Supplementary Materials

Structure–property modeling scheme based on optimized microstructural information by two-point statistics and principal component analysis

Xiaobing Hu, Jiajun Zhao, Yiming Chen, Yujian Wang, Junjie Li*, Qingfeng Wu, Zhijun Wang, Jincheng Wang**

State Key Laboratory of Solidification Processing, Northwestern Polytechnical University, Xi'an 710072, China

S1. Application of our method in simulated dataset for Al-Cu alloys

The multiphase field model is a powerful tool for simulating dendritic growth during solidification[1,2]. The reason we used this model to simulate the microstructures of binary Al-Cu alloys is to simplify the calculation because only autocorrelations of each microstructure should be considered. For instance, if the ternary eutectic microstructures of a Al-Cu-Ag alloy are generated, at least autocorrelations and cross-correlations of two thermodynamic phases should be calculated to ensure the integrity of statistical information[3–6]. Further research will be conducted on this case. In the multiphase field model, the phase state $\phi_1 = -1$ is chosen to represent the liquid phase, whereas ϕ_i $(i \neq 1)$ denotes the solid phase with various crystalline orientations. The sum constraint $\sum_{j=1\ldots n} \phi_j = 1$ is maintained throughout. The governing equation of the phase field is expressed as follows:

$$\frac{\partial \phi_i}{\partial t} = -\frac{2}{N}\sum_{j\neq i}^{n} s_i s_j M_{ij}\left(\frac{\delta f}{\delta \phi_i} - \frac{\delta f}{\delta \phi_j} + \Delta g_{ij}\right) \tag{1}$$

where

$$\frac{\delta f}{\delta \phi_i} = \begin{cases} \sum_{j\neq 1}^{n}\left(\frac{\varepsilon_{SL}^2}{2}\nabla^2\phi_j + w_{SL}\phi_j\right) & \text{for i} = 1 \\ \frac{\varepsilon_{SL}^2}{2}\nabla^2\phi_1 + w_{SL}\phi_1 + \sum_{j\neq 1,i}^{n}\left(\frac{\varepsilon_{SS}^2}{2}\nabla^2\phi_j + w_{SS}\phi_j\right) & \text{for i} \neq 1 \end{cases} \tag{2}$$

$$\Delta g_{ij} = \begin{cases} 0 & \text{for } i \neq 1 \text{and } j \neq 1 \\ 6\phi_1(1-\phi_1)\Delta S(T_m - T - mc_L) & \text{for } i = 1 \text{and } j \neq 1 \\ -6\phi_1(1-\phi_1)\Delta S(T_m - T - mc_L) & \text{for } i \neq 1 \text{and } j = 1 \end{cases} \quad (3)$$

The parameters ε and w with subscript *SL* or *SS* are determined from the interface energy σ and interface width 2ξ of the solid–liquid interface or solid–solid interface:

$$\varepsilon_{SL} = \frac{4}{\pi}\sqrt{\xi_{SL}\sigma_{SL}}\ ; \quad w_{SL} = \frac{2\sigma_{SL}}{\xi_{SL}} \quad (4)$$

$$\varepsilon_{SS} = \frac{4}{\pi}\sqrt{\xi_{SS}\sigma_{SS}}\ ; \quad w_{SS} = \frac{2\sigma_{SS}}{\xi_{SS}} \quad (5)$$

By adjusting the ratio between σ_{SL} and σ_{SS}, the wetting property of grain boundary (GB) can be adjusted. The formation of GB is not considered here, and we set $\sigma_{SS} = 2.5\sigma_{SL}$ and $\xi_{SL} = \xi_{SS}$.

In the case of a fourfold anisotropic solid–liquid interfacial energy, the term ε_{SL} should be extended to the following:

$$\varepsilon_{SL} = \varepsilon_0(1-3\gamma_4)\left[1 + \frac{4\gamma_4}{1-3\gamma_4}\frac{(\partial\phi_1/\partial x)^4 + (\partial\phi_1/\partial y)^4}{[(\partial\phi_1/\partial x)^2 + (\partial\phi_1/\partial y)^2]^2}\right] \quad (6)$$

where γ_4 is the anisotropy coefficient of interface energy, which represents the magnitude of the anisotropy.

