

Article

Not peer-reviewed version

Directional Entropy Bands for Surface Characterization of Polymer Crystallization

[Elyar Tourani](#) , [Brian J. Edwards](#) * , [Bamin Khomami](#) *

Posted Date: 29 July 2025

doi: 10.20944/preprints202507.2422.v1

Keywords: polymer crystallization; directional entropy bands; local order parameters; nucleation; surface detection; molecular dynamics; phase transition



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a Creative Commons CC BY 4.0 license, which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Article

Directional Entropy Bands for Surface Characterization of Polymer Crystallization

Elyar Tourani ¹, Brian J. Edwards ^{1*} and Bamin Khomami ^{1,*}

Materials Research and Innovation Laboratory, Department of Chemical and Biomolecular Engineering, University of Tennessee, Knoxville, TN 37996, USA

* Correspondence: bedward1@utk.edu (B.J.E.); bkhomami@utk.edu (B.K.)

Abstract

Molecular dynamics (MD) simulations provide atomistic insights into nucleation and crystallization in polymers, yet interpreting their complex spatiotemporal data remains a challenge. Existing order parameters face limitations, such as failing to account for directional alignment or lacking sufficient spatial resolution, preventing them to accurately capture the anisotropic and heterogeneous characteristics of nucleation or the surface study of polymer crystallization. We introduce a novel set of local order parameters, namely, *directional entropy bands*, that extend scalar entropy-based descriptors by capturing first-order angular moments of the local entropy field around each particle. We compare these against conventional metrics (entropy, crystallinity index, and Smooth overlap of atomic positions (SOAP) descriptors) in equilibrium MD simulations of polymer crystallization. We show that (i) scalar entropy bands outperform SOAP in phase separation at single-snapshot resolution, and (ii) directional extensions (dipole projections and gradient estimates) robustly highlight the evolving crystal–melt interface, enabling earlier nucleation detection and quantitative surface profiling. UMAP embeddings of these 24–30D feature vectors reveal a continuous melt–surface–core manifold, confirmed by supervised boundary classification. Our approach is efficient and directly interpretable, offering a practical framework for studying polymer crystallization kinetics and surface growth phenomena.

Keywords: polymer crystallization; directional entropy bands; local order parameters; nucleation; surface detection; molecular dynamics; phase transition

1. Introduction

Crystallization in polymers is fundamentally more complex than in small molecules due to chain connectivity and topological constraints, which necessitate both conformational and spatial ordering. Upon cooling, polymers transition from coiled to semicrystalline states, yet a persistent amorphous phase often remains [1–5]. This transformation is driven by weak van der Waals interactions, whereas excluded volume effects hinder the entropic collapse into crystalline configurations [6]. Classical models such as Lauritzen-Hoffman (LH) [7] and Sadler-Gilmer (SD) [8] have described lamellar growth, but several aspects remain unresolved: the role of local structure in the melt prior to nucleation, the dynamic reentry of chains into crystalline regions, and how heterogeneous liquid environments influence growth pathways [9–11].

Molecular simulations have helped clarify non-equilibrium crystallization mechanisms in short [12–14] and entangled chains [15–20], often using local order parameters. These descriptors are designed to extract physically meaningful structure from disordered polymer configurations using rotation- and translation-invariant representations [21,22]. Angle-based alignment metrics, commonly applied to quantify segmental order, classify crystalline domains by comparing the principal axes of adjacent segments [12,14,23–25]. However, angle-based metrics lack universal thresholds and are undefined at chain ends, and spatial averaging, though useful for noise reduction, often obscures interfacial and precursor structures critical to nucleation and growth.

A range of translational and orientational order parameters (OPs) have been developed for identifying structural motifs in small-molecule systems, including Common Neighbor Analysis (CNA) [26], Bond Angle Analysis (BAA) [27], the Centrosymmetry Parameter (CNP) [28], Voronoi tessellation [29,30], and Steinhardt's Bond-Orientational Order (BOO) parameters [31]. Although many of these descriptors were originally designed to detect crystalline phases such as fcc, bcc, and icosahedral structures in atomic systems [21], select methods, particularly Voronoi and BOO, have shown promise for application to polymeric systems. Recently [], we systematically assessed these OPs in the context of polymer crystallization. Voronoi analysis adapts well to local density variations, but lacks information on orientational symmetry. In contrast, BOO parameters capture rotationally invariant local motifs using spherical harmonics [32] and have been used to study crystallization, glass transitions, and interface behavior in soft matter systems [33–36]. However, conventional BOO calculations are highly sensitive to how neighbors are defined, by distance, number, or topology, and can fail to resolve subtle structural distinctions, such as variations in Wyckoff positions [37]. Methods such as solid angle nearest neighbors (SANN) [38] and polyhedral template matching (PTM) [39] have addressed some of these limitations in small-molecule systems, but their effectiveness in polymers remains limited. In particular, the search for unique order parameters to identify and classify potentially short-lived precursor states has not yet been successful.

Spectrum-based descriptors, such as Smooth Overlap of Atomic Positions (SOAP) [40], Behler-Parrinello symmetry functions, and bispectrum coefficients, have extended the capabilities of traditional order parameters by mapping local environments into high-dimensional feature spaces. Unlike conventional BOO parameters, which summarize local symmetry using spherical harmonics, these descriptors retain richer structural information and enable advanced machine learning workflows [41–43]. To the best of our knowledge, this study presents the first systematic application and evaluation of SOAP descriptors in polymer crystallization, as their ability to capture chain connectivity, local packing anisotropy, and precursor states has not been previously explored.

Machine learning techniques have been increasingly used to extract structure-property relationships from high-dimensional data in a wide range of materials applications [22,44–47]. The Spectrum-based descriptors have been widely used in atomic systems to investigate crystallization, phase behavior, and self-assembly through unsupervised learning and clustering techniques, often combined with dimensionality reduction [48–50]. In particular, Spellings and Glotzer [41] introduced a fingerprint that avoids spherical harmonic averaging, using odd l values and variable neighborhood sizes to enhance rotational sensitivity. Adorf et al. [42] further integrated bispectrum features with BOO, bond angle, and distance metrics for manifold learning via HDBSCAN and classifier training. More recently, Bhardwaj et al. [51] applied self-supervised autoencoders to analyze conformational fingerprints constructed from shell-averaged angle-based OPs, enabling the detection of local order in entangled polymer systems.

Parallel to geometric and symmetry-based descriptors, local thermodynamic-like parameters, specifically entropy and enthalpy, have shown strong potential for identifying phase transitions at the particle level. Building on the early statistical mechanics work by Nettleton and Green [52], Piaggi et al. [53,54] introduced localized entropy and enthalpy measures derived from smoothed radial distribution functions, successfully distinguishing liquid and crystalline regions in atomic systems. Unlike ensemble-averaged thermodynamic quantities, these local estimates capture spatial variations relevant to nucleation. Nafar Sefiddashti et al. [19,55] extended this framework to polymer crystallization under elongational flow, demonstrating that configurational entropy could signal the onset of crystallization without requiring system-specific hyperparameters. However, challenges remain in quiescent systems, where radial distribution-based entropy may misclassify folded regions due to their proximity but lack of geometric symmetry. Although folds often appear entropically ordered, they lack orientational coherence and scalar entropy values cannot fully resolve this distinction.

Recognizing these limitations, in our previous study [56], we developed a machine learning framework that quantitatively assessed a broad set of geometric, entropic, and symmetry-based order

parameters in polymer crystallization. Our results showed that scalar entropy outperformed other descriptors in the detection of early nucleation events, but its effectiveness diminished in the resolution of the surface structure or the distinction between competing internal motifs. To address this, we introduced the crystallinity index (C-index) [56] to evaluate the OP sensitivity in the nucleation stages and identified entropy as one of the most informative indicators of the early stages. These insights motivate a new direction: moving beyond scalar thermodynamic variables toward spatially resolved entropy-based features that retain physical interpretation while capturing configurational anisotropy.

To overcome the limitations of scalar entropy and enhance local structural resolution, we introduce a new class of thermodynamic-like descriptors, namely, directional entropy bands. This approach extends the idea of local entropy by partitioning the radial space around each atom into concentric shells, or “bands”, and computing the entropy contributions within each band separately. These directional bands capture anisotropic variations in local configurational order and enable a more detailed distinction between folded segments, crystal surfaces, amorphous regions, and precursor structures. By combining radial decomposition with angular moments and dimensionality reduction, this method produces spatially resolved fingerprints that are both interpretable and physically grounded. Importantly, we demonstrate that simple averages of directional entropy bands, when combined with local enthalpy or spatial orientation metrics, can match or exceed the discriminative power of high-dimensional spectral descriptors like SOAP, particularly in early-stage crystallization.

The data summarize the performance of these descriptors against conventional OPs and spectrum-based ML fingerprints across nucleation regimes; (ii) benchmark the performance of these descriptors against conventional OPs and spectrum-based ML fingerprints across nucleation regimes, (iii) propose a simple yet effective averaging scheme that combines entropy bands with angular information to enhance the identification of anisotropic structures, such as folds and crystal edges, and (iv) provide a machine learning workflow that integrates unsupervised

The remainder of the paper is organized as follows. Section II presents the theoretical foundation and computational methodology, including the calculation of entropy bands, angular expansions, and ML workflows. Section III evaluates the performance of entropy bands through dimensionality reduction, clustering, and classification, comparing them with conventional and spectrum-based OPs. Section IV discusses the implications of directional entropy in identifying precursor structures and interfacial features. Finally, Section V summarizes the key findings and provides directions for future work.

2. Materials and Methods

2.1. Molecular Dynamics Simulation Details

Molecular dynamics simulations were performed using the LAMMPS package [57,58], employing the Siepmann-Karaboni-Smit (SKS) unified atom (UA) model [59] to represent polyethylene chains. In this model, each CH₂ group is treated as a single interaction site, while the CH₃ terminal groups serve as head and tail beads. To improve numerical stability and eliminate the need for rigid bond constraints in integration, the original constraint-based bonds were replaced with harmonic potentials [60–64]. Although some uncertainty remains about the ability of UA force fields to reproduce true crystal microstructures in the ground state of polyethylene [65], the SKS model remains well suited to study the dynamics of nucleation and crystallization [19,66].

Non-bonded interactions were modeled using a standard 12–6 Lennard-Jones (LJ) potential:

$$U_{LJ}(r_{ij}) = 4\epsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^6 \right], \quad (1)$$

where ϵ_{ij} is the potential well depth and σ_{ij} is the zero-crossing distance. The parameter values were set as $\epsilon/k_B = 47$ K and $\sigma = 3.93$ Å for the CH₂ and $\epsilon/k_B = 114$ K for the CH₃ groups. Interactions between different atom types were calculated using the Lorentz-Berthelot mixing rules: $\epsilon_{ij} = \sqrt{\epsilon_i \epsilon_j}$

and $\sigma_{ij} = (\sigma_i + \sigma_j)/2$. LJ interactions within chains were excluded for atoms separated by fewer than four bonds, and all LJ interactions were truncated at a cutoff of $2.5\sigma_{\text{CH}_2}$.

Bonded interactions include harmonic bond stretching:

$$U_{\text{str}}(l) = \frac{k_l}{2}(l - l_0)^2,$$

with equilibrium bond length $l_0 = 1.54 \text{ \AA}$ and force constant $k_l/k_B = 452,900 \text{ K/\AA}^2$; harmonic angle bending:

$$U_{\text{bend}}(\theta) = \frac{k_\theta}{2}(\theta - \theta_0)^2,$$

with equilibrium angle $\theta_0 = 114^\circ$ and $k_\theta/k_B = 62,500 \text{ K/rad}^2$; and torsional interactions defined by:

$$U_{\text{tor}}(\phi) = \sum_{m=0}^3 a_m \cos^m \phi,$$

with coefficients $a_0/k_B = 1010$, $a_1/k_B = -2019$, $a_2/k_B = 136.4$, and $a_3/k_B = 3165 \text{ K}$ [19,59,67].

Simulations were performed under isothermal and isobaric (NPT) conditions at 1 atm with orthorhombic periodic boundaries. Temperature and pressure were controlled using a Nose–Hoover thermostat and barostat. The system consisted of 60 polyethylene chains of n-pentacontahectane (C150), modeled at a quench temperature of 300 K, which corresponds to approximately 25% supercooling. This setup is consistent with previous nucleation studies for long-chain polymers [14,17]. Before quenching, the system was equilibrated at 550 K for 200 ns to ensure thermal homogeneity, then cooled to 300 K, where a single nucleation event was observed. The density evolved from 0.76 g/cm^3 at 450 K to 0.94 g/cm^3 at the end of crystallization.

2.2. Order Parameters

2.2.1. Smooth Overlap of Atomic Positions (SOAP) Descriptor

The SOAP descriptor is a density-based atomic environment characterization that has demonstrated considerable effectiveness in predicting crystalline properties [40,43] and classifying structures [68–70]. In a single-species system, specifically a united-atom potential energy model where each site represents an effective methylene group, the single-species SOAP method can be applied. This method describes the local atomic environment using a continuous density function, which is defined as

$$\rho(\vec{r}) = \sum_i \exp\left(-\frac{1}{2\sigma^2}|\vec{r} - \vec{R}_i|^2\right). \quad (2)$$

where \vec{R}_i represents the positions of the atoms and σ is the Gaussian width. The summation encompasses all atoms i within the local region of interest. To characterize the environment surrounding a selected center (often taken as $\vec{r} = 0$ in the atom of interest), one can expand $\rho(\vec{r})$ using a complete basis of radial functions and spherical harmonics as

$$\rho(\vec{r}) = \sum_{n=1}^{n_{\text{max}}} \sum_{l=0}^{l_{\text{max}}} \sum_{m=-l}^l c_{nlm} g_n(r) Y_{lm}(\theta, \phi), \quad (3)$$

where $g_n(r)$ represents a set of selected radial basis functions (such as orthonormal radial basis functions, Gaussian-type orbitals, or other suitable sets), $Y_{lm}(\theta, \phi)$ denotes the spherical harmonics and c_{nlm} are the expansion coefficients that encapsulate the geometric structure of the local atomic environment. The expansion coefficients c_{nlm} are calculated,

$$c_{nlm} = \int_V \rho(\vec{r}) g_n(r) Y_{lm}^*(\theta, \phi) dV. \quad (4)$$

This choice of basis is a matter of convention and is computationally more efficient but does not alter the final descriptor. The SOAP descriptor is derived by creating rotationally invariant combinations of the coefficients, commonly known as the power spectrum.

$$p_{nm'l} = \sum_{m=-l}^l c_{nlm}^* c_{n'l'm} . \quad (5)$$

The values of $p_{nm'l}$ then serve as the components of the SOAP descriptor for the local environment around the selected atom. This captures the essential geometry of the local atomic environment in a manner that is independent of the chosen coordinate frame. Consequently, the descriptor for each particle has dimensions of $n_{\max} \times n'_{\max} \times l_{\max}$.

