Visualization of free energy landscape in spinodal decomposition using persistent homology combined with unsupervised machine learning

Tokyo Univ. of Science[1], ○Alexandre Lira Foggiatto[1], Hirotaka Aoki[1], Sotaro Kunii[1], Masato Kotsugi[1]
E-mail: a.foggiatto@rs.tus.ac.jp

Although, nowadays, the amount of available image data is enormous, the discoveries are still limited by the analysis process which is not improving at the same rate as the data extraction. Materials informatics is an emergent field that deals with this problem by combining computer and material science. Many prediction models and material discovery has been done using this methodology, but there is still a lack of application in more fundamental researches. Persistent homology is a powerful tool to extract topological features from the microstructure, such as the size, shapes, and connectivity of holes. It can also be combined with machine learning to find hiding correlations between variables in the data set.

In this work, we prepared spinodal decomposition images using phase-field calculation and applied persistent homology to study the relationship between free energy and morphology. The image sets were prepared using different values for the gradient energy coefficient (κ), a parameter that controls the probability of phase separation. We prepared 5 sets of 400 images for each value of κ and we computed the total free energy (Fig. 1) at each step to confirm the convergence of the calculation. After concluding the simulation, persistent homology was applied to image sets to obtain persistent diagrams (PD). The PD maps the topological features of an image in a set of points (”birth”,”death”) that can be later analyzed using machine learning or other informatics technique. In our case, an unsupervised machine learning algorithm called kernel principal component analysis (PCA) was implemented to search for non-linear correlations between the PDs (Fig. 2). The PCA scatter plot shows that the data can be separated and clustered regarding the principal components based on the input values. One surprising result is that the contribution of these two variables is near 1.0, which implies that a large amount of image data could successfully be embedded in a low-dimensional manifold. One may concern about the extremities of Fig. 2. However, the free energy and image features are almost similar in these regions resulting in closer points in the principal components.

Concretely, the separation of κ and continuity of energy value were successfully visualized. Thus, we believe that this approach can be used in experimental images, and it might be helpful to improve the analysis and to extract information of hidden parameters.

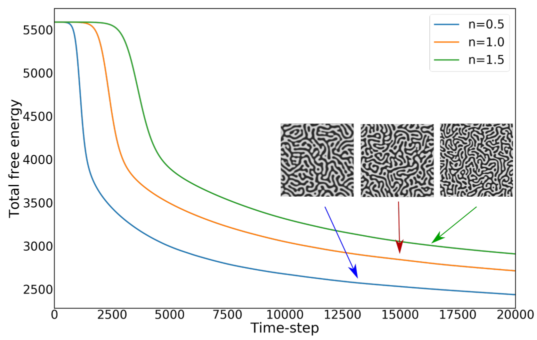

Figure 1: Total free energy as a function of time for different values of n∗κ.

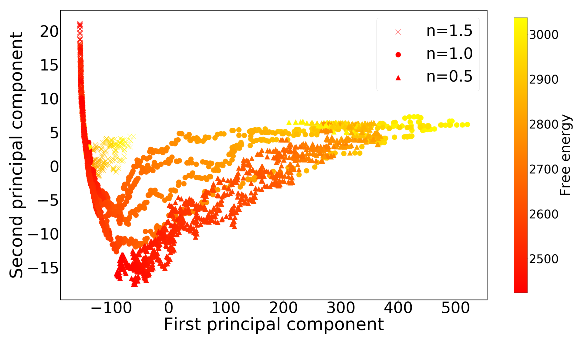

Figure 2: PCA decomposition as a function of the total free energy.