostic protocol, and training/evaluation configuration (Table 1), including the single-GPU hardware setting and runtimes. An anonymized core implementation, minimal Multi-K configuration, synthetic smoke test, and batch-contract documentation for the macro-model interface accompany this submission as supplementary material for independent verification.

REFERENCES

B. J. Alder and T. E. Wainwright. Studies in Molecular Dynamics. I. General Method. *The Journal of Chemical Physics*, 31(2):459–466, 1959. doi: 10.1063/1.1730376.

Albert P. Bartók, Mike C. Payne, Risi Kondor, and Gábor Csányi. Gaussian Approximation Potentials: The Accuracy of Quantum Mechanics, Without the Electrons. *Physical Review Letters*, 104:136403, 2010. doi: 10.1103/PhysRevLett.104.136403.

Ilyes Batatia, Dávid Péter Kovács, Gregor N. C. Simm, Christoph Ortner, and Gábor Csányi. MACE: Higher Order Equivariant Message Passing Neural Networks for Fast and Accurate Force Fields. In *Advances in Neural Information Processing Systems*, volume 35, pp. 11423–11436, 2022. URL https://proceedings.neurips.cc/paper_files/paper/2022/hash/4a36c3c51af11ed9f34615b81edb5bbc-Abstract-Conference.html.

Peter Battaglia, Razvan Pascanu, Matthew Lai, Danilo Jimenez Rezende, and Koray Kavukcuoglu. Interaction Networks for Learning About Objects, Relations and Physics. In *Advances in Neural Information Processing Systems*, volume 29, pp. 4502–4510, 2016. URL https://proceedings.neurips.cc/paper/2016/hash/3147da8ab4a0437c15ef51a5cc7f2dc4-Abstract.html.

Simon Batzner, Albert Musaelian, Lixin Sun, Mario Geiger, Jonathan P. Mailoa, Mordechai Kornbluth, Nicola Molinari, Tess E. Smidt, and Boris Kozinsky. E(3)-Equivariant Graph Neural Networks for Data-Efficient and Accurate Interatomic Potentials. *Nature Communications*, 13:2453, 2022. doi: 10.1038/s41467-022-29939-5.

Jörg Behler and Michele Parrinello. Generalized Neural-Network Representation of High-Dimensional Potential-Energy Surfaces. *Physical Review Letters*, 98:146401, 2007. doi: 10.1103/PhysRevLett.98.146401.

Troels Arnfred Bojesen. Policy-Guided Monte Carlo: Reinforcement-Learning Markov Chain Dynamics. *Physical Review E*, 98:063303, 2018. doi: 10.1103/PhysRevE.98.063303.

A. B. Bortz, M. H. Kalos, and J. L. Lebowitz. A New Algorithm for Monte Carlo Simulation of Ising Spin Systems. *Journal of Computational Physics*, 17(1):10–18, 1975. doi: 10.1016/0021-9991(75)90060-1.

Abhijit Chatterjee and Dionisios G. Vlachos. An Overview of Spatial Microscopic and Accelerated Kinetic Monte Carlo Methods. *Journal of Computer-Aided Materials Design*, 14:253–308, 2007. doi: 10.1007/s10820-006-9042-9.

Yongqiang Cheng and Evan Ma. Atomic-level structure and structure–property relationship in metallic glasses. *Progress in Materials Science*, 56(4):379–473, 2011. doi: 10.1016/j.pmatsci.2010.12.002.

John D. Chodera and Frank Noé. Markov State Models of Biomolecular Conformational Dynamics. *Current Opinion in Structural Biology*, 25:135–144, 2014. doi: 10.1016/j.sbi.2014.04.002.

Kamal Choudhary, Kevin F. Garrity, Andrew C. E. Reid, Brian DeCost, Adam J. Biacchi, Angela R. Hight Walker, Zachary Trautt, Jason Hattrick-Simpers, A. Gilad Kusne, Andrea Centrone, et al. The Joint Automated Repository for Various Integrated Simulations (JARVIS) for Data-Driven Materials Design. *npj Computational Materials*, 6:173, 2020. doi: 10.1038/s41524-020-00440-1.

Gerhard Cordier, Axel Gudat, Rüdiger Kniep, and Albrecht Rabenau. LiCaN and Li_4SrN_2, Derivatives of the Fluorite and Lithium Nitride Structures. *Angewandte Chemie International Edition in English*, 28(12):1702–1703, 1989. doi: 10.1002/anie.198917021.

Stefano Curtarolo, Wahyu Setyawan, Gus L. W. Hart, Michal Jahnatek, Roman V. Chepulskii, Richard H. Taylor, Shidong Wang, Junkai Xue, Kesong Yang, Ohad Levy, et al. AFLOW: An Automatic Framework for High-Throughput Materials Discovery. *Computational Materials Science*, 58:218–226, 2012. doi: 10.1016/j.commatsci.2012.02.005.

Artem D. Dembitskiy, Innokentiy S. Humonen, Roman A. Eremin, Dmitry A. Aksyonov, Stanislav S. Fedotov, and Semen A. Budennyy. Benchmarking machine learning models for predicting lithium ion migration. *npj Computational Materials*, 11(1):131, 2025. doi: 10.1038/s41524-025-01571-z.

Kristen A. Fichthorn and Yangzheng Lin. A Local Superbasin Kinetic Monte Carlo Method. *The Journal of Chemical Physics*, 138(16):164104, 2013. doi: 10.1063/1.4801869.

Kristen A. Fichthorn and William H. Weinberg. Theoretical Foundations of Dynamical Monte Carlo Simulations. *The Journal of Chemical Physics*, 95(2):1090–1096, 1991. doi: 10.1063/1.461138.

Thomas Garnier and Maylise Nastar. Coarse-Grained Kinetic Monte Carlo Simulation of Diffusion in Alloys. *Physical Review B*, 88:134207, 2013. doi: 10.1103/PhysRevB.88.134207.

Johannes Gasteiger, Janek Groß, and Stephan Günnemann. Directional Message Passing for Molecular Graphs. In *International Conference on Learning Representations*, 2020. URL https://openreview.net/forum?id=B1eWbxStPH.

Daniel T. Gillespie. A General Method for Numerically Simulating the Stochastic Time Evolution of Coupled Chemical Reactions. *Journal of Computational Physics*, 22(4):403–434, 1976. doi: 10.1016/0021-9991(76)90041-3.

Daniel T. Gillespie. Exact Stochastic Simulation of Coupled Chemical Reactions. *The Journal of Physical Chemistry*, 81(25):2340–2361, 1977. doi: 10.1021/j100540a008.