To compute the SOAP vectors, we focused on the united-atom sites of the polymer chains. Based on the RDF plot of the orthorhombic reference crystal structure, a cutoff radius of 6.0 Å was selected (slightly beyond the third peak), for inclusion in the environment space of the neighbor list of each site. The SOAP parameters used in this study were chosen as $n_{\max} = 8$, $l_{\max} = 8$ to balance resolution and computational cost. The width of the Gaussian smearing functions, which controls the resolution of atomic position encoding, was set to 0.1 Å. All other parameters were kept at their default values, as implemented in the DDescribe library. [71].

2.2.2. Thermodynamic-Like Parameters for Phase Transitions

Local thermodynamic descriptors, specifically local enthalpy and configurational entropy, have proven effective in distinguishing between crystalline and amorphous environments on the atomic scale [52–54]. The local enthalpy for atom i is defined by distributing the potential energy and pressure volume contribution of the system into the atomic components:

$$H_i = U_i + \frac{pV}{N}, \quad (6)$$

where U_i is the potential energy of the atom i , p is the pressure, V is the volume of the system, and N is the total number of atoms. Then a neighbor-averaged local enthalpy is computed as:

$$\bar{H}_i = \frac{\sum_j H_j + H_i}{N_{\text{neigh}} + 1}. \quad (7)$$

The local entropy estimate is based on the two-body excess entropy term S_2 from Kirkwood's expansion of the configurational entropy, which captures up to 90% of the total entropy in dense fluids [53,72,73]:

$$S_2 = -2\pi\rho k_B \int_0^\infty [g(r) \ln g(r) - g(r) + 1] r^2 dr, \quad (8)$$

where ρ is the numbers density and $g(r)$ is the radial distribution function. For each atom i , the local entropy S_i is calculated using a smoothed atom-specific radial distribution function $g^i(r)$ truncated at a finite cutoff r_m :

$$S_i = -2\pi\rho_N k_B \int_0^{r_m} [g^i(r) \ln g^i(r) - g^i(r) + 1] r^2 dr, \quad (9)$$

where

$$g^i(r) = \frac{1}{4\pi\rho_N r^2} \sum_j \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(r - r_{ij})^2}{2\sigma^2}\right), \quad (10)$$

and r_{ij} is the distance between the atom i and the neighbor j , with the smoothing parameter σ . The neighborhood-averaged local entropy is expressed as:

$$\bar{S}_i = \frac{\sum_j S_j f(r_{ij}) + S_i}{\sum_j f(r_{ij}) + 1}, \quad (11)$$

where $f(r_{ij})$ is a switching function that decays from 1 to 0 as $r_{ij} \rightarrow r_c$ [19,54].

This scalar entropy framework has been widely applied to characterize polymer crystallization under flow conditions [19,74–79]. However, coarse scalar averages provide only isotropic and low-resolution information and are often insufficient in quiescent systems, where radial-distribution-based entropy may misclassify folded regions. These regions can appear entropically ordered due to local packing proximity, yet they lack geometric symmetry or orientational coherence.

To overcome this limitation, we introduce an array descriptor termed *entropy bands*, which computes the local entropy within concentric radial shells (Figure 1a). For each atom, the spherical environment is divided into fixed-thickness radial bands ΔR , and the entropy within each shell is independently averaged, resulting in an entropy fingerprint vector \vec{S}_i . Shell boundaries were selected on the basis of the RDF of a reference orthorhombic crystal structure, with a maximum cutoff of $R_{\max} = 10.0 \text{ \AA}$, previously validated as optimal for phase separation [19]. The fingerprint vector typically includes 6–8 bands, depending on the resolution ΔR , which we varied between 1.0 and 1.5 \AA . The value $\Delta R = 1.5 \text{ \AA}$ was selected for the final analysis due to its superior ability to distinguish folded and aligned structures, as illustrated in Figure 1b.

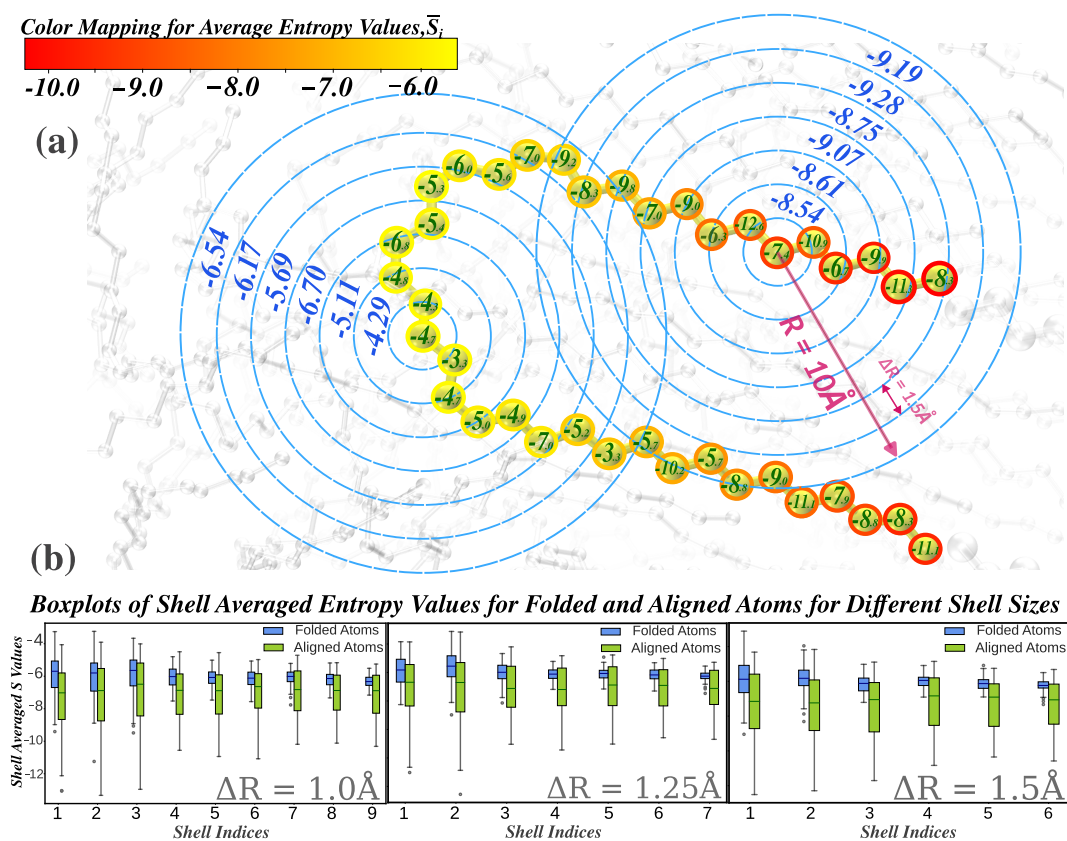


Figure 1. Definition of entropy bands: (a) Entropy fingerprint vectors for random particles, with a color bar indicating scalar average entropy. Particle labels show non-averaged entropy values, and shell labels represent entropy from distinct radial shells ($\Delta R = 1.5$); (b) Box plots of shell-averaged entropy values for randomly selected folded and aligned atoms across varying shell thicknesses (ΔR), highlighting the resolution impact of ΔR .

2.2.3. Directional Entropy Bands

Although entropy bands provide improved spatial resolution compared to scalar entropy, they still represent an isotropic average within each radial shell and are therefore limited in their ability to capture orientational features. This isotropy hinders their effectiveness in distinguishing anisotropic structures, such as chain folds, crystallite edges, and precursor motifs, which often exhibit direction-dependent ordering. To overcome this limitation, we introduce the *directional entropy bands*, an extension of the entropy band framework that incorporates angular information through geometric projections of local entropy values.

Rather than explicitly computing a directional radial distribution function $g(r, \theta, \phi)$, which is sparse and noisy at the atomic scale, we employ a more tractable and physically interpretable approach. For each atom i and the radial band $[r_1, r_2]$, the following quantities are calculated based on scalar entropy values of neighboring atoms:

- The **radial entropy average**:

$$\bar{S}_{i,r_1 \rightarrow r_2} = \frac{1}{N_{\text{shell}}} \sum_{j \in \text{shell}} S_j,$$

where S_j are the values of the scalar entropy of the neighboring atoms within the shell.

- The **directional entropy projections** along the Cartesian axes:

$$\langle S_j \cdot \hat{r}_{ij,x} \rangle, \quad \langle S_j \cdot \hat{r}_{ij,y} \rangle, \quad \langle S_j \cdot \hat{r}_{ij,z} \rangle,$$

where \hat{r}_{ij} is the unit displacement vector from the atom i to the neighbor j . These dipole-like projections reflect the first-order directional anisotropy of the local entropy field and are analogous to $l = 1$ spherical harmonic components.

- The **max-based entropy gradient** within the shell:

$$dS_{r_1 \rightarrow r_2}^{\text{max}} = \max_j \left(\frac{S_j - S_i}{\|\vec{r}_j - \vec{r}_i\|} \right),$$

which estimates the steepest local change in entropy and is particularly sensitive to interfacial regions and phase boundaries.

The full directional entropy fingerprint for each atom is constructed by concatenating the outputs from all radial shells, including: (i) scalar entropy averages, (ii) directional projections, and (iii) max-based gradients.

To further enhance the ability to identify surface atoms, those located at sharp boundaries between crystalline and amorphous regions, we define a feature of **surface weighted entropy gradient**:

$$f_i = \exp\left(-\frac{(\bar{S}_i - S^*)^2}{2\sigma_s^2}\right) \cdot \max_{r_1 \rightarrow r_2} dS_{r_1 \rightarrow r_2}^{\text{max}},$$

where \bar{S}_i is the mean scalar entropy of the atom i (imported or calculated), S^* is a characteristic surface entropy (typically around -5.8), and σ_s defines the width of the entropy window. This feature acts as a smooth, localized filter: it suppresses gradient contributions from atoms deep within crystalline or amorphous regions while enhancing those near interfacial zones. The scalar f_i has proven effective in isolating surface atoms even in systems with fluctuating entropy distributions and is used later in dimensionality reduction and classification analyses.

All calculations were performed using atomic coordinates and entropy values extracted from molecular dynamics snapshots in fixed time steps. Neighbor identification was handled using a periodic cKDTree implementation with minimum image convention. The radial shells ranged from 0.38 to 2.67 Å, divided into six bands based on the radial distribution function of the orthorhombic crystal reference. A Gaussian broadening parameter of $\sigma = 0.1$ Å was used in the estimation of entropy, consistent with the prior thermodynamic-like descriptors [19,54].

The resulting directional entropy bands, along with the surface-weighted gradient f_i , enable fine-grained resolution of structural anisotropy in polymeric systems. As demonstrated later, these descriptors provide a meaningful separation of crystalline, amorphous, and interfacial atoms in both unsupervised and supervised frameworks.

2.2.4. Machine Learning Workflow

A structured machine learning (ML) workflow was applied to rigorously evaluate the performance of directional entropy bands (DEB) and related order parameters in classifying crystalline, amorphous, and surface regions. The ML pipeline involved both unsupervised and supervised methods.

The high-dimensional features of the DEB and SOAP descriptors were separately projected onto low-dimensional manifolds using UMAP [80], with PCA [81] applied beforehand only when the SOAP descriptors were utilized. UMAP was chosen for its efficiency and ability to preserve local topological structure, which makes it well suited for interpreting structural phase organization. It was applied using a Manhattan distance metric, with 10 neighbors and a minimum distance of zero; detailed tuning procedures are described in our previous study [56]. HDBSCAN [82] was then used to identify structural phases and transition regions in the reduced space, with cluster size and sample thresholds set between 0.5–5% and 0.5–2% of the system, respectively, based on a validated parameter selection framework [56].

For supervised analysis, logistic regression and XGBoost classifiers were used, trained on labels assigned based on defined thermodynamic and structural criteria (crystal, melt, and surface). Model performance was rigorously evaluated using stratified cross-validation and metrics such as the Area Under the Receiver Operating Characteristic Curve (ROC-AUC). Forward Feature Selection (FFS) was systematically employed to identify the most informative DEB features for specific binary classification tasks, including surface versus melt and surface versus crystal distinctions. Further interpretation was achieved through SHapley Additive exPlanations (SHAP), allowing a detailed understanding of feature contributions.

Further details of ML implementation, parameter selection, and evaluation criteria are presented along with the results in Section 3.

3. Results and Discussion

To assess the effectiveness of structural order parameters in distinguishing amorphous, crystalline, and interfacial phases during polymer crystallization, we analyze both high-dimensional symmetry-based descriptors and physically interpretable thermodynamic metrics. We begin with the Smooth Overlap of Atomic Positions (SOAP), a geometric fingerprint that encodes local environments via rotationally invariant density projections. As a high-dimensional symmetry-aware descriptor, SOAP has demonstrated strong performance in capturing subtle structural motifs in both crystalline and disordered materials [22,40,43]. Given its sensitivity to local geometric arrangements, we hypothesize that SOAP may also capture interfacial complexity in polymers, particularly in regions where traditional order parameters fail to resolve transitional environments. In the following, after discussing the results of the SOAP method, we will evaluate entropy-based descriptors, beginning with scalar entropy bands and then we add the directional variants which integrate the local anisotropy.

Throughout this study, we focus more on four representative simulation time points to be consistent with the literature [51,56]: pre-nucleation (t_{pre}), transitional (t_{tr}), intermediate growth (t_{mid}), and steady-state crystallization (t_{ss}). These time points were defined in our previous work [56], based on the evolution of system density and cluster size.

3.1. Smooth Overlap of Atomic Positions (SOAP) Descriptors

To the best of our knowledge, the use of SOAP in polymeric systems and the findings of the ML analyses to be presented are novel contributions. Therefore, this work is a test application of SOAP descriptors to distinguish structural phases in a polymeric system. In the present case, each SOAP fingerprint vector has 324 components, generated using the parameters $r_{cut} = 6.0$, $n_{max} = 8$, $l_{max} = 8$, and $\sigma = 0.1$.

To assess how best to reduce these high-dimensional data, dimensionality reduction was performed at a representative intermediate time step (t_{mid}). In Figure 2, three approaches are compared: (a), (b), (c) PCA for 15 components followed by UMAP to 2D, (d), (e), (f) PCA for 25 components

followed by UMAP and (g), (h), (i) direct UMAP without PCA. For PCA-based reductions, the total variance retained was 95.59% with 15 components and 97.60% with 25 components (shown in panels (a) and (d), respectively). The retention of the variance for the PCA projections is summarized in Appendix Figure A2. In panel (a), the 2D embedding after PCA(15)+UMAP is shown. When colored by p_2 values (panel b), two major clusters emerge, albeit with considerable noise within the crystalline cluster that p_2 shows noncrystal particles as well. Interestingly, when plotted by the dihedral backbone conformation (panel c), the entire orange cluster consists of trans conformations, whereas the blue cluster includes both trans and gauche (with approximately 64% trans and 31% gauche).

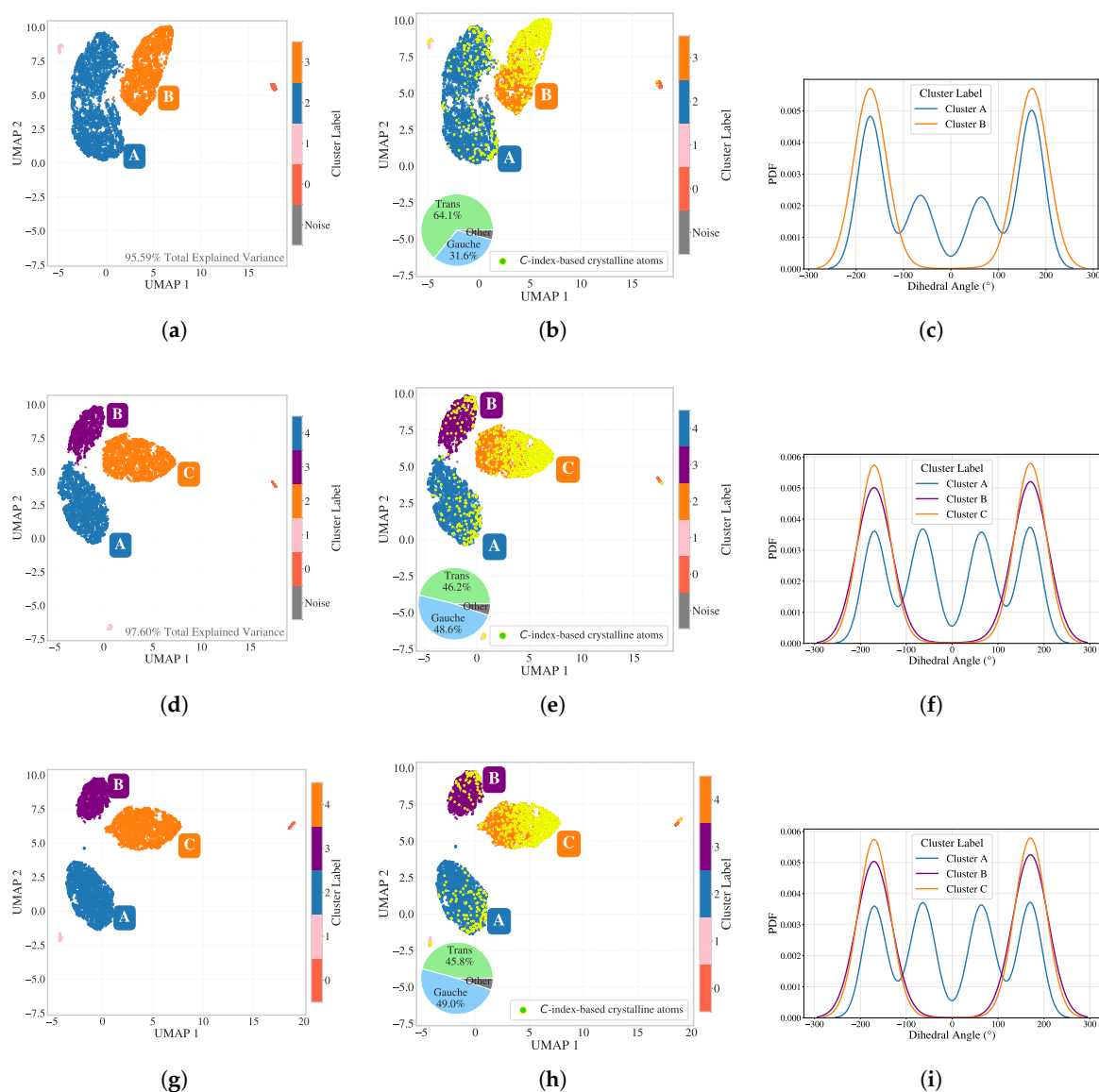


Figure 2. Comparison of SOAP dimensionality reduction pipelines at t_{mid} . **(a)–(c)** PCA to 15 components + UMAP + HDBSCAN: (a) 2D embedding with 95.59% explained the retained variance; (b) colored by p_2 ; (c) Dihedral distribution for different clusters. Dihedral proportions for cluster A (64% trans, 31% gauche) are shown in inset pie chart of panel (b). **(d)–(f)** PCA to 25 components + UMAP + HDBSCAN: (d) embedding with 97.60% explained the retained variance; (e) colored by p_2 ; (f) Dihedral distribution for different clusters with dihedral proportions (46% trans, 49% gauche) shown in inset pie chart in panel (e). **(g)–(i)** Direct UMAP + HDBSCAN: cleanest separation and consistent dihedral interpretation.

Panel (d) shows PCA(25)+UMAP, where the embedding reveals three clusters. In panel (e), the classification p_2 highlights the crystalline cluster, now more distinct from the rest. The dihedral

coloring (f) reveals that both the orange and purple clusters correspond to trans conformations, while the blue cluster retains mixed conformations with around 46% trans and 49% gauche.

Panel (g) presents the direct UMAP result, which shows even greater separation among the clusters. Again, panel (h) highlights the p_2 -classified crystalline particles concentrated in the orange cluster. Cluster A includes 46% trans and 49% gauche dihedral configurations. Panel (i) confirms the trans and gauche conformation separation. Given cleaner separation and the potential to retain subtle features, the direct UMAP pipeline was selected for SOAP dimensionality reduction. Since the resulting UMAP embeddings showed well-separated clusters, it was not necessary to perform a detailed hyperparameter benchmark for HDBSCAN in this case.

Pie charts were added next to cluster A in the embeddings to display the proportion of trans versus gauche conformations. And small peripheral clusters across all UMAP embeddings were found to correspond to chain-end particles or those immediately adjacent to them.

All analyses were performed on per-particle SOAP descriptors computed in the quiescent state without alignment or centering, ensuring consistent treatment of local chain geometry across the system. Researchers adopting different SOAP parameterizations that lead to even higher-dimensional feature spaces are advised to repeat this PCA+UMAP comparison. UMAP performance may be degraded in extremely high dimensions [80], and overaggressive PCA may ignore informative variance.

To investigate the temporal stability of SOAP-based classifications, we extended each particle's descriptor by concatenating SOAP vectors across 20 consecutive time steps centered around t_{mid} , ranging from $t_{mid} - 10\Delta t$ to $t_{mid} + 10\Delta t$. This resulted in high-dimensional fingerprints of size $324 \times 20 = 6480$ per particle. Dimensionality reduction was performed using PCA (retaining 25 components) followed by UMAP to embed the data into two dimensions, which was subsequently clustered using HDBSCAN (Figure 3a). The resulting clusters aligned closely with those classified as crystalline by the scalar parameter p_2 (panel b), although the stacked temporal representation precludes the analysis of instantaneous conformational states such as dihedrals.

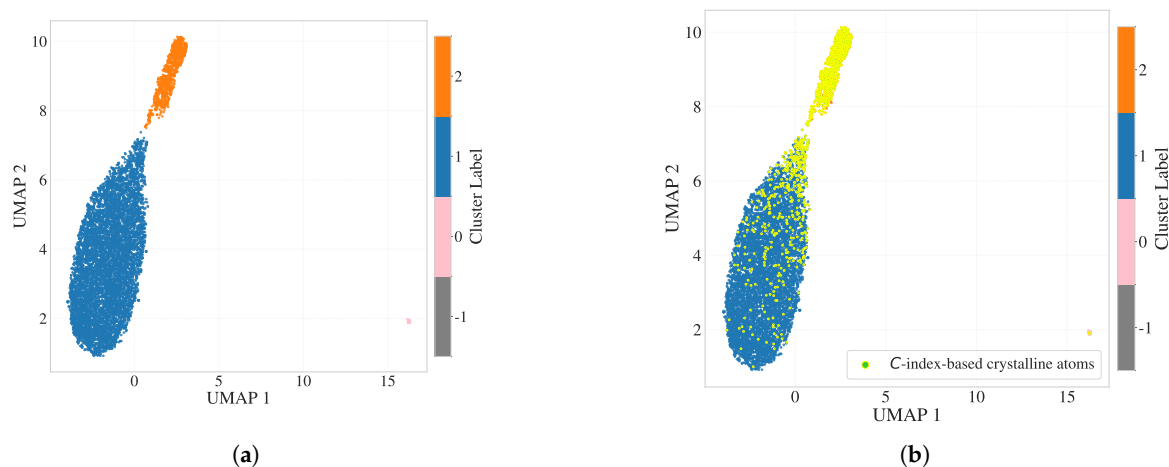


Figure 3. Dimensionality reduction and clustering of SOAP descriptors incorporating 20 consecutive timesteps around t_{mid} . **(a)** 2D UMAP embedding of PCA(25)-reduced SOAP vectors (6480 dimensions per particle) with HDBSCAN clustering. **(b)** Particles are colored based on p_2 -based crystallinity classification. The crystalline cluster closely aligns with the cluster identified by HDBSCAN, as the yellow-green points correspond with the orange points identified through clustering in the 2D embedding. However, instantaneous dihedral information was not preserved in this analysis due to the concatenation of temporal data.

A similar strategy was applied in our previous study for bond orientation order (BOO) descriptors using 5, 10 and 20 time-step windows, where longer temporal windows led to progressively improved separability between phases [56]. This finding echoes the observations by Adorf et al. [42], who demonstrated that incorporating time into high-dimensional spherical harmonic descriptors significantly improves clustering resolution.

However, despite its benefits, we do not adopt temporally stacked descriptors in the remainder of this study, as the goal here is to study OPs that achieve high performance on the same-time structures so that we can move towards creating better and parsimonious models of high temporal resolution dynamics of structures, e.g. nucleation and surface growth. Simple concatenation of features over time and use of opaque models would not help in this regard.

3.2. Directional Entropy Bands: Resolving Local Configurational Order

3.2.1. Manifold Learning and Comparison of Band-Averaged and Scalar Entropy

To investigate whether entropy bands provide enhanced structural resolution over scalar entropy, we applied UMAP to the entropy band vectors combined with local enthalpy (see Appendix A.1 for a comparison of feature scaling methods, including skewness correction of thermodynamic-like parameters prior to dimensionality reduction). Figure 4 presents the resulting embeddings in three representative crystallization stages: transitional (t_{tr}), intermediate (t_{mid}) and steady state (t_{ss}). Panels (a)–(c) show the UMAP projections colored by the scalar entropy threshold ($\bar{S}_i^* = -5.8$). These embeddings highlight the absence of sharp phase separation, especially as crystallization proceeds (panel c). Interestingly, many atoms classified as crystalline by scalar entropy appear in amorphous regions of the entropy band manifold — most notably in folded or bridged regions — suggesting that scalar entropy fails to capture key configurational nuances. The red circles in panel (b) mark such discrepancies at t_{mid} .

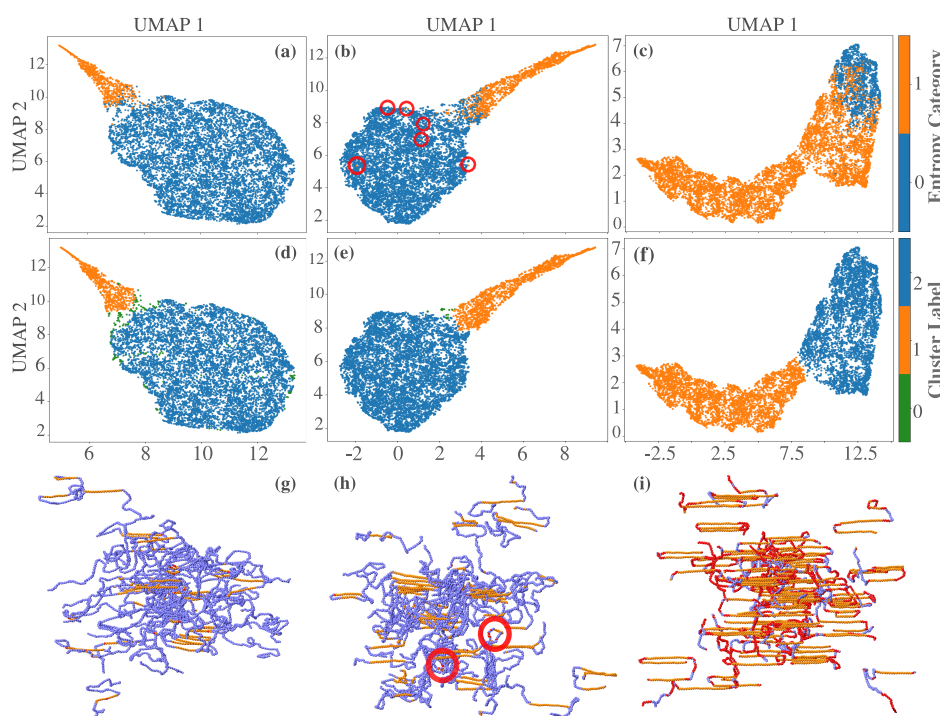


Figure 4. UMAP analysis of shell-based entropy vectors (*entropy bands*) combined with enthalpy values for three key crystallization stages. (a)–(c) UMAP embeddings for t_{tr} , t_{mid} , and t_{ss} , respectively, colored by scalar entropy. Red circles in (b) highlight discrepancies where scalar entropy classifies particles as crystalline, yet their embedding position suggests otherwise. (d)–(f) HDBSCAN clustering applied to the UMAP embeddings of entropy bands, identifying two major structural populations. (g)–(i) Structural reconstructions of the clusters at t_{tr} , t_{mid} , and t_{ss} , respectively. Misclassified particles from panel (b) appear in folded or bridged segments in panel (h), demonstrating the improved sensitivity of entropy bands in capturing configurational complexity beyond scalar entropy metrics. In panel (i), red particles represent disordered regions—primarily folded segments—that are incorrectly classified as crystalline by scalar entropy (panel c), but correctly identified as amorphous by scalar entropy bands. Orange particles correspond to atoms consistently classified as crystalline by both methods.