Justin Gilmer, Samuel S. Schoenholz, Patrick F. Riley, Oriol Vinyals, and George E. Dahl. Neural Message Passing for Quantum Chemistry. In *Proceedings of the 34th International Conference on Machine Learning*, volume 70 of *Proceedings of Machine Learning Research*, pp. 1263–1272. PMLR, 2017. URL https://proceedings.mlr.press/v70/gilmer17a.html.

Danijar Hafner, Timothy Lillicrap, Ian Fischer, Ruben Villegas, David Ha, Honglak Lee, and James Davidson. Learning Latent Dynamics for Planning from Pixels. In *Proceedings of the 36th International Conference on Machine Learning*, volume 97 of *Proceedings of Machine Learning Research*, pp. 2555–2565. PMLR, 2019. URL https://proceedings.mlr.press/v97/hafner19a.html.

Danijar Hafner, Timothy Lillicrap, Jimmy Ba, and Mohammad Norouzi. Dream to Control: Learning Behaviors by Latent Imagination. In *International Conference on Learning Representations*, 2020. URL https://openreview.net/forum?id=S1lOTC4tDS.

Danijar Hafner, Wilson Yan, and Timothy Lillicrap. Training Agents Inside of Scalable World Models. *arXiv preprint arXiv:2509.24527*, 2025. doi: 10.48550/arXiv.2509.24527.

Graeme Henkelman and Hannes Jónsson. Long Time Scale Kinetic Monte Carlo Simulations Without Lattice Approximation and Predefined Event Table. *The Journal of Chemical Physics*, 115(21): 9657–9666, 2001. doi: 10.1063/1.1415500.

Graeme Henkelman, Blas P. Uberuaga, and Hannes Jónsson. A Climbing Image Nudged Elastic Band Method for Finding Saddle Points and Minimum Energy Paths. *The Journal of Chemical Physics*, 113(22):9901–9904, 2000. doi: 10.1063/1.1329672.

Anubhav Jain, Shyue Ping Ong, Geoffroy Hautier, Wei Chen, William Davidson Richards, Stephen Dacek, Shreyas Cholia, Dan Gunter, David Skinner, Gerbrand Ceder, et al. Commentary: The Materials Project: A Materials Genome Approach to Accelerating Materials Innovation. *APL Materials*, 1:011002, 2013. doi: 10.1063/1.4812323.

George Em Karniadakis, Ioannis G. Kevrekidis, Lu Lu, Paris Perdikaris, Sifan Wang, and Liu Yang. Physics-Informed Machine Learning. *Nature Reviews Physics*, 3:422–440, 2021. doi: 10.1038/s42254-021-00314-5.

Diederik P. Kingma and Max Welling. Auto-Encoding Variational Bayes. In *International Conference on Learning Representations*, 2014. doi: 10.48550/arXiv.1312.6114.

Thomas N. Kipf, Ethan Fetaya, Kuan-Chieh Wang, Max Welling, and Richard S. Zemel. Neural Relational Inference for Interacting Systems. In *Proceedings of the 35th International Conference on Machine Learning*, volume 80 of *Proceedings of Machine Learning Research*, pp. 2688–2697. PMLR, 2018. URL https://proceedings.mlr.press/v80/kipf18a.html.

Kun Li, Honghui Shang, Yunquan Zhang, Shigang Li, Baodong Wu, Dong Wang, Libo Zhang, Fang Li, Dexun Chen, and Zhiqiang Wei. OpenKMC: A KMC Design for Hundred-Billion-Atom Simulation Using Millions of Cores on Sunway TaihuLight. In *Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis*, pp. 1–16. ACM, 2019. doi: 10.1145/3295500.3356165. Article 68, 16 pages.

Yen Ting Lin, Song Feng, and William S. Hlavacek. Scaling Methods for Accelerating Kinetic Monte Carlo Simulations of Chemical Reaction Networks. *The Journal of Chemical Physics*, 150(24): 244101, 2019. doi: 10.1063/1.5096774.

Zhirong Liu. Accelerating Kinetics with Time-Reversal Path Sampling. *Molecules*, 28(24):8147, 2023. doi: 10.3390/molecules28248147.

Siewert J. Marrink, H. Jelger Risselada, Serge Yefimov, D. Peter Tieleman, and Alex H. de Vries. The MARTINI Force Field: Coarse Grained Model for Biomolecular Simulations. *The Journal of Physical Chemistry B*, 111(27):7812–7824, 2007. doi: 10.1021/jp071097f.

Amil Merchant, Simon Batzner, Samuel S. Schoenholz, Muratahan Aykol, Gowoon Cheon, and Ekin D. Cubuk. Scaling Deep Learning for Materials Discovery. *Nature*, 624:80–85, 2023. doi: 10.1038/s41586-023-06735-9.

Albert Musaelian, Simon Batzner, Anders Johansson, Lixin Sun, Cameron J. Owen, Mordechai Kornbluth, and Boris Kozinsky. Learning Local Equivariant Representations for Large-Scale Atomistic Dynamics. *Nature Communications*, 14:579, 2023. doi: 10.1038/s41467-023-36329-y.

W. G. Noid, Jhih-Wei Chu, Gary S. Ayton, Vinod Krishna, Sergei Izvekov, Gregory A. Voth, Avisek Das, and Hans C. Andersen. The Multiscale Coarse-Graining Method. I. A Rigorous Bridge Between Atomistic and Coarse-Grained Models. *The Journal of Chemical Physics*, 128:244114, 2008. doi: 10.1063/1.2938860.

Shyue Ping Ong, William Davidson Richards, Anubhav Jain, Geoffroy Hautier, Michael Kocher, Shreyas Cholia, Dan Gunter, Vincent L. Chevrier, Kristin A. Persson, and Gerbrand Ceder. Python Materials Genomics (pymatgen): A Robust, Open-Source Python Library for Materials Analysis. *Computational Materials Science*, 68:314–319, 2013. doi: 10.1016/j.commatsci.2012.10.028.

Tobias Pfaff, Meire Fortunato, Alvaro Sanchez-Gonzalez, and Peter W. Battaglia. Learning Mesh-Based Simulation with Graph Networks. In *Proceedings of the 38th International Conference on Machine Learning*, volume 139 of *Proceedings of Machine Learning Research*. PMLR, 2021. URL https://openreview.net/forum?id=roNqYL0_XP.

Steve Plimpton. Fast Parallel Algorithms for Short-Range Molecular Dynamics. *Journal of Computational Physics*, 117(1):1–19, 1995. doi: 10.1006/jcph.1995.1039.