To further investigate these structural groupings, we applied HDBSCAN to the UMAP representations (panels d–f). The resulting clusters broadly align with amorphous and crystalline regions, and boundary overlaps diminish as crystallization matures. Compared to scalar entropy, these clusters offer a more refined partitioning of the phase space, particularly in regions involving interfacial or partially ordered configurations.

The spatial distribution of these clusters is visualized in panels (g)–(i), showing particle-level reconstructions at t_{tr} , t_{mid} and t_{ss} , respectively. In the early stages (panel g), the differences in clustering are minimal. By t_{mid} (panel h), the entropy bands correctly classify folded-chain atoms misidentified by scalar entropy. In steady state (panel i), the entropy bands robustly distinguish the folded, bridged, and chain-end segments, underscoring their improved sensitivity to configurational heterogeneity.

While $\bar{S}_i^{(b)}$ represents a spatial average within concentric radial shells and \bar{S}_i captures total local entropy, the two differ in their ability to resolve phase interfaces. Band-averaged entropy demonstrates improved separability between amorphous and crystalline regions, as evidenced by UMAP manifolds and HDBSCAN clustering results. However, this separation remains incomplete: the manifold does not exhibit fully disjoint clusters, and HDBSCAN classification near phase boundaries is sensitive to hyperparameters and density fluctuations. These ambiguities particularly affect interfacial or partially ordered atoms, where scalar descriptors, even band-averaged ones, struggle to provide consistent phase assignments.

To address this limitation, we now turn to the full directional entropy vector $\vec{S}_i^{(b)}$, which retains first-order angular information about the local entropy field. This vectorized representation is sensitive to configurational anisotropy and directional gradients, key signatures of folded structures, interfacial asymmetries, and structural motifs that are often overlooked by scalar metrics. In the next subsection, we examine whether $\vec{S}_i^{(b)}$ offers improved phase discrimination and interface resolution, particularly compared to the crystallinity index.

3.2.2. Comparison of Directional Entropy Bands (DEB), $\vec{S}_i^{(b)}$, with the Crystallinity Index (C-Index) and Phase Boundaries

To evaluate the spatial and phase resolution capacity of the directional entropy bands, we applied UMAP to feature sets that comprise the full directional entropy vector $\vec{S}_i^{(b)}$. These features include the band-averaged entropy, directional projections ($\bar{S}_{ix}^{(b)}$, $\bar{S}_{iy}^{(b)}$, $\bar{S}_{iz}^{(b)}$), the normalized angular moment ($\bar{S}_{im}^{(b)}$), and the surface-weighted entropy gradient f_i . Full details of the structure of the feature are provided in Section 2.2.3.

Figure 5 shows the embeddings of $\vec{S}_i^{(b)}$ in the same three key crystallization stages: transitional (t_{tr}), intermediate (t_{mid}), and steady state (t_{ss}). Panels (a)–(c) display UMAP projections colored by the scalar entropy (\bar{S}_i). These embeddings reveal smooth, though not sharply delineated, phase separation, particularly at later crystallization stages (panel c), where increasing interfacial heterogeneity becomes prominent. This outcome is intentional, as the set of characteristics includes the surface-weighted entropy gradient f_i , which improves sensitivity to transitional and boundary atoms. Excluding f_i leads to sharper melt–crystal separation in the manifold, but at the expense of interface resolution, which is essential to characterize crystal growth dynamics.

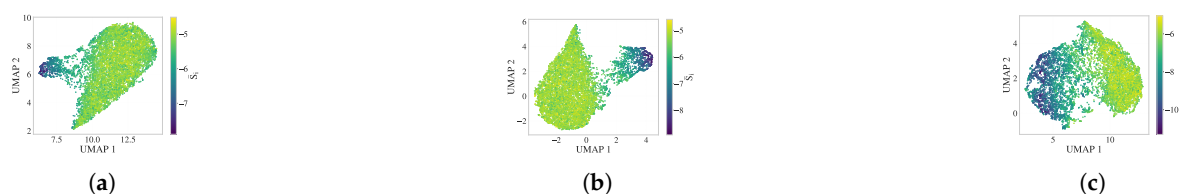


Figure 5. UMAP embeddings of the directional entropy vector $\vec{S}_i^{(b)}$ at three crystallization stages: (a) t_{tr} , (b) t_{mid} , and (c) t_{ss} , colored by scalar entropy \bar{S}_i . The embeddings reveal smooth phase separation and reflecting the influence of interfacial and directional features.

The surface-weighted entropy gradient parameter f_i plays a critical role in enhancing the sensitivity to interfacial structure. In this study, we selected $\sigma_s = 1.5$ for f_i as a representative value that balances resolution and stability at different stages of crystallization. Qualitatively similar results are obtained for σ_s values in the range of 0.1 to 2.0, with minor changes in the position of the inferred interface. Specifically, increasing σ_s tends to bias the interfacial boundary slightly towards the crystalline side, reflecting a stronger surface weighting. This choice ultimately depends on the desired trade-off between strict geometric separation and the inclusion of interfacial disorder. A comparison of UMAP embeddings and C-index disagreement maps across the range $\sigma_s \in [0.1, 2.0]$ is provided in Appendix Figure A4.

In general, directional components and f_i together introduce sufficient angular and boundary sensitivity to delineate physically meaningful subpopulations - corresponding to melt, crystalline cores, and structurally complex interface regions - that are not easily distinguishable using scalar entropy alone or any other order parameter.

To further assess the effectiveness of $\vec{S}_i^{(b)}$ in resolving local phase identity, we focus on the intermediate crystallization stage (t_{mid}), where phase boundaries are the most structurally uncertain. Figure 6 compares the directional entropy vectors with the C-index in four panels.

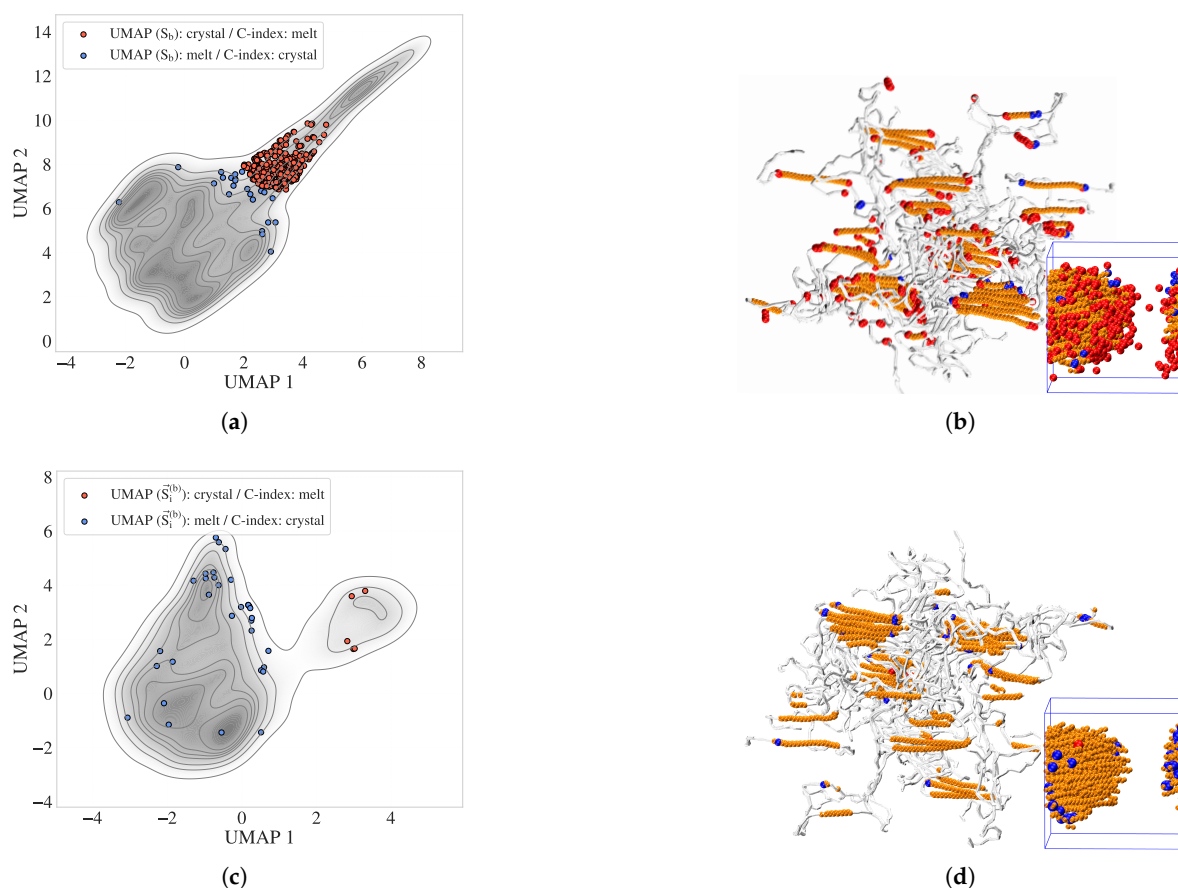


Figure 6. Comparison of scalar entropy bands and directional entropy vectors with the Crystallinity Index (C-index) at t_{mid} . **(a)** Contour map of scalar entropy bands ($S_i^{(b)}$), with discordant atoms highlighted: red atoms are classified as crystalline by entropy bands but amorphous by the C-index, and blue atoms vice versa. **(b)** VMD snapshot of the same atoms from (a), showing their spatial distribution. Orange atoms represent crystalline regions where both methods agree; red and blue atoms lie predominantly at interfaces and folded chain segments. **(c)** UMAP embedding of the directional entropy vector $\vec{S}_i^{(b)}$, showing reduced phase disagreement relative to (a). **(d)** VMD visualization of (c), confirming improved agreement with local structure and enhanced resolution of interfacial atoms.

Panel (a) presents a map of the scalar entropy band ($\bar{S}_i^{(b)}$), with atoms highlighted in red or blue to indicate disagreement with the C index: red atoms are labeled crystalline by the entropy bands but amorphous by the C index, while blue atoms show the opposite discrepancy. (b) Maps these discordant atoms onto a VMD [83]-rendered snapshot of the polymer configuration. Orange atoms correspond to regions where both descriptors agree on crystallinity, while red and blue atoms lie predominantly at interfaces or folded chain regions. Panel (c) shows a contour plot of the manifold built from $\bar{S}_i^{(b)}$, colored by discordance with the C-index. Compared with panel (a), the number of conflicting atoms is significantly reduced. (d) Improved consistency between DEB predictions and local structure from C-index results, with fewer discordant atoms (red/blue) on the surface of the crystalline regions.

These comparisons illustrate that directional entropy vectors align better with the physical morphology and reduce classification ambiguity at phase boundaries.

To further validate the ability of the directional entropy bands to resolve the interfacial structure, we adopted a geometric surface reference ('silver standard') derived from alpha shapes constructed on the C-index-defined crystalline cluster. This approach identifies atoms located on the geometric boundary of the cluster at t_{mid} , independent of entropy-based criteria.

We examined the use of alpha shapes to define geometric boundaries on the C-index-defined crystalline cluster. Based on Appendix Figure A3, a suitable range for the alpha parameter is $\alpha \in [0.3, 0.7]$, where the balance between over-smoothing (too small α) and more sharp surface (too large α) is optimal. For the results shown in this section, we selected $\alpha = 0.7$ to include a larger number of atoms on the crystal–melt boundary as surface atoms. This choice facilitates both a richer visual interpretation of interfacial behavior in UMAP space and a broader set of labels for subsequent classification tasks.

Figure 7 presents a complementary comparison in both feature space and real space. Panel (a) shows the UMAP embedding of $\bar{S}_i^{(b)}$, where the silver standard surface atoms are colored red. These atoms form a distinct transitional group located between melt-like and crystalline clusters in the feature space, consistent with their physical interfacial role. (b) VMD rendering of the crystalline cluster (cyan), overlaid with red atoms denoting the alpha shape surface and blue atoms denoting additional DEB-inferred surface atoms to the alpha shape. The DEB-identified surface population closely matches or slightly extends the alpha-shaped boundary, suggesting that it captures both geometric and entropic signatures of the crystal–melt interface.

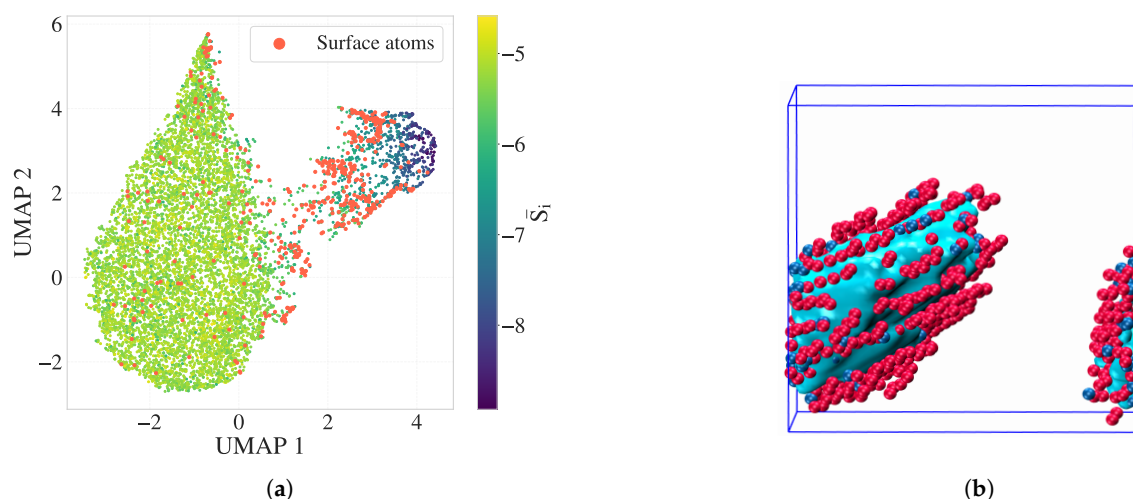


Figure 7. Surface atom identification using the C-index alpha shape (silver standard) and directional entropy bands at t_{mid} . **(a)** UMAP embedding of $\bar{S}_i^{(b)}$, with surface atoms highlighted in red. These atoms form a transitional group between melt-like and crystalline clusters. **(b)** VMD snapshot of the crystalline cluster (cyan), overlaid with silver-standard surface atoms (red) and DEB-inferred surface atoms (blue). The strong correspondence illustrates the interface-resolving capability of directional entropy bands.

This correspondence highlights the potential of directional entropy bands to encode interfacial structure with high fidelity, capturing both geometric and entropic signatures of surface atoms. These

findings motivate a deeper investigation into whether such features can be leveraged in supervised modeling tasks to systematically identify surface environments.

3.3. Model Explanation via Supervised Classification

3.3.1. Motivation and Approach

To further explore the structural information encoded in directional entropy bands, we adopt a supervised classification approach, not to demonstrate the superiority of any particular ML model, but to assess whether entropy-based features alone can reproduce known phase assignments and provide interpretable insights into interfacial structure. As target labels, we use the C-index derived in our previous work [56], which combines clustering results from a broad set of conventional order parameters.

Our objective is to determine whether the features of the directional entropy band, derived solely from differential entropy estimates, are sufficiently expressive to learn phase distinctions and surface localization in a physically interpretable way. In particular, our aim is to: Fit supervised models to establish the predictive power of entropy bands without relying on geometric or symmetry-based descriptors, Examine label uncertainty near interfacial regions to assess robustness of entropy-based boundaries, Identify the most informative subset of components of the entropy band by analyzing the importance of characteristics.

This framework allows us to assess the potential of entropy-derived descriptors to serve as early, physically grounded indicators of interfacial structure, complementing or even anticipating more complex symmetry-based features. In particular, in our previous work [56], we showed that scalar entropy outperforms traditional order parameters in detecting early stage crystallization in entangled polymer chains. Building on that result, the present framework investigates whether directional entropy band features, enhanced with interfacial sensitivity, can improve both interpretation and predictive performance in resolving phase boundaries and surface environments.

3.3.2. Data Preparation

The supervised classification task used band descriptors of directional entropy, including band-averaged entropy ($\bar{S}_i^{(b)}$), directional projections ($\bar{S}_{ix}^{(b)}$, $\bar{S}_{iy}^{(b)}$, $\bar{S}_{iz}^{(b)}$), max-based gradient estimate ($\bar{S}_{im}^{(b)}$), and the surface-weighted entropy gradient (f_i). All continuous features were standardized to zero mean and unit variance for the random forest model. The data set was divided into training subsets (80%) and testing subsets (20%) using stratified sampling to maintain class balance.

3.3.3. Model Selection and Training

We evaluated both a Random Forest (RF) classifier and a Logistic Regression (LogReg) model to assess the separability of surface atoms based solely on entropy-derived descriptors. The RF model, known for capturing nonlinear feature interactions and robustness to colinearity, was used as a high-capacity baseline. The hyperparameters (number of estimators and maximum tree depth) were tuned through 5-fold stratified cross-validation using grid search to maximize the mean area under the receiver operating characteristic curve (AUC).

To complement this, we also trained a Logistic Regression model, which provides a linear and additive mapping between input features and predicted probabilities. This model allows for direct interpretation of the feature weights, making it especially suited to identifying combinations of entropy band components that align with a physically meaningful interfacial structure. Our goal here is not to optimize predictive performance per se, but to evaluate whether a linear combination of entropy-based features is sufficient to distinguish surface environments, a valuable step toward building mechanistic understanding of nucleation and growth processes.

The surface atom labels used for classification were generated using alpha shapes constructed in the C-index-defined crystalline cluster, as described in Section 3.2.2. Although alpha-shape surfaces are geometric proxies and not exact thermodynamic boundaries, they serve as a useful silver standard for assessing the extent to which directional entropy bands capture interfacial structure.

3.3.4. Performance Evaluation: ROC Analysis

Figure 8 presents the ROC curves for the surface class using both the RF and LogReg models trained on three sets of characteristics: p_2 , the set of conventional scalar OPs (excluding entropy) and the proposed DEB descriptors. Despite the silver-standard ground truth labels (i.e., the C index clusters) being derived using the OPs set (including scalar entropy), DEB descriptors alone achieve comparable or even superior predictive performance. For surface classification, DEB reaches AUCs of 0.92 (LogReg) and 0.96 (RF), outperforming p_2 (0.87 and 0.84) and closely matching the OPs group (0.92 and 0.97), which aggregates more variables.

These results underscore a key contribution of this work: DEB descriptors, built solely from differential entropy without relying on high-order symmetry terms, are not only physically interpretable but also highly predictive of interfacial environments. The fact that DEB matches the performance of a broader OP set, even though the ground-truth clustering was derived from those OPs, highlights its robustness. Furthermore, in light of our previous findings that entropy is a sensitive early stage indicator of nucleation, these entropy-based features may offer a more timely and mechanistically meaningful route toward understanding crystallization front propagation and interfacial dynamics in polymeric systems.

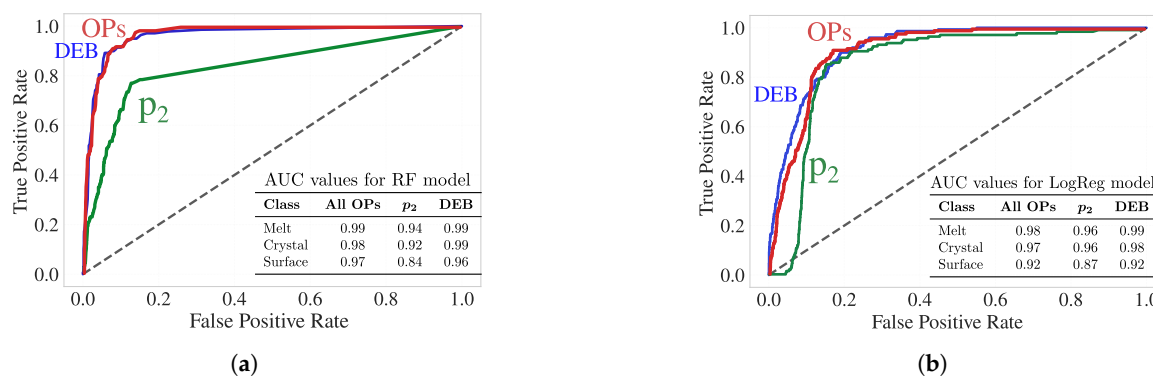


Figure 8. ROC curves for surface classification using (a) RF, and (b) LogReg models, trained on three feature sets: p_2 , the set of conventional scalar OPs (excluding entropy), and DEB. Only the surface-class ROC curves are shown for clarity. AUC values for melt, crystal, and surface classes are reported in the inset tables. Despite the silver truth labels being derived from the full OP set (including entropy), the DEB features alone exhibit competitive or superior surface-prediction performance.

3.3.5. Uncertainty of Cluster_Label

To quantify the uncertainty of the label in interfacial regions, we manually inspected 600 atoms selected near the boundaries of the UMAP cluster (surface atoms) in three representative time steps (200 atoms per timestep). Approximately 10 atoms ($\sim 1.7\%$) were clearly mislabeled, while more than 100 atoms ($\sim 17\%$) exhibited ambiguous labeling due to their proximity to phase interfaces. Conservatively, we adopt an estimate of uncertainty $\sim 2\%$ for clearly mislabeled atoms, acknowledging a potential upper bound of $\sim 18\%$ if ambiguous cases are considered. This manually estimated uncertainty provides a practical ceiling for classification AUCs, beyond which additional predictive gains may reflect overfitting to label noise rather than improved physical resolution of the interfacial structure.

3.3.6. Forward Feature Selection and Model Interpretation

Using both logistic regression and XGBoost classifiers, we applied forward feature selection (FFS) with five-fold stratified cross-validation to identify key directional entropy band descriptors for surface atom classification. The descriptors were added one at a time until the AUC of the test set plateaued or reached the noise floor determined by manual inspection (Section 3.3.5).

As shown in Figure 9, the performance of the classifier was rapidly saturated, usually after incorporating only 3 to 5 directional features of the entropy band. For example, the XGBoost model

reached an AUC of 0.955 for the surface vs. rest classification (including both melt and crystal), while logistic regression plateaued at 0.985 for surface vs. melt and 0.89 for surface vs. crystal. Given our estimated labeling uncertainty of 1–2%, this corresponds to a practical AUC ceiling of approximately 0.98–0.99¹, so the inclusion of additional features led to only marginal gains. This plateau behavior indicates potential overfitting to boundary noise and confirms that a small, physically significant subset of entropy band components suffices for accurate surface classification. Including additional features risks fitting spurious correlations rather than capturing meaningful structure, especially in interfacial regions where labels are intrinsically uncertain due to geometric or thermodynamic ambiguity.

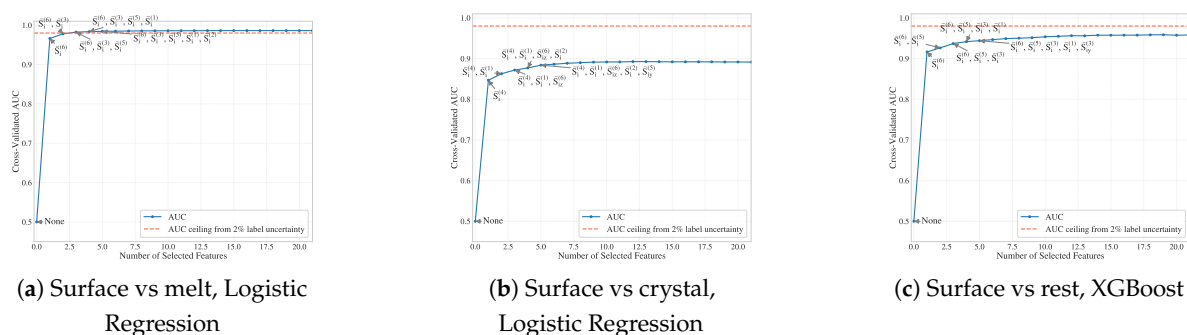


Figure 9. Forward feature selection (FFS) results using ROC-AUC as the performance metric. Dashed red and gray lines denote uncertainty ceilings derived from 2% clearly mislabeled atoms and 18% ambiguous labels, respectively.

Logistic regression was also applied separately to classify surface atoms against either melt or crystal environments. This decomposition avoids confounding effects in three-class linear models and reflects the inherent physical asymmetry between the two interfacial types. In the main model here, constructed with $\alpha = 0.7$ for alpha-shape surfaces and $\sigma = 1.5$ for \vec{f}_i , we observed that surface–crystal boundaries were harder to distinguish, consistent with tighter packing and reduced entropy gradients near the crystalline domains. These surface atoms are more deeply embedded and less geometrically distinct from the bulk, whereas the surface–melt interface exhibits broader spatial separation and more pronounced configurational contrast.

To examine whether the observed asymmetry is due to model resolution or label definition rather than intrinsic structural complexity, we repeated the classification using hyperparameters: $\alpha = 0.3$ for the alpha-shaped surface and $\sigma = 0.2$ for the weighting of the entropy gradient in f_i . This configuration yields a more spacious cluster and a thinner interfacial layer (i.e., fewer surface atoms). As shown in Appendix Figure A5, the surface–crystal classification performance improved to AUC > 0.91, suggesting that the earlier difficulty was partially due to label ambiguity in those definitions. However, the modest gain in AUC (approximately three percentage points) implies that learning the surface–crystal boundary remains intrinsically more challenging using linear decision boundaries (here, logistic regression). This refinement also came at the cost of surface melt classification, which decreased to AUC ≈ 0.95 , probably due to the reduced spatial distinction when the surface region is narrowly defined with lower values α .

These results suggest that surface atom classification using DEB descriptors is consistently robust across different interface definitions. The set of DEB components required to reach performance saturation may vary depending on the classifier and hyperparameters, but the overall predictive power remains strong. Importantly, we observe that AUC values plateau after inclusion of only a few DEB features, underscoring their relevance and sufficiency for identifying interfacial atoms in polymeric systems with varying structural resolution.

¹ Assuming misclassifications occur near the class boundary, the AUC can be approximated as $AUC \approx 1 - (\Delta \times 2250) / (4500 \times 4500)$ under a 50/50 class balance. For example, $\Delta \sim 90$ produces AUC ≈ 0.99 , and $\Delta \sim 180$ produces AUC ≈ 0.98 .

To capture non-linear interactions among entropy band descriptors, we also trained an XGBoost classifier for surface vs. rest (melt + crystal). This model achieved a high AUC (≥ 0.985) using only 3-4 features, indicating a strong discriminative power of the directional entropy bands. SHAP values were computed to interpret these non-linear decision boundaries, focusing on surface atoms and stratifying results by radial and longitudinal surface geometries. Together, these results confirm that directional entropy bands are physically grounded, compact, and effective variables for surface detection and provide a foundation for future modeling of interfacial dynamics.