Aneesur Rahman. Correlations in the Motion of Atoms in Liquid Argon. *Physical Review*, 136: A405–A411, 1964. doi: 10.1103/PhysRev.136.A405.

M. Raissi, P. Perdikaris, and G. E. Karniadakis. Physics-Informed Neural Networks: A Deep Learning Framework for Solving Forward and Inverse Problems Involving Nonlinear Partial Differential Equations. *Journal of Computational Physics*, 378:686–707, 2019. doi: 10.1016/j.jcp.2018.10.045.

Víctor Garcia Satorras, Emiel Hoogeboom, and Max Welling. E(n) Equivariant Graph Neural Networks. In *Proceedings of the 38th International Conference on Machine Learning*, volume 139 of *Proceedings of Machine Learning Research*, pp. 9323–9332. PMLR, 2021. URL https://proceedings.mlr.press/v139/satorras21a.html.

John Schulman, Filip Wolski, Prafulla Dhariwal, Alec Radford, and Oleg Klimov. Proximal Policy Optimization Algorithms. *arXiv preprint arXiv:1707.06347*, 2017. doi: 10.48550/arXiv.1707.06347.

Kristof T. Schütt, Pieter-Jan Kindermans, Huziel Enoc Sauceda Felix, Stefan Chmiela, Alexandre Tkatchenko, and Klaus-Robert Müller. SchNet: A Continuous-Filter Convolutional Neural Network for Modeling Quantum Interactions. In *Advances in Neural Information Processing Systems*, volume 30, pp. 991–1001, 2017. URL https://proceedings.neurips.cc/paper/2017/hash/303ed4c69846ab36c2904d3ba8573050-Abstract.html.

F. Soisson, A. Barbu, and G. Martin. Monte Carlo Simulations of Copper Precipitation in Dilute Iron-Copper Alloys During Thermal Ageing and Under Electron Irradiation. *Acta Materialia*, 44 (9):3789–3800, 1996. doi: 10.1016/1359-6454(95)00447-5.

Frédéric Soisson and Chu-Chun Fu. Cu-Precipitation Kinetics in α-Fe from Atomistic Simulations: Vacancy-Trapping Effects and Cu-Cluster Mobility. *Physical Review B*, 76:214102, 2007. doi: 10.1103/PhysRevB.76.214102.

Mads R. Sørensen and Arthur F. Voter. Temperature-Accelerated Dynamics for Simulation of Infrequent Events. *The Journal of Chemical Physics*, 112(21):9599–9606, 2000. doi: 10.1063/1.481576.

Hao Tang, Boning Li, Yixuan Song, Mengren Liu, Haowei Xu, Guoqing Wang, Heejung Chung, and Ju Li. Reinforcement Learning-Guided Long-Timescale Simulation of Hydrogen Transport in Metals. *Advanced Science*, 11(5):2304122, 2024. doi: 10.1002/advs.202304122.

Aidan P. Thompson, H. Metin Aktulga, Richard Berger, Dan S. Bolintineanu, W. Michael Brown, Paul S. Crozier, Pieter J. in 't Veld, Axel Kohlmeyer, Stan G. Moore, Trung Dac Nguyen, et al. LAMMPS – A Flexible Simulation Tool for Particle-Based Materials Modeling at the Atomic, Meso, and Continuum Scales. *Computer Physics Communications*, 271:108171, 2022. doi: 10.1016/j.cpc.2021.108171.

Loup Verlet. Computer "Experiments" on Classical Fluids. I. Thermodynamical Properties of Lennard-Jones Molecules. *Physical Review*, 159:98–103, 1967. doi: 10.1103/PhysRev.159.98.

Arthur F. Voter. Parallel Replica Method for Dynamics of Infrequent Events. *Physical Review B*, 57: R13985–R13988, 1998. doi: 10.1103/PhysRevB.57.R13985.

Arthur F. Voter, Francesco Montalenti, and Timothy C. Germann. Extending the Time Scale in Atomistic Simulation of Materials. *Annual Review of Materials Research*, 32:321–346, 2002. doi: 10.1146/annurev.matsci.32.112601.141541.

Han Wang, Linfeng Zhang, Jiequn Han, and Weinan E. DeePMD-kit: A Deep Learning Package for Many-Body Potential Energy Representation and Molecular Dynamics. *Computer Physics Communications*, 228:178–184, 2018. doi: 10.1016/j.cpc.2018.03.016.

Logan Ward, Alexander Dunn, Alireza Faghaninia, Nils E. R. Zimmermann, Saurabh Bajaj, Qi Wang, Joseph Montoya, Jiming Chen, Kyle Bystrom, Maxwell Dylla, et al. Matminer: An Open Source Toolkit for Materials Data Mining. *Computational Materials Science*, 152:60–69, 2018. doi: 10.1016/j.commatsci.2018.05.018.

Tian Xie and Jeffrey C. Grossman. Crystal Graph Convolutional Neural Networks for an Accurate and Interpretable Prediction of Material Properties. *Physical Review Letters*, 120:145301, 2018. doi: 10.1103/PhysRevLett.120.145301.

Richard J. Zamora, Arthur F. Voter, Danny Perez, Nandakishore Santhi, Susan M. Mniszewski, Sunil Thulasidasan, and Stephan J. Eidenbenz. Discrete Event Performance Prediction of Speculatively Parallel Temperature-Accelerated Dynamics. *SIMULATION: Transactions of The Society for Modeling and Simulation International*, 92(12):1065–1086, 2016. doi: 10.1177/0037549716674806.

Linfeng Zhang, Jiequn Han, Han Wang, Roberto Car, and Weinan E. Deep Potential Molecular Dynamics: A Scalable Model with the Accuracy of Quantum Mechanics. *Physical Review Letters*, 120:143001, 2018. doi: 10.1103/PhysRevLett.120.143001.

Yusheng Zhao and Luke L. Daemen. Superionic conductivity in lithium-rich anti-perovskites. *Journal of the American Chemical Society*, 134(36):15042–15047, 2012. doi: 10.1021/ja305709z.