3.3.7. Feature Importance and Local Structure Patterns

To interpret the trained non-linear XGBoost classifier for surface detection, we calculated SHAP values over the subset of surface atoms (Cluster_Label = 1). The SHAP summary plot (Figure 10a) highlights that the most influential features are the band-averaged entropies $\bar{S}^{(b)}_i$, particularly from the outermost and midrange shells ($b = 6, 5$, and 3), along with directional components such as $\bar{S}^{(3)}_{ix}$ and $\bar{S}^{(3)}_{iy}$. This confirms the central role of band-resolved entropy profiles in characterizing interfacial order. In particular, both positive and negative SHAP values appear across bands, indicating nonmonotonic interactions between local order and surface classification. The heatmap (Figure 10b) further reveals structured, non-random patterns in SHAP values across atoms, with feature contributions varying coherently between instances. These results demonstrate that directional entropy bands capture reproducible interfacial features with high discriminative value, validating their use in subsequent modeling of surface and transition dynamics.

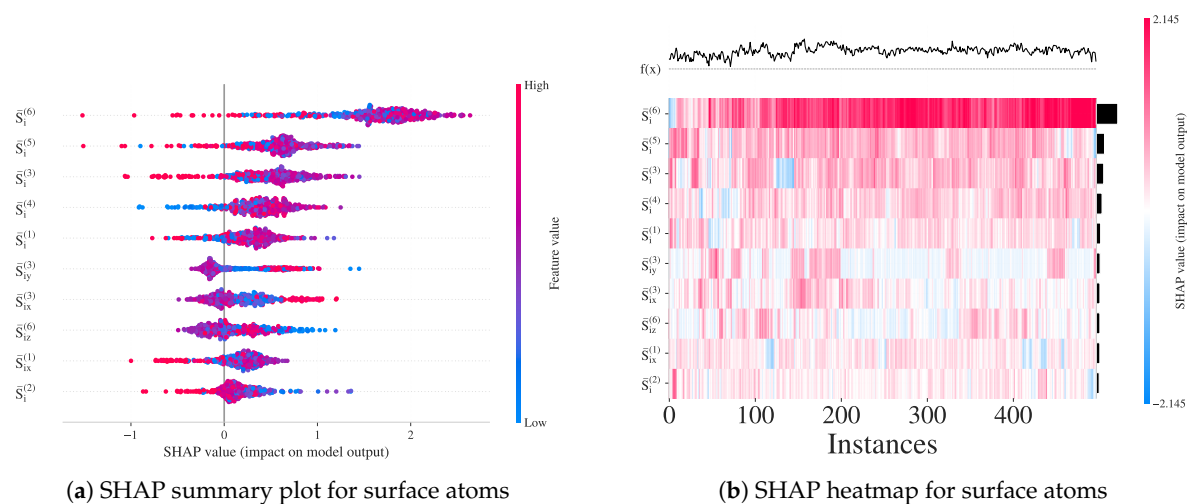


Figure 10. SHAP value analysis for the XGBoost model trained to distinguish surface atoms. (a) Summary plot showing the contribution of each feature across atoms labeled as surface (Cluster_Label = 1). (b) Heatmap visualization of SHAP values across surface atoms, illustrating consistent patterns in the importance of band entropy features.

4. Conclusions

We introduced *directional entropy bands* (DEB): shell-resolved, angularly projected extensions of local entropy that augment scalar entropy with first-order directional moments and radial entropy gradients. In equilibrium polymer crystallization simulations, these descriptors (24-30D) produce a continuous melt-surface-core manifold in UMAP space and separate interfacial atoms with fidelity higher than conventional order parameters (scalar entropy, q_6 , p_2 and SOAP) at single snapshot resolution. Forward feature selection and binary classification demonstrate that (i) surface vs. melt discrimination attains near-noise-ceiling AUC (~ 0.98) with **three** band-averaged entropy features, (ii) surface vs. crystal separation is intrinsically harder (plateau AUC ~ 0.90), reflecting structural proximity of surface and crystalline core, and (iii) a nonlinear model (XGBoost) surface vs rest reaches AUC $\gtrsim 0.955$ using only 3-4 DEB features, confirming that added performance arises from feature interactions rather than large feature sets. SHAP analysis restricted to surface atoms shows that the

dominant contributions arise from the outer and mid shells ($b = 6, 5, 3$) with secondary directional components, validating that the anisotropic interfacial order is captured by low-order angular entropy projections without recourse to high-dimensional spectrum fingerprints. The ability of a compact family of entropy-centric physically interpretable variables to match or exceed rich fingerprints for interfacial detection underscores that the directional organization of local entropy is a primary driver of the crystallization surface signal.

Future Directions

The present work is a first step toward a surface-centric thermodynamic-kinetic modeling framework. Extensions include: (1) **temporal predictions**—contrasting same-time detection with lagged prediction to isolate features whose signal precedes growth rather than merely correlates with established surfaces, and forecasting future surface formation (melt \rightarrow surface attachment) to identify components with predictive (possibly causal) leverage; (2) **early nucleation signaling**—quantifying advance warning times of DEB vs. conventional OPs for emergent crystalline regions; (3) **surface taxonomy**—unsupervised and supervised classification of anisotropic surface classes (facet orientations, fold/edge motifs) using DEB subsets; (4) **coarse thermodynamic modeling**—fitting effective interfacial free-energy or Gibbs-like functionals in reduced DEB space across chain lengths, thermal protocols, and flow conditions; These directions aim to translate the high discriminative power and ease of interpretation of directional entropy bands into predictive, mechanistic models of polymer crystal growth kinetics.

Funding: This research was funded by the Materials Research and Innovation Laboratory (MRAIL) at the Department of Chemical and Biomolecular Engineering, University of Tennessee, Knoxville. The APC was waived by the journal.

Data Availability Statement: The data that support the findings of this study are available from the corresponding author upon reasonable request.

Acknowledgments: The authors acknowledge the Texas Advanced Computing Center (TACC) at The University of Texas at Austin for providing high-performance computing resources used in large-scale molecular dynamics simulations. Post-processing and data analysis were performed using computational infrastructure provided by the ISAAC at the University of Tennessee, Knoxville. Financial support was provided by the Materials Research and Innovation Laboratory (MRAIL) at the University of Tennessee, Knoxville.

Conflicts of Interest: The authors declare that there are no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

MD	Molecular Dynamics
PE	Polyethylene
UA	United Atom
OP	Order Parameter
DEB	Directional Entropy Bands
UMAP	Uniform Manifold Approximation and Projection
SOAP	Smooth Overlap of Atomic Positions
HDBSCAN	Hierarchical Density-Based Spatial Clustering of Applications with Noise
ROC	Receiver Operating Characteristic
AUC	Area Under the Curve (of the ROC)
SHAP	SHapley Additive exPlanations

Appendix A

Appendix A.1 Effect of Skewness Correction on UMAP Embedding

To assess the effect of skewed feature distributions - particularly in thermodynamic-like parameters such as enthalpy, entropy, and entropy bands during the early stages of nucleation, a Yeo-Johnson power transform was tested against standard mean-variance scaling (see Figure A1). The Yeo-Johnson method is appropriate for features with both positive and negative values. It helped reduce skewness; however, its effect on the UMAP manifold structure is minimal. Therefore, the standard scaling was retained for the remainder of the analysis to maintain simplicity and consistency.

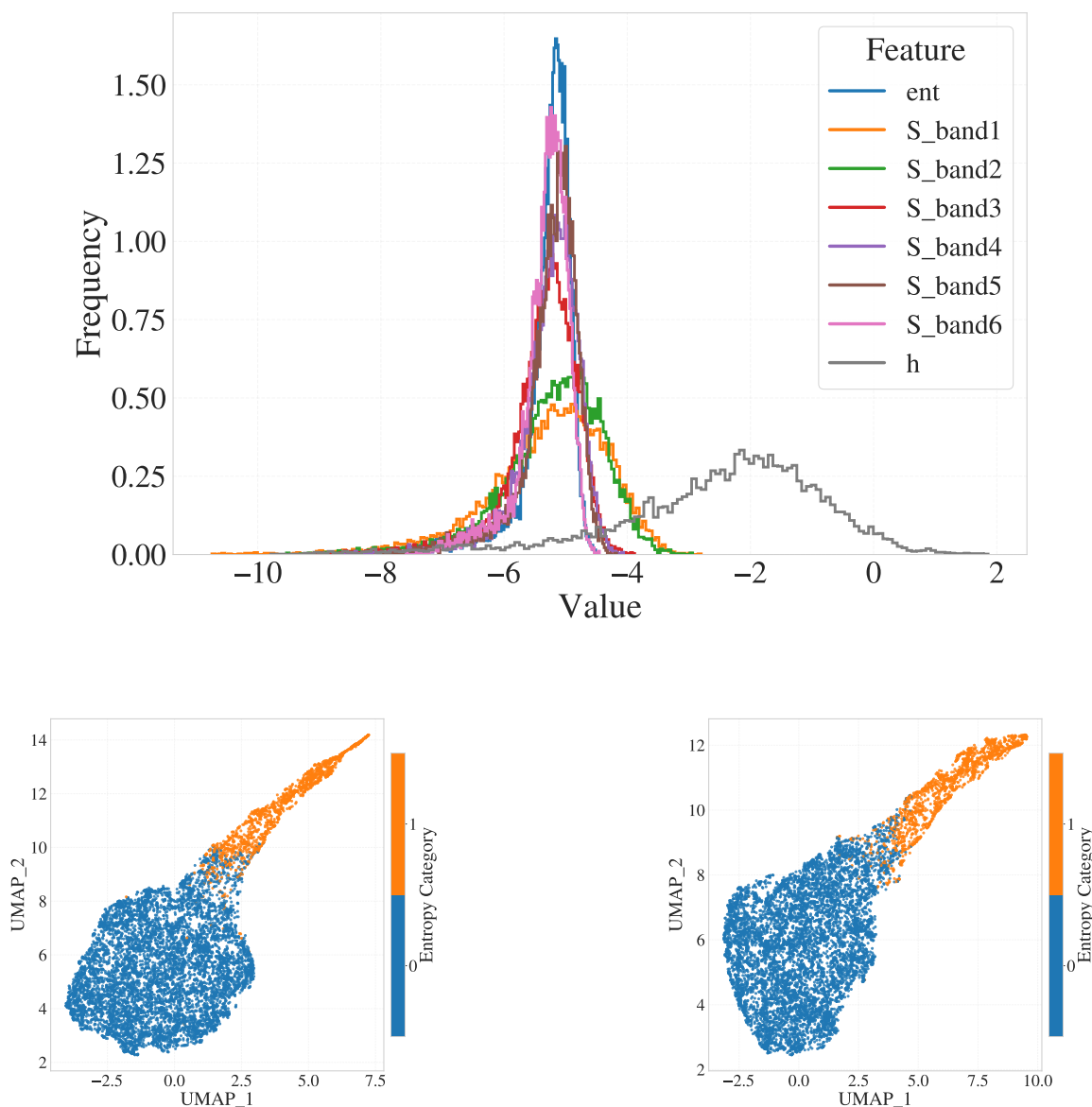


Figure A1. (a) Histograms of raw features at an early crystallization stage: scalar entropy (ent), six entropy-band components (S_{band1} to S_{band6}), and enthalpy (H). Notice the skew in several bands and in enthalpy. (b) 2D UMAP embeddings (for the transitional snapshot t_{tr}), colored by the discrete entropy category. Left: after standard scaling; right: after Yeo-Johnson power transform. While the power transform slightly reduces skew in the latent space, the overall manifold structure and cluster separation remain very similar, so we proceed with standard scaling for simplicity.

Appendix A.2 PCA for SOAP Descriptors

In the Scree plot, Figure A2(a), each point on the y-axis is the percentage of total variance explained by that component alone. We see that once we pass about the components 10~20, each subsequent principal component captures only a tiny additional fraction of the variance.

In the Cumulative Explained Variance plot, Figure A2(b), 15 components already explain more than 95% of the total variance, so taking more than 15 dimensions from the feature space usually offers diminishing returns. However, a detailed study of this is conducted in the main text, Fig. 4.

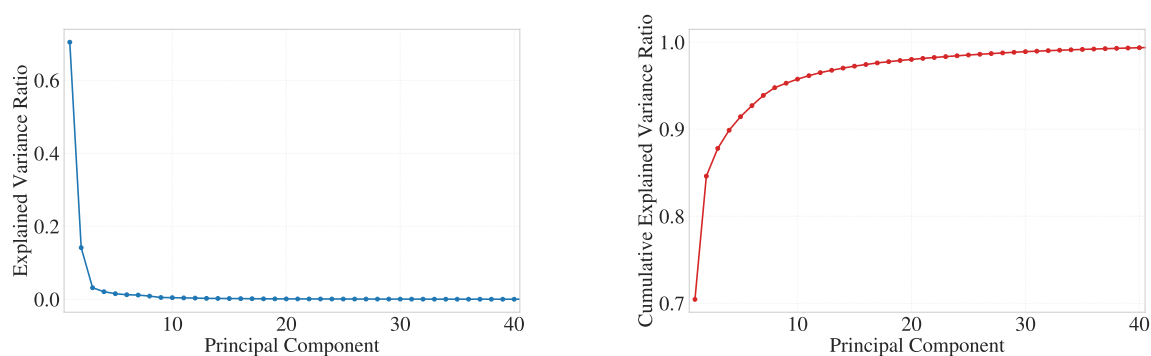


Figure A2. (a) Scree plot showing the individual variance contributions of principal components, illustrating the point of diminishing returns beyond the first 15 components. (b) Cumulative Explained Variance plot indicating where more than 95% of the total variance has been captured, supporting the selection of the first 15 components as a practical cutoff for our analyses. These plots validate our approach to reducing dimensionality in the SOAP descriptor space.

Appendix A.3 Selection of Alpha for Geometric Surface Definition

To determine an appropriate alpha value for geometric surface detection, we evaluated the density of alpha-shaped crystalline clusters over time using varying values of the alpha-shape parameter $\alpha \in [0.01, 1.0]$. For each α , the volume of the alpha shape was calculated and the effective density was estimated using a corrected count of enclosed atoms, normalized by a simulation-specific scaling factor.

Figure A3 shows the resulting density profiles as a function of the number of enclosed particles. At small α values, the alpha shape becomes more permissive, enclosing fewer boundary atoms and void space, systematically leading to lower densities. Based on this analysis, we find that α values in the range of 0.3 to 0.7 yield physically consistent densities near the expected crystalline value for the system under study. For visualization and classification tasks in the main text, we selected $\alpha = 0.7$ to capture a larger set of interfacial atoms without excessive overfitting.

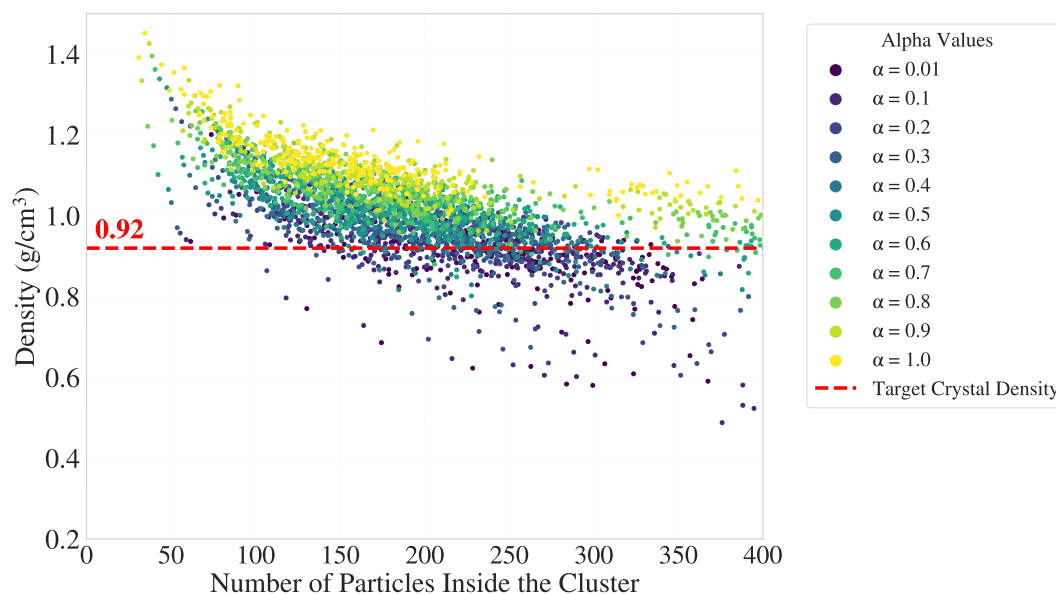


Figure A3. Density of alpha-shaped crystalline clusters as a function of the number of enclosed particles, evaluated across multiple alpha values throughout the growth trajectory. Each point corresponds to one cluster at a given timestep and α . The red dashed line indicates the equilibrium crystalline density (0.92 g/cm^3), obtained from an independently equilibrated bulk simulation using the same force field and model.

Appendix A.4 UMAP Response to Surface Smoothing Parameter, σ_s

To assess the effect of the surface weighting parameter σ_s on the directional entropy gradient f_i , we generated UMAP embeddings at t_{mid} using directional entropy vectors $\vec{S}_i^{(b)}$ calculated with varying values of σ_s . In all cases, the geometric reference alpha shape (with $\alpha = 0.7$) is used to highlight the surface atoms in red.

As the value of σ_s increases, the feature transitions in the UMAP space appear smoother and the surface regions become slightly wider. At lower σ_s values (e.g. 0.1-0.2), surface detection sharpens but can miss transitional atoms. At higher values (e.g. 1.5 to 2.0), interfacial atoms blend more deeply into the feature space, slightly reducing surface-crystal separation. Based on this comparison, $\sigma_s = 0.5, 1.0,$ or 1.5 offers a balanced resolution of surface structure and interfacial continuity.

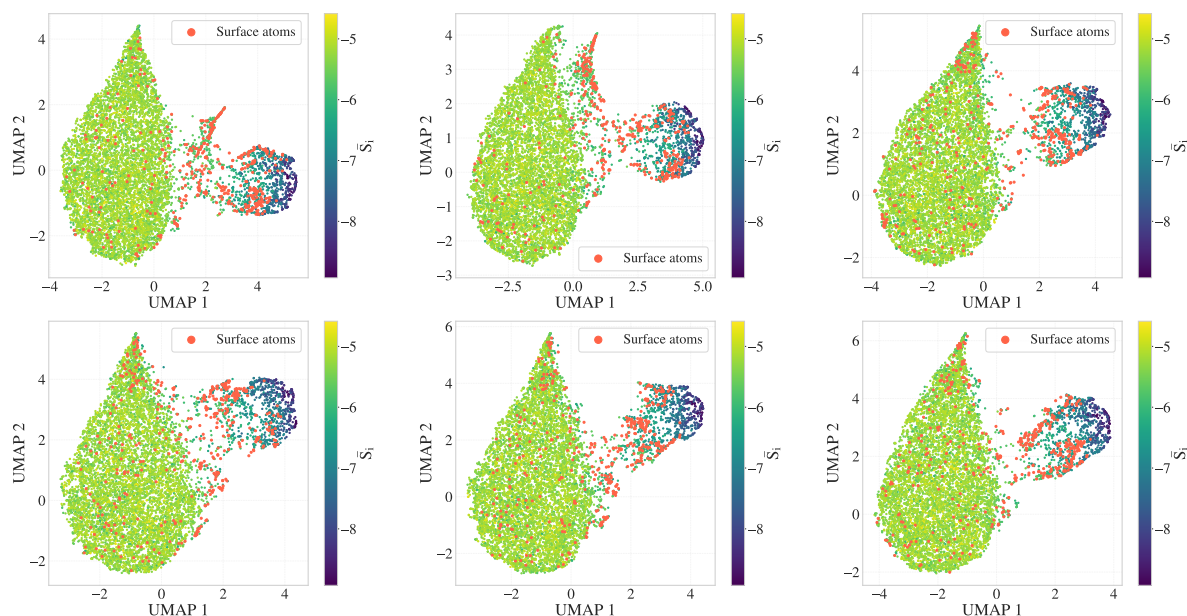


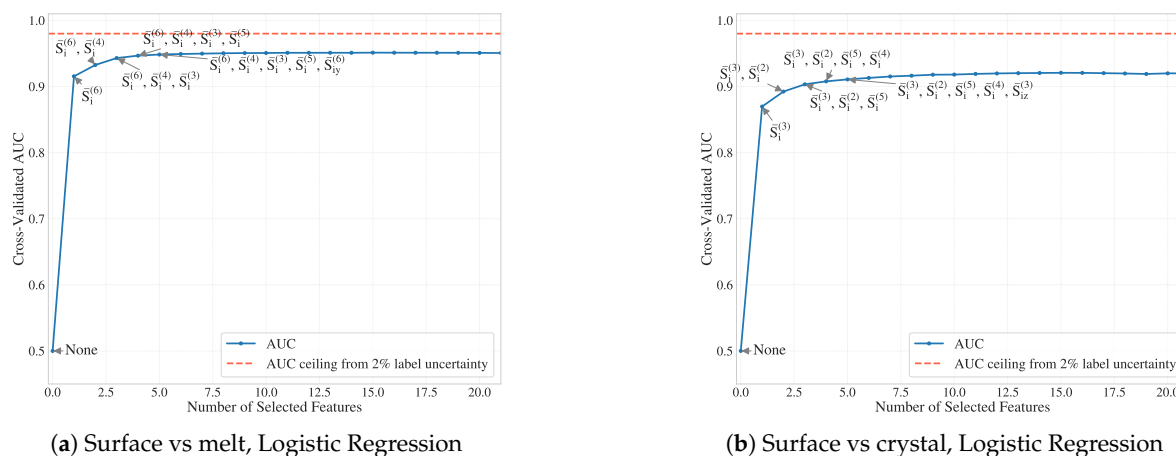
Figure A4. UMAP embeddings of $\vec{S}_i^{(b)}$ at t_{mid} for six values of the entropy gradient smoothing parameter σ_s : (Top row, left to right) $\sigma_s = 0.1, 0.2, 0.5$; (Bottom row, left to right) $\sigma_s = 1.0, 1.5, 2.0$. Red dots mark atoms located on the alpha shape surface ($\alpha = 0.7$). Lower σ_s yields sharper separation but may underdetect soft interfaces; higher σ_s smooths directional gradients and broadens the transitional region. The selected value $\sigma_s = 1.0$ achieves balanced sensitivity and stability.

Appendix A.5 Forward Feature Selection under Alternative Interface Definitions

As shown in Figure A5, these modifications change the relative difficulty of classification between the two interfaces. In panel (a), the performance of logistic regression for surface-melt classification decreases slightly (AUC plateau ~ 0.86), probably due to the reduced spatial separation between melt and surface atoms under tighter surface definitions. In contrast, panel (b) shows improved classification of surface-crystal atoms (AUC > 0.98), supporting the interpretation that the lower AUC in the original model was partially due to the ambiguity of the label. However, the improvement is modest (from ~ 0.89 to > 0.98), implying that learning the surface-crystal boundary remains intrinsically more challenging using linear decision boundaries (here, logistic regression).

To evaluate the sensitivity of surface classification performance to the choice of interface-defining parameters, we repeated the forward feature selection (FFS) analysis using a tighter alpha shape parameter ($\alpha = 0.3$) and a more localized entropy gradient weighting ($\sigma = 0.2$) for computing \vec{f}_i . This configuration produces thinner and more distinct surface layers at the cost of potentially excluding some transitional atoms.

As shown in Figure A5, these changes alter the relative ease of classification between the surface melt and surface-crystal interfaces. In panel (a), the logistic regression model classifies surface versus melt atoms with slightly reduced accuracy (AUC plateau ~ 0.955), likely due to reduced boundary resolution under the new surface definition. In contrast, panel (b) shows significantly improved classification of surface-crystal separation (AUC > 0.91), indicating a clearer distinction when the embedded surface set are slightly excluded from the crystalline atoms using a smaller alpha value. As with previous configurations, the performance saturates after including only a small subset of directional entropy band components, demonstrating that the DEB representation is robust to surface definition changes and consistently encodes meaningful interfacial information.



(a) Surface vs melt, Logistic Regression

(b) Surface vs crystal, Logistic Regression

Figure A5. Forward feature selection using DEB descriptors under stricter surface definitions. (a) Logistic regression model distinguishing surface from melt atoms; (b) logistic regression distinguishing surface from crystalline atoms; (c) These results correspond to interface definitions based on $\alpha = 0.3$ (alpha shape) and $\sigma = 0.2$ (surface-weighted entropy gradient). While the melt interface becomes harder to resolve under this configuration, surface–crystal classification improves significantly. In all cases, performance saturates after a few DEB features, highlighting their robustness across surface extraction parameters.

References

1. Keller, A. Morphology of crystallizing polymers. *Nature* **1952**, *169*, 913–914.
2. Armitstead, K.; Goldbeck-Wood, G.; Keller, A. Polymer crystallization theories. *Macromol. Syn. Order Adv. Prop.* **1992**, pp. 219–312.
3. Gedde, U.L.F. *Polymer Physics*; Springer Science & Business Media, 1995.
4. Sawyer, L.; Grubb, D.T.; Meyers, G.F. *Polymer microscopy*; Springer Science & Business Media, 2008. <https://doi.org/10.1007/978-0-387-72628-1>.
5. Luo, C.; Kröger, M.; Sommer, J.U. Molecular dynamics simulations of polymer crystallization under confinement: Entanglement effect. *Polymer* **2017**, *109*, 71–84.
6. Yamamoto, T. Molecular dynamics simulations of steady-state crystal growth and homogeneous nucleation in polyethylene-like polymer. *J. Chem. Phys.* **2008**, *129*, 184903.
7. Hoffman, J.D.; Davis, G.T.; Lauritzen, J.I., The Rate of Crystallization of Linear Polymers with Chain Folding. In *Treatise on Solid State Chemistry: Vol. 3. Crystalline and Noncrystalline Solids*; Hannay, N.B., Ed.; Springer US: Boston, MA, 1976; pp. 497–614.
8. Sadler, D.; Gilmer, G. A model for chain folding in polymer crystals: rough growth faces are consistent with the observed growth rates. *Polymer* **1984**, *25*, 1446–1452.
9. Olmsted, P.D.; Poon, W.C.; McLeish, T.; Terrill, N.; Ryan, A. Spinodal-assisted crystallization in polymer melts. *Phys. Rev. Lett.* **1998**, *81*, 373.
10. Strobl, G. From the melt via mesomorphic and granular crystalline layers to lamellar crystallites: A major route followed in polymer crystallization? *Europ. Phys. J. E* **2000**, *3*, 165–183.
11. Tang, X.; Yang, J.; Xu, T.; Tian, F.; Xie, C.; Li, L. Local structure order assisted two-step crystal nucleation in polyethylene. *Phys. Rev. Mat.* **2017**, *1*, 073401.
12. Yi, P.; Rutledge, G.C. Molecular simulation of crystal nucleation in n-octane melts. *J. Chem. Phys.* **2009**, *131*, 134902.
13. Yamamoto, T. Molecular dynamics simulation of polymer ordering. II. Crystallization from the melt. *J. Chem. Phys.* **2001**, *115*, 8675–8680.
14. Anwar, M.; Turci, F.; Schilling, T. Crystallization mechanism in melts of short n-alkane chains. *J. Chem. Phys.* **2013**, *139*, 214904.
15. Muthukumar, M.; Welch, P. Modeling polymer crystallization from solutions. *Polymer* **2000**, *41*, 8833–8837.
16. Yamamoto, T. Molecular dynamics of polymer crystallization revisited: Crystallization from the melt and the glass in longer polyethylene. *J. Chem. Phys.* **2013**, *139*, 054903.
17. Yi, P.; Locker, C.R.; Rutledge, G.C. Molecular dynamics simulation of homogeneous crystal nucleation in polyethylene. *Macromolecules* **2013**, *46*, 4723–4733.