A TRAINING DETAILS AND PARAMETERS

Data / Teacher	Value	Student / Optim.	Value
Teacher	atom-vacancy KMC	**Student**	Mirror macro-step world model
Lattice	40^3 BCC	**Candidates**	$\mathcal{C}_k(X_t)$; 8 V seeds
Cu / V density	0.0134 / 0.0002	**Path input**	teacher path summary
Temperature	300 K	**Latents**	global 32×16; patch/path 64/32
Event support	2NN hops	**Epochs / batch**	120 / 32
Train / val	2000 / 400 segments per k	**Optimizer**	Adam, lr 10^{-4}
Split policy	independent envs	**Weight decay**	10^{-5}
Filtering	no-op skipped	**Expected / realized time**	1.0 / 0.25
Time targets	τ_{exp} main; τ_{real} aux	**Edit / reward / latent / proj.**	1.0 / 0.5 / 0.5 / 0.5
Aux. losses	proj. / reward / priors	**KL / prior edit / prior latent**	0.05 / 0.25 / 0.25
Compute	NVIDIA A100 GPUs	**Runtime**	35.5 min train; 23.3 / 47.1 s eval

Table 1: Training and evaluation configuration.

The reported training and evaluation runs use NVIDIA A100 GPUs. Timing values are taken from persisted run logs for the reported configuration.

B IMPLEMENTATION DETAILS BY COMPONENT

Figure 5 shows the full pipeline, and Table 2 consolidates the component-level inputs, outputs, and dimensions used throughout Section 4.

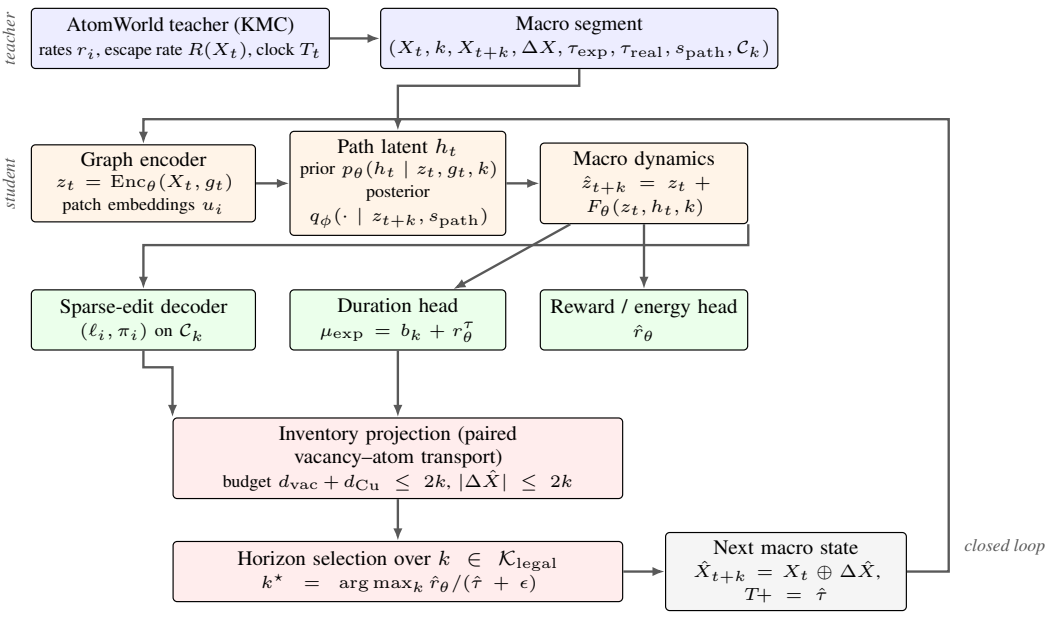


Figure 5: AtomWorld-Mirror architecture. The teacher supplies segments with reachable candidate supports, path summaries, and continuous-time labels. Mirror encodes the configuration and active patch, draws a horizon-conditioned path latent from the posterior during training and the prior during rollout, predicts the macro latent, and decodes a sparse edit, accumulated duration, and reward/energy quantity. Inventory projection precedes the downstream losses, so the projected state consumed during rollout is also the state used for optimization.

Component	**Input**	**Output**	**Rollout-time**
Graph encoder	X_t, global context g_t	global latent z_t (32×16), site embeddings u_i	yes
Candidate builder	X_t, 2NN event support, k	$\mathcal{C}_k(X_t)$ from 8 vacancy seeds	yes
Path posterior q_ϕ	z_t, z_{t+k}, s_{path}, k	path latent h_t (dim 32)	no (training only)
Path prior p_θ	z_t, g_t, k	path latent h_t (dim 32)	yes
Macro dynamics F_θ	z_t, h_t, k	residual to $\hat{z}_{t+k}$	yes
Edit decoder D_θ	u_i, patch context (dim 64), $\hat{z}_{t+k}$, h_t, k	change logit ℓ_i, type distribution π_i over {Fe, Cu, Vac}	yes
Inventory projection	raw edit, $\mathcal{C}_k$, budget $2k$	projection-closed $\Delta\hat{X}_{t:t+k}$	yes
Duration head	z_t, $\hat{z}_{t+k}$, h_t, g_t, edit context ρ_t, k	μ_{exp}, $\log\sigma_{\mathrm{exp}}$ on top of $b_k(X_t)$	yes
Realized-time head	same as duration head	μ_{real}, σ_{real} (lognormal)	yes
Reward head	z_t, $\hat{z}_{t+k}$, k	$\hat{r}_\theta$ (reward scale 10.0)	yes
Horizon selector	$\{\hat{r}_\theta, \hat{\tau}\}_{k\in\mathcal{K}_{\mathrm{legal}}}$	$k^\star$ by Eq. 24	yes

Table 2: Component-level implementation details. Rollout-time availability marks whether a component's inputs are computable from X_t and k alone.

C MACRO-STEP PSEUDOCODE

Algorithm 1 restates the rollout forward path of Section 4 as executable pseudocode. Colored comments mark the three stages the implementation separates: latent encoding of the configuration and active patches, per-horizon prediction followed by inventory projection, and horizon selection

with clock advance. Every quantity consumed inside the loop is computable from the current state and the horizon alone, the contract recorded in the rollout-time column of Table 2.

Algorithm 1 Mirror macro-step rollout over the critical evolution backbone. Colored comments mark encoding, projection-closed prediction, and horizon selection with clock advance.

Require: configuration X_0, clock T_0, trained Mirror θ, horizon set $\mathcal{K} = \{1, \dots, 1024\}$, candidate builder $\mathcal{C}_k(\cdot)$
Ensure: backbone trajectory $\mathcal{B}$

$t \leftarrow 0$; $m \leftarrow 0$; $X_t \leftarrow X_0$; $T \leftarrow T_0$; $\mathcal{B} \leftarrow [(X_0, T_0)]$
while the rollout budget remains **do**
  // encode $z_t \leftarrow \mathrm{Enc}_\theta(X_t, g_t)$; site embeddings u_i per site
  $\mathcal{K}_{\mathrm{legal}} \leftarrow \emptyset$
  for $k \in \mathcal{K}$ **do**
    $\mathcal{C}_k \leftarrow \mathrm{CandidateSupport}(X_t, k)$; construct active-patch context $c_{k,i}$
    $h_k \sim p_\theta(h \mid z_t, g_t, k)$; $\hat{z}_{t+k} \leftarrow z_t + F_\theta(z_t, h_k, k)$
    $(\ell_i, \pi_i) \leftarrow D_\theta(u_i, c_{k,i}, \hat{z}_{t+k}, h_k, k)$ over $i \in \mathcal{C}_k$
    $\Delta\hat{X}_k \leftarrow \mathrm{Decode}(\ell, \pi)$; *// inventory projection*
    $\Delta\hat{X}_k \leftarrow \mathrm{Project}(\Delta\hat{X}_k, \mathcal{C}_k, 2k)$ ▷ paired vacancy–atom transport
    $\hat{r}_k \leftarrow \mathrm{Reward}_\theta(z_t, \hat{z}_{t+k}, k)$; $\hat{\tau}_k \leftarrow \mathrm{Duration}_\theta(z_t, \hat{z}_{t+k}, h_k, g_t, k)$
    if $\Delta\hat{X}_k$ satisfies reachability and inventory **then** add k to $\mathcal{K}_{\mathrm{legal}}$ **else** discard k **end if**
  end for
  if $\mathcal{K}_{\mathrm{legal}} = \emptyset$ **then** record the stop reason and **break**
  // horizon selection $k^\star \leftarrow \arg\max_{k \in \mathcal{K}_{\mathrm{legal}}} \hat{r}_k / (\hat{\tau}_k + \epsilon)$
  $X_t \leftarrow X_t \oplus \Delta\hat{X}_{k^\star}$; $T \leftarrow T + \hat{\tau}_{k^\star}$; $t \leftarrow t + k^\star$; $m \leftarrow m + 1$; append (X_t, T) to $\mathcal{B}$
end while
return $\mathcal{B}$

D GENERALITY ACROSS MATERIAL SYSTEMS AND TEACHER TYPES

The four additional systems extend the three physical commitments across chemistry, bonding character, and teacher type. RPV steel keeps the atom-vacancy KMC teacher of Section 4.1 and raises the element count to 3 (Fe–Cu–Ni) and 10 (adding Mn, Cr, Al, Si, P, Co, Mo). Cu–Zr metallic glass and the Li_3N-based anti-perovskite solid electrolyte use molecular dynamics, with a macro step ending at a cage-swap topology event and its duration given by the elapsed MD time between consecutive events. The MD systems carry no vacancy species. A snap-to-lattice bijection places one atom at each reference site, conserving inventory by construction; species-level edits appear as paired swaps, the direct analog of KMC Cu–vacancy transport.

Each table reports the teacher as a reference alongside AtomWorld-Mirror. The decoder restricts edits to $\mathcal{C}_k$, and the projection conserves species counts. The time columns report log-MAE and the per-segment scale ratio of Eq. 25.

D.1 RPV STEEL, MULTI-COMPONENT (KMC TEACHER)

System	Model	violation% · F1	log-MAE · scale
RPV 3-element	KMC	0% · 1.00	0.00 · 1.00
	Mirror	**0% · 0.77**	**0.21 · 1.05**
RPV 10-element	KMC	0% · 1.00	0.00 · 1.00
	Mirror	**0% · 0.78**	**0.32 · 1.06**

Table 3: RPV steel: physical validity and time calibration at 3 and 10 elements.

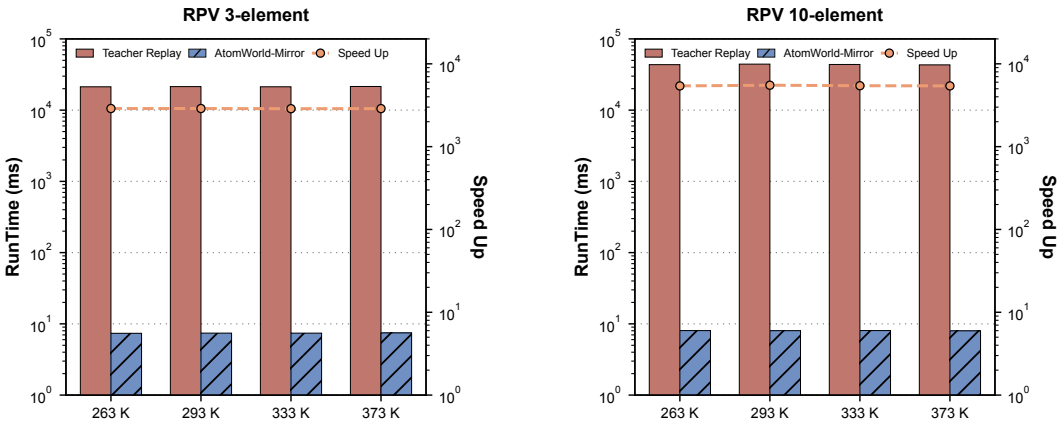


Figure 6: End-to-end Speed Up over the KMC teacher for the RPV steel systems at 263, 293, 333, and 373 K. Bars show runtime and the line shows Speed Up.

Increasing the RPV composition from 3 to 10 elements roughly doubles the Speed Up. Added chemistry raises the teacher's per-event cost, while macro-step cost remains tied to the candidate patch.

D.2 Cu–Zr metallic glass (MD teacher)

Model	violation% · F1	log-MAE · scale
MD teacher	0% · 1.00	0.00 · 1.00
Mirror	**0% · 0.84**	**0.26 · 1.10**

Table 4: Cu–Zr metallic glass: physical validity and time calibration.

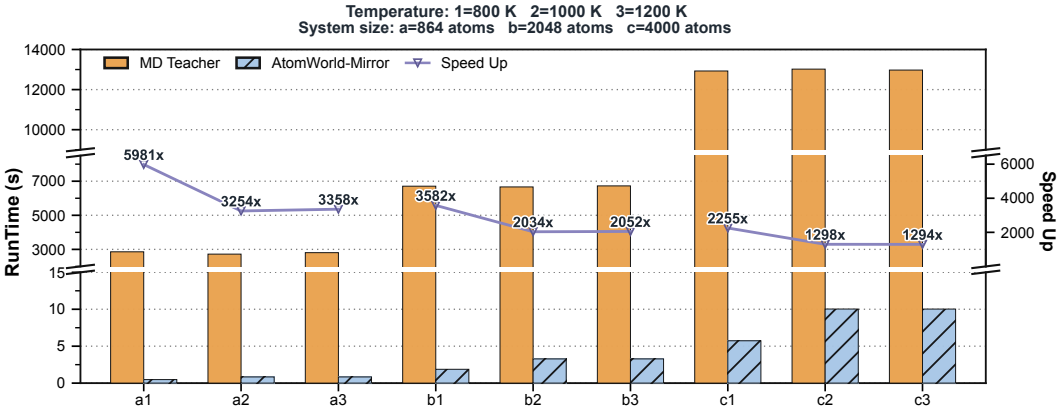


Figure 7: End-to-end Speed Up over the MD teacher for the Cu–Zr metallic glass across temperature and system size.