18. Anwar, M.; Schilling, T. Crystallization of polyethylene: A molecular dynamics simulation study of the nucleation and growth mechanisms. *Polymer* **2015**, *76*, 307–312.
19. Nafar Sefiddashti, M.H.; Edwards, B.J.; Khomami, B. A thermodynamically inspired method for quantifying phase transitions in polymeric liquids with application to flow-induced crystallization of a polyethylene melt. *Macromolecules* **2020**, *53*, 10487–10502.
20. Hussain, M.A.; Yamamoto, T.; Adil, S.F.; Yao, S. Coarse-grained simulations on polyethylene crystal network formation and microstructure analysis. *Polymers* **2024**, *16*, 1007.
21. Tanaka, H.; Tong, H.; Shi, R.; Russo, J. Revealing key structural features hidden in liquids and glasses. *Nature Rev. Phys.* **2019**, *1*, 333–348.
22. Musil, F.; Grisafi, A.; Bartók, A.P.; Ortner, C.; Csányi, G.; Ceriotti, M. Physics-inspired structural representations for molecules and materials. *Chem. Rev.* **2021**, *121*, 9759–9815.
23. Esselink, K.; Hilbers, P.; Van Beest, B. Molecular dynamics study of nucleation and melting of n-alkanes. *J. Chem. Phys.* **1994**, *101*, 9033–9041.
24. Yi, P.; Rutledge, G.C. Molecular simulation of bundle-like crystal nucleation from n-eicosane melts. *J. Chem. Phys.* **2011**, *135*, 024903.
25. Nicholson, D.A.; Rutledge, G.C. Analysis of nucleation using mean first-passage time data from molecular dynamics simulation. *J. Chem. Phys.* **2016**, *144*, 134105.
26. Faken, D.; Jónsson, H. Systematic analysis of local atomic structure combined with 3D computer graphics. *Comp. Mat. Sci.* **1994**, *2*, 279–286.
27. Ackland, G.; Jones, A. Applications of local crystal structure measures in experiment and simulation. *Phys. Rev. B* **2006**, *73*, 054104.
28. Kelchner, C.L.; Plimpton, S.J.; Hamilton, J.C. Dislocation nucleation and defect structure during surface indentation. *Phys. Rev. B* **1998**, *58*, 11085.
29. Finney, J. Random packings and the structure of simple liquids. I. The geometry of random close packing. *Proc. Roy. Soc. London: A. Math. Phys. Sci.* **1970**, *319*, 479–493.
30. Tanemura, M.; Hiwatari, Y.; Matsuda, H.; Ogawa, T.; Ogita, N.; Ueda, A. Geometrical analysis of crystallization of the soft-core model in an FCC crystal formation. *Progr. Theo. Phys.* **1978**, *59*, 323–324.
31. Steinhardt, P.J.; Nelson, D.R.; Ronchetti, M. Bond-orientational order in liquids and glasses. *Phys. Rev. B* **1983**, *28*, 784–805.
32. Nelson, D.R.; Halperin, B. Dislocation-mediated melting in two dimensions. *Phys. Rev. B* **1979**, *19*, 2457–2484.
33. Chakravarty, C.; Debenedetti, P.G.; Stillinger, F.H. Lindemann measures for the solid-liquid phase transition. *J. Chem. Phys.* **2007**, *126*.
34. Tanaka, H.; Kawasaki, T.; Shintani, H.; Watanabe, K. Critical-like behaviour of glass-forming liquids. *Nature Mat.* **2010**, *9*, 324–331.
35. Ikeda, A.; Miyazaki, K. Glass transition of the monodisperse gaussian core model. *Phys. Rev. Lett.* **2011**, *106*, 015701.
36. Musil, F.; Grisafi, A.; Bartók, A.P.; Ortner, C.; Csányi, G.; Ceriotti, M. Physics-inspired structural representations for molecules and materials. *Chem. Rev.* **2021**, *121*, 9759–9815.
37. Mickel, W.; Kapfer, S.C.; Schröder-Turk, G.E.; Mecke, K. Shortcomings of the bond orientational order parameters for the analysis of disordered particulate matter. *J. Chem. Phys.* **2013**, *138*.
38. van Meel, J.A.; Filion, L.; Valeriani, C.; Frenkel, D. A parameter-free, solid-angle based, nearest-neighbor algorithm. *J. Chem. Phys.* **2012**, *136*.
39. Larsen, P.M.; Schmidt, S.; Schiøtz, J. Robust structural identification via polyhedral template matching. *Model. Sim. Mat. Sci. Eng.* **2016**, *24*, 055007.
40. Bartók, A.P.; Kondor, R.; Csányi, G. On representing chemical environments. *Phys. Rev. B* **2013**, *87*, 184115.
41. Spellings, M.; Glotzer, S.C. Machine learning for crystal identification and discovery. *AIChE J.* **2018**, *64*, 2198–2206.
42. Adorf, C.S.; Moore, T.C.; Melle, Y.J.; Glotzer, S.C. Analysis of self-assembly pathways with unsupervised machine learning algorithms. *J. Phys. Chem. B* **2019**, *124*, 69–78. <https://doi.org/10.1021/acs.jpcc.9b09621>.
43. De, S.; Bartók, A.P.; Csányi, G.; Ceriotti, M. Comparing molecules and solids across structural and alchemical space. *Phys. Chem. Chem. Phys.* **2016**, *18*, 13754–13769.
44. Wang, J.; Ferguson, A. Nonlinear machine learning in simulations of soft and biological materials. *Mol. Sim.* **2018**, *44*, 1090–1107.
45. Butler, K.T.; Davies, D.W.; Cartwright, H.; Isayev, O.; Walsh, A. Machine learning for molecular and materials science. *Nature* **2018**, *559*, 547–555.

46. Barakati, K.; Liu, Y.; Nelson, C.; Ziatdinov, M.; Zhang, X.; Takeuchi, I.; Kalinin. Reward driven workflows for unsupervised explainable analysis of phases and ferroic variants from atomically resolved imaging data. *Advanced Materials* **2024**, p. 2418927.
47. Asoudeh, M.; Nguyen, N.; Raith, M.; Denman, D.S.; Anozie, U.C.; Mokhtarnejad, M.; Khomami, B.; Skotty, K.M.; Isaac, S.; Gebhart, T.; et al. PEGylated nanoparticles interact with macrophages independently of immune response factors and trigger a non-phagocytic, low-inflammatory response. *Journal of Controlled Release* **2024**, *366*, 282–296.
48. Wang, L. Discovering phase transitions with unsupervised learning. *Phys. Rev. B* **2016**, *94*, 195105.
49. Jadrlich, R.B.; Lindquist, B.A.; Truskett, T.M. Unsupervised machine learning for detection of phase transitions in off-lattice systems. I. Foundations. *J. Chem. Phys.* **2018**, *149*, 194109.
50. Boattini, E.; Dijkstra, M.; Filion, L. Unsupervised learning for local structure detection in colloidal systems. *J. Chem. Phys.* **2019**, *151*, 154901.
51. Bhardwaj, A.; Sommer, J.U.; Werner, M. Nucleation patterns of polymer crystals analyzed by machine learning models. *Macromolecules* **2024**, *57*, 9711–9724.
52. Nettleton, R.; Green, M. Expression in terms of molecular distribution functions for the entropy density in an infinite system. *J. Chem. Phys.* **1958**, *29*, 1365–1370.
53. Piaggi, P.M.; Valsson, O.; Parrinello, M. Enhancing entropy and enthalpy fluctuations to drive crystallization in atomistic simulations. *Phys. Rev. Lett.* **2017**, *119*, 015701.
54. Piaggi, P.M.; Parrinello, M. Entropy based fingerprint for local crystalline order. *J. Chem. Phys.* **2017**, *147*, 114112.
55. Nafar Sefiddashti, M.H.; Edwards, B.J.; Khomami, B. Flow-induced phase separation and crystallization in entangled polyethylene solutions under elongational flow. *Macromolecules* **2020**, *53*, 6432–6451.
56. Tourani, E.; Edwards, B.J.; Khomami, B. Machine Learning Workflow for Analysis of High-Dimensional Order Parameter Space: A Case Study of Polymer Crystallization from Molecular Dynamics Simulations, 2025, [arXiv:physics.comp-ph/2507.17980].
57. Plimpton, S. Fast parallel algorithms for short-range molecular dynamics. *J. Comp. Phys.* **1995**, *117*, 1–19.
58. Thompson, A.P.; Aktulga, H.M.; Berger, R.; Bolintineanu, D.S.; Brown, W.M.; Crozier, P.S.; In't Veld, P.J.; Kohlmeyer, A.; Moore, S.G.; Nguyen, T.D. LAMMPS—a flexible simulation tool for particle-based materials modeling at the atomic, meso, and continuum scales. *Comp. Phys. Comm.* **2022**, *271*, 108171.
59. Siepmann, J.I.; Karaborni, S.; Smit, B. Simulating the critical behaviour of complex fluids. *Nature* **1993**, *365*, 330–332.
60. Moore, J.; Cui, S.; Cochran, H.; Cummings, P. A molecular dynamics study of a short-chain polyethylene melt: I. steady-state shear. *J. Non-Newtonian Fluid Mech.* **2000**, *93*, 83–99.
61. Baig, C.; Edwards, B.J.; Keffer, D.J.; Cochran, H.D.; Harmandaris, V. Rheological and structural studies of linear polyethylene melts under planar elongational flow using nonequilibrium molecular dynamics simulations. *J. Chem. Phys.* **2006**, *124*, 084902.
62. Cui, S.; Cummings, P.; Cochran, H. Multiple time step nonequilibrium molecular dynamics simulation of the rheological properties of liquid n-decane. *J. Chem. Phys.* **1996**, *104*, 255–262.
63. Baig, C.; Edwards, B.J.; Keffer, D.J.; Cochran, H.D. Rheological and structural studies of liquid decane, hexadecane, and tetracosane under planar elongational flow using nonequilibrium molecular-dynamics simulations. *J. Chem. Phys.* **2005**, *122*, 184906.
64. Ionescu, T.C.; Baig, C.; Edwards, B.J.; Keffer, D.J.; Habenschuss, A. Structure formation under steady-state isothermal planar elongational flow of n-eicosane: A comparison between simulation and experiment. *Phys. Rev. Lett.* **2006**, *96*, 037802.
65. Zhang, M.C.; Guo, B.H.; Xu, J. A review on polymer crystallization theories. *Crystals* **2016**, *7*, 4. <https://doi.org/10.3390/cryst7010004>.
66. Yamamoto, T. Molecular dynamics simulation of stretch-induced crystallization in polyethylene: Emergence of fiber structure and molecular network. *Macromolecules* **2019**, *52*, 1695–1706.
67. Nafar Sefiddashti, M.; Edwards, B.; Khomami, B. Individual chain dynamics of a polyethylene melt undergoing steady shear flow. *J. Rheol.* **2015**, *59*, 119–153.
68. Gasparotto, P.; Ceriotti, M. Recognizing molecular patterns by machine learning: An agnostic structural definition of the hydrogen bond. *J. Chem. Phys.* **2014**, *141*, 174110.
69. Yang, J.; De, S.; Campbell, J.E.; Li, S.; Ceriotti, M.; Day, G.M. Large-scale computational screening of molecular organic semiconductors using crystal structure prediction. *Chem. Mat.* **2018**, *30*, 4361–4371.

70. Helfrecht, B.A.; Cersonsky, R.K.; Fraux, G.; Ceriotti, M. Structure-property maps with Kernel principal covariates regression. *Mach. Learn.: Sci. Techn.* **2020**, *1*, 045021.
71. Himanen, L.; Jäger, M.O.; Morooka, E.V.; Canova, F.F.; Ranawat, Y.S.; Gao, D.Z.; Rinke, P.; Foster, A.S. DScribe: Library of descriptors for machine learning in materials science. *Comp. Phys. Comm.* **2020**, *247*, 106949.
72. Baranyai, A.; Evans, D.J. Direct entropy calculation from computer simulation of liquids. *Phys. Rev. A* **1989**, *40*, 3817–3822.
73. Laird, B.B.; Haymet, A. Calculation of the entropy from multiparticle correlation functions. *Phys. Rev. A* **1992**, *45*, 5680–5689.
74. Nafar Sefiddashti, M.H.; Edwards, B.J.; Khomami, B. A method for calculating the nonequilibrium entropy of a flowing polymer melt via atomistic simulation. *J. Chem. Phys.* **2021**, *155*, 111101.
75. Edwards, B.J.; Nafar Sefiddashti, M.H.; Khomami, B. Nonequilibrium thermodynamics of polymeric liquids via atomistic simulation. *Entropy* **2022**, *24*, 175.
76. Nafar Sefiddashti, M.H.; Edwards, B.J.; Khomami, B. Atomistic simulation of flow-induced microphase separation and crystallization of an entangled polyethylene melt undergoing uniaxial elongational flow and the role of Kuhn segment extension. *Polymers* **2023**, *15*, 1831.
77. Nafar Sefiddashti, M.H.; Edwards, B.J.; Khomami, B. 122.) Flow-induced phase phenomena in an entangled polyethylene/benzene solution under uniaxial elongational flow. *Rheo. Acta* **2024**, *63*, 113–133.
78. Baig, C.; Edwards, B.J. Atomistic simulation of flow-induced crystallization at constant temperature. *Europhys. Lett.* **2010**, *89*, 36003.
79. Baig, C.; Edwards, B.J. Atomistic simulation of crystallization of a polyethylene melt in steady uniaxial extension. *J. Non-Newtonian Fluid Mech.* **2010**, *165*, 992–1004.
80. McInnes, L.; Healy, J.; Melville, J. Umap: Uniform manifold approximation and projection for dimension reduction. *arXiv preprint arXiv:1802.03426* **2018**.
81. Hotelling, H. Analysis of a complex of statistical variables into principal components. *Journal of educational psychology* **1933**, *24*, 417. <https://doi.org/10.1037/h0071325>.
82. McInnes, L.; Healy, J.; Astels, S. hdbscan: Hierarchical density based clustering. *J. Open Source Softw.* **2017**, *2*, 205.
83. Humphrey, W.; Dalke, A.; Schulten, K. VMD: visual molecular dynamics. *Journal of molecular graphics* **1996**, *14*, 33–38.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.