Cu–Zr provides the largest Speed Up in this study, with edit F1 0.84 and duration scale 1.10. The MD teacher resolves femtosecond integration steps between rare cage swaps, allowing each macro step to replace a long stretch of integration. Higher temperature increases cage-swap frequency and shortens the interval covered by one macro step.

D.3 Li_3N-BASED ANTI-PEROVSKITE SOLID ELECTROLYTE (MD TEACHER)

Model	violation% · F1	log-MAE · scale
MD teacher	0% · 1.00	0.00 · 1.00
Mirror	**0% · 0.71**	**0.27 · 1.07**

Table 5: Li_3N-based anti-perovskite solid electrolyte: physical validity and time calibration.

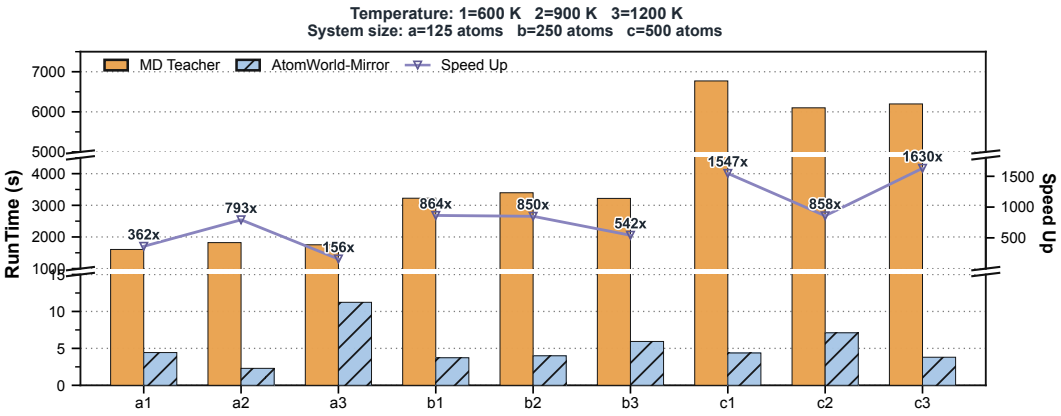


Figure 8: End-to-end Speed Up over the MD teacher for the Li_3N-based anti-perovskite solid electrolyte across temperature and system size.

Mirror reaches edit F1 0.71 on the electrolyte, with duration scale 1.07. Ionic migration produces well-separated topology events, keeping the duration target sharp.

E SEPARATING THE MACRO-STEP FORMULATION FROM LEARNED DYNAMICS

We isolate the macro-transition formulation and its three physical constraints through three matched comparisons.

The first comparison holds the latent backbone fixed while removing the macro-step structure. A Dreamer-style variant omits reachability, inventory, and continuous-time constraints. A direct endpoint regressor predicts X_{t+k} over the full lattice without candidate restriction, path latent, or duration head, reusing the main model's graph encoder for representation parity.

Model	Physically valid sparse edits	log-MAE · scale	ρ_T
Teacher KMC	violation 0% · F1 1.00	0.00 · 1.00	1.00
Dreamer-style (no constraints)	reach 45.2% · inv 59.2%	4.08 · 0.35	0.35
Endpoint regression	change-F1 0.42 · type-acc 0.45	N/A	N/A
AtomWorld-Mirror	**violation 0% · F1 0.91**	**0.13 · 1.00**	**1.01**

Table 6: Constraint and formulation baselines against AtomWorld-Mirror on the shared RPV steel aging held-out split (Cu 0.0134, 300 K, $k \in \{1, \ldots, 1024\}$). The table quantifies the contributions of reachability, inventory conservation, and duration modeling.

At matched computational cost, the constrained macro-step formulation delivers reachable edits, inventory preservation, and calibrated physical time. AtomWorld-Mirror matches the Dreamer-style variant within 0.5% in wall-clock time, raises reachable edits from 45.2% to 100%, and raises the

cumulative clock ratio from 0.35 to 1.01. AtomWorld-Mirror reaches change-F1 0.91 with path-conditioned duration modeling, compared with 0.42 for the endpoint regressor, which has no duration output.

F TEACHER-FORCED LONG-TRAJECTORY PROTOCOL

At segment m, from teacher state X_{t_m}, the model predicts a projected macro candidate. When several horizons are legal, the planner selects one using Eq. 24. The teacher advances the matching microscopic dynamics to provide the reference next state, reward, expected time, and realized time. The model contributes $(\Delta\hat{X}_m, \hat{\tau}_{\mathrm{exp},m}, \hat{r}_m)$, and the evaluator reports cumulative expected-time curves, selected-horizon histograms, stop reasons, and cumulative energy-error summaries. The teacher-forced diagnostic uses the teacher state as input at every segment, isolating duration-scale error, horizon-selection bias, and reward misalignment from autonomous rollout drift. The per-segment scale ratio of Eq. 25 is the time-calibration summary reported in the per-system tables of Appendix D. The cumulative clock ratio ρ_T of Eq. 26 is reported in Table 6 and Table 8.

G ADAPTIVE HORIZON SELECTION

At each macro step, the selector ranks legal horizons after projection. The resulting histogram records local configuration difficulty through selected horizons and projection stop events. When every candidate is rejected, the segment is recorded with its stop reason.

H MACRO-STEP INFERENCE OVER $k \in \{1, \ldots, 1024\}$

Macro-step inference cost is set by the candidate patch defined by the event neighborhood. Cost per macro step stays near-constant in the total number of atoms, and a checkpoint trained at 40^3 applies to larger configurations without retraining.

We measure the prior forward pass plus inventory projection on held-out active patches drawn from KMC trajectories over the operating range $k \in \{1, \ldots, 1024\}$. One macro step covers 480 teacher micro events on average and advances the physical clock by 12.98 seconds of accumulated expected time. Table 7 reports representative measurements within this range.

Horizon k	**128**	**256**	**512**	**1024**
Coverage	0.889	0.906	0.780	0.649

Aggregate measurement	**Value**
Samples / batch size	96 / 32
Wall time per pass	134.7 s
Throughput	0.713 segments/s
Mean micro events per macro step	480.0
Mean expected duration per macro step	12.98 s

Table 7: Measured macro-step inference within the operating range $k \in \{1, \ldots, 1024\}$ on NVIDIA A100 GPUs. The table reports representative horizons; coverage is the fraction of attempted segments that yield a usable macro sample after terminal and no-op filtering.

Longer horizons replace more micro-event replay per macro step. Coverage changes from 0.889 at $k = 128$ to 0.649 at $k = 1024$, making the efficiency-coverage relationship explicit.

I 54-BILLION-ATOM RPV STEEL AGING BENCHMARK

We evaluate a 50-year autonomous closed-loop rollout of Mirror on the full 5.4×10^{10}-atom RPV steel aging system at $1.34\,\mathrm{at.\%}$ Cu concentration across four temperatures. OpenKMC runs on 5.2

million cores of the Sunway TaihuLight supercomputer, whose many-core processors are based on SW26010, while Mirror runs on NVIDIA A100 GPUs. The benchmark records physical validity, cumulative time alignment, and end-to-end Speed Up against OpenKMC.

Model	Physically valid edits	Single-segment time	Cumulative clock
OpenKMC	violation 0% · F1 1.00	log-MAE 0.00 · scale 1.00	ρ_T 1.00
Mirror	**violation 0% · F1 0.91**	**log-MAE 0.27 · scale 1.17**	ρ_T **0.89**

Table 8: Physical validity and time alignment at the 5.4×10^{10}-atom scale in the RPV steel aging benchmark.

At the 5.4×10^{10}-atom scale, Mirror retains physically valid edits with F1 0.91 and reaches a cumulative clock ratio $\rho_T = 0.89$.

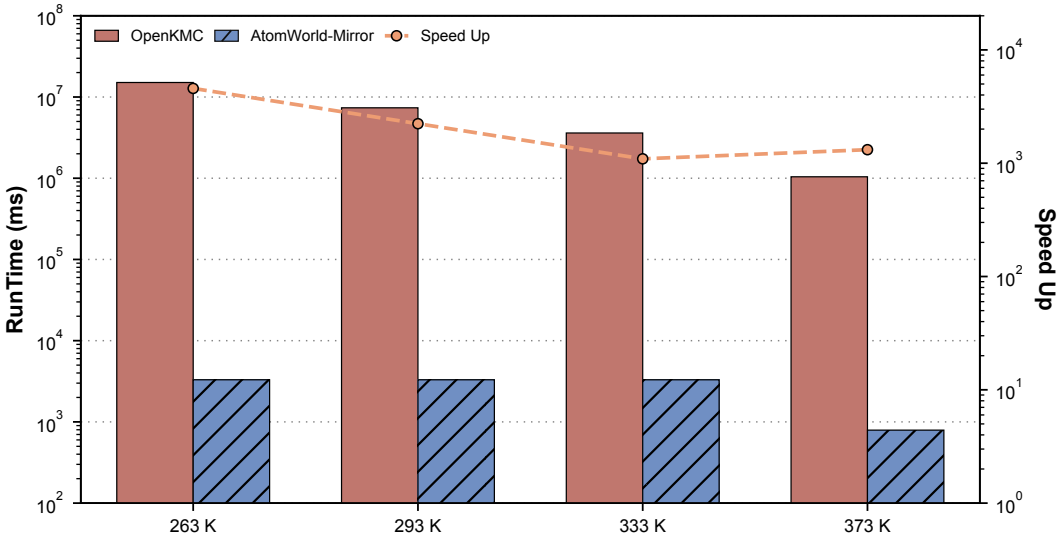


Figure 9: End-to-end Speed Up of Mirror over OpenKMC for the 5.4×10^{10}-atom RPV steel aging benchmark over a 50-year physical timescale. Bars show runtime and the line shows Speed Up.

J RPV STEEL AGING VISUALIZATION

Figure 10 shows Cu-cluster aging in a vacancy-mediated KMC trajectory on a 40^3 BCC lattice. The full boxes show the simulated lattice, and the partial boxes enlarge a representative sub-volume to resolve local cluster growth. Colors encode Cu-cluster sizes computed on the full periodic BCC lattice from the teacher configuration.

K LIMITATIONS AND FUTURE WORK

AtomWorld-Mirror targets long-horizon materials evolution in regimes where explicit microscopic replay consumes most of the computational budget before structurally consequential states are reached. The study evaluates this setting with controlled simulator teachers and establishes extensions to additional teacher types and materials systems. Future work will extend the framework to broader simulator-teacher combinations and optimize distributed deployment.

The five-system evaluation spans binary and multi-component RPV steel aging, Cu–Zr metallic glass, and Li_3N-based anti-perovskite solid electrolyte across KMC and MD teachers. The controlled RPV steel aging setting establishes the primary edit benchmark, and the MD systems demonstrate transfer

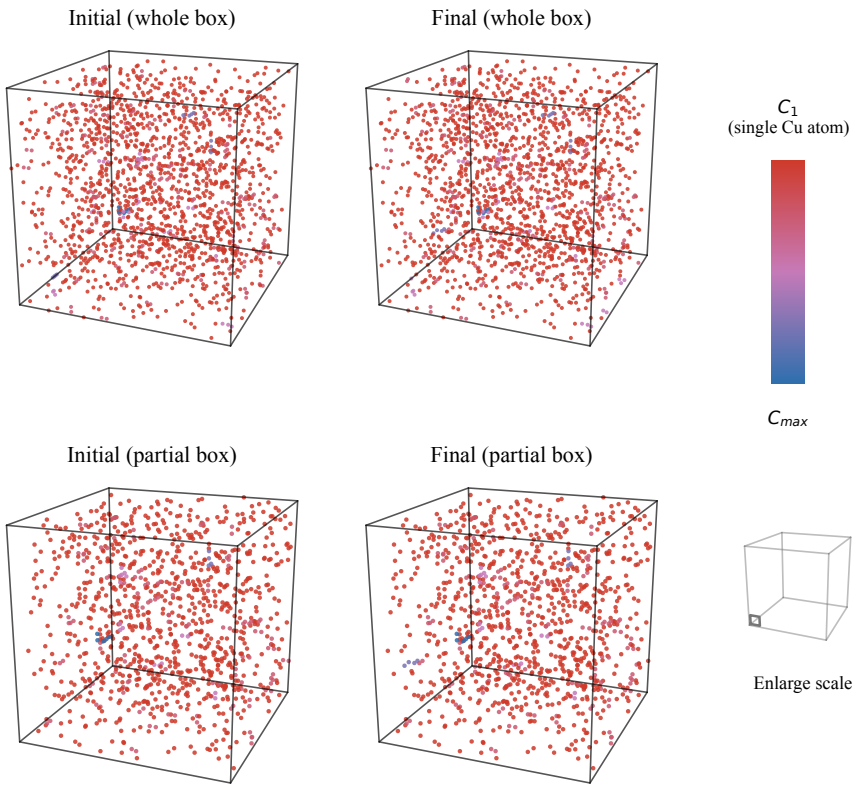


Figure 10: Cu-cluster visual evolution from initial to final states in an RPV steel aging simulation. Each point denotes a Cu atom in the teacher-simulated lattice. Colors use a shared red–pink-purple–blue scale for the actual 1NN/2NN-connected Cu cluster size C_i, with red denoting single-Cu clusters and blue denoting the maximum cluster size in this trajectory.

of the state-time interface across bonding character and teacher type. The current operating range reaches $k = 1024$. Extending candidate-support construction and conservation projection to systems with larger species inventories, together with topology criteria for continuously evolving structures, defines the next methodological step.

L EXPANDED RELATED-WORK DISCUSSION

Materials simulation spans several representations of physical evolution. Coarse-grained molecular dynamics, coarse-grained KMC, Monte Carlo, and Markov state models reduce atomic, event, or state resolution through coarse variables, aggregated events, or metastable-state representations (Marrink et al., 2007; Noid et al., 2008; Garnier & Nastar, 2013; Chodera & Noé, 2014). Classical molecular dynamics advances atomic configurations through short integration steps (Alder & Wainwright, 1959; Rahman, 1964; Verlet, 1967; Plimpton, 1995; Thompson et al., 2022). High-throughput materials databases (Jain et al., 2013; Curtarolo et al., 2012) and structure/property analysis toolchains (Ong et al., 2013; Ward et al., 2018; Choudhary et al., 2020) standardize structures, computed properties, and analysis workflows. These abstractions reduce computational cost. AtomWorld-Mirror adds a predictive backbone representation that selects future structural waypoints and assigns the accumulated physical duration of each sparse transition.

Learning-based materials simulation develops reusable models of energies, forces, local structures, and interacting-system dynamics (Merchant et al., 2023). Neural-network and Gaussian approximation potentials (Behler & Parrinello, 2007; Bartók et al., 2010), graph neural networks for molecules and materials (Schütt et al., 2017; Xie & Grossman, 2018; Batatia et al., 2022; Gilmer et al., 2017; Gasteiger et al., 2020), and equivariant force fields (Batzner et al., 2022; Musaelian et al., 2023) amortize local physical computations across atoms and molecules. Deep potential molecular dynamics and related packages turn these learned local models into scalable dynamical simulators (Zhang et al., 2018; Wang et al., 2018). Physics-informed learning (Raissi et al., 2019; Karniadakis et al., 2021) and graph-network simulators (Battaglia et al., 2016; Kipf et al., 2018; Pfaff et al., 2021) extend learned dynamics toward differential-equation, mesh-based, and interacting-system settings. AtomWorld-Mirror complements these approaches with a sparse macro state-time prediction target that explicitly represents backbone transitions and accumulated physical duration.

Step-Wise atomistic simulation provides the continuous-time teacher used in our experiments. KMC samples local events from physical rates and advances a residence-time clock (Bortz et al., 1975; Gillespie, 1977; Fichthorn & Weinberg, 1991), supplying local event semantics, reachable trajectories, and continuous-time labels (Chatterjee & Vlachos, 2007). High-performance KMC systems such as OpenKMC use engineered parallel execution to reach very large systems (Li et al., 2019). Established KMC accelerators include superbasin KMC and adaptive or self-learning KMC, which consolidate repeated events or construct transition mechanisms during simulation (Fichthorn & Lin, 2013; Henkelman & Jónsson, 2001). Parallel-replica, temperature-accelerated, rate-scaling, and policy-guided methods further alter event traversal or event selection (Voter, 1998; Sørensen & Voter, 2000; Zamora et al., 2016; Lin et al., 2019; Schulman et al., 2017; Tang et al., 2024; Bojesen, 2018). These methods operate on event-level traversal, event aggregation, or trajectory selection. AtomWorld-Mirror operates at the macro-transition level: the simulator teacher supplies the state-time reference, and Mirror learns sparse backbone transitions with accumulated physical duration. Speed Up follows from this predictive representation of atomistic dynamics.

Rate Scaling KMC reduces the sampling frequency of fast event channels by scaling their rates, allowing a fixed event budget to cover more of the slower evolution (Lin et al., 2019). Event selection and waiting times are then computed from the scaled rates, so the resulting clock follows the modified process; agreement with the original kinetics depends on the scaling regime and the observable being evaluated. This approach retains explicit event-level updates and targets the cost of repeatedly sampling fast dynamics. AtomWorld-Mirror changes the prediction unit to a reachable macro edit with a jointly predicted duration, amortizing the intervening event sequence into a learned state-time transition.

Table 9 organizes the comparison by the quantity changed and the physical clock preserved. Trajectory accelerators modify the accepted event sequence. Analytical accelerators retain event-level execution and aggregate fast events inside basins. Rate-scaling methods retain event-level execution while modifying fast-channel rates and their associated waiting-time distribution. AtomWorld-Mirror

predicts a sparse edit between backbone states, assigns its accumulated duration as a path-conditioned quantity, and enforces reachability and inventory at every macro step. Structural and temporal quantities remain explicit at the macro level.

Family	What it changes	Physical clock	Examples
Trajectory accelerators	the accepted event sequence	not directly traceable once event order is altered	RL/PPO policies for KMC, neural surrogate event samplers (Schulman et al., 2017; Tang et al., 2024; Bojesen, 2018)
Analytical accelerators	rates or grouping of fast events; event-level execution retained	preserved, every effective event still resolved	superbasin KMC, adaptive KMC (Fichthorn & Lin, 2013; Henkelman & Jónsson, 2001)
Rate-scaling accelerators	fast-event rates; event-level execution retained	sampled from scaled rates; fidelity to original kinetics depends on scaling regime	Rate Scaling KMC (Lin et al., 2019)
AtomWorld-Mirror	**the prediction unit: sparse macro edit between backbone states**	**predicted explicitly as path-conditioned accumulated duration**	**this work**

Table 9: Where AtomWorld-Mirror sits among KMC acceleration families.