The diffusion equation with an anti-trapping current is expressed as[2,7]

$$\frac{\partial c}{\partial t} = \nabla\phi_1 D_L \nabla c_L + \nabla\left(\frac{\varepsilon_{SL}}{\sqrt{2w_{SL}}}(c_L - c_S)\sqrt{\phi_1(1-\phi_1)}\frac{\partial\phi_1}{\partial t}\frac{\nabla\phi_1}{|\nabla\phi_1|}\right) \quad (7)$$

where D_L is the liquid diffusion coefficient. The diffusion in the solid phase is neglected here. $c=\phi_1 c_L+(1-\phi_1)c_S$ is a mixture concentration of solid and liquid.

This multiphase field model was successfully applied to dendritic growth competition during directional solidification in our previous works[8,9]. Additional details on the model are presented in [10].

In the present study, the growth of multi-dendrites with various orientations during isothermal solidification of Al-Cu binary alloy was simulated inside a square domain of size $1000\mathrm{d}x \times 1000\mathrm{d}x$. The degree of nucleation supercooling (ΔT), the number of primary grains (n), and anisotropy coefficient of interface energy (γ_4) were chosen as the process parameters that control microstructures. The parameters used in the simulations are listed in Table S1. Six cases with different combinations of process parameters were simulated, as shown in Figure S1.

Table S1. Parameters and physical properties[11].

Parameters	**Variable**	**Value**
Melting point of Al	T_m/K	933.6
Interface energy	σ/Jm^{-2}	0.24
Entropy of fusion	$\Delta S/Jm^{-3}K^{-1}$	1.01×10^6
Liquidus slope	$m/Kwt\%^{-1}$	-2.6
Partition coefficient	k	0.14
Anisotropy strength	γ_4	0.01, 0.04
Initial concentration	$c_0/wt\%$	3.0
Liquid diffusivity	D_L/m^2s^{-1}	3×10^{-9}
Undercooling	$\Delta T/K$	4, 8, 12
Grid size	$dx/\mu m$	0.5
The number of primary grains	n	0.01, 0.04

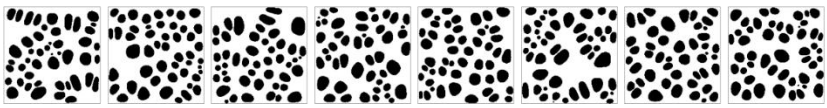
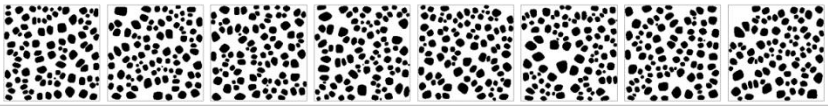
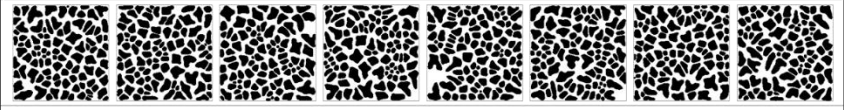
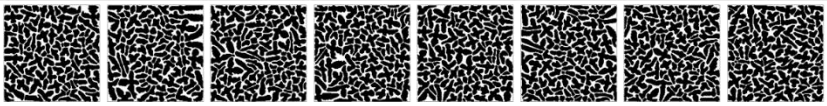
Labels	Parameters			Microstructures produced by PFM
	n	γ_4	ΔT	
#1	50	0.01	4	
#2	50	0.01	8	
#3	50	0.01	12	
#4	100	0.04	4	
#5	100	0.04	8	
#6	100	0.04	12	

Figure S1. Microstructures produced by PFM. n, γ_4, and ΔT represent the number of primary grains, anisotropy coefficient of interfacial free energy, and degree of nucleation supercooling, respectively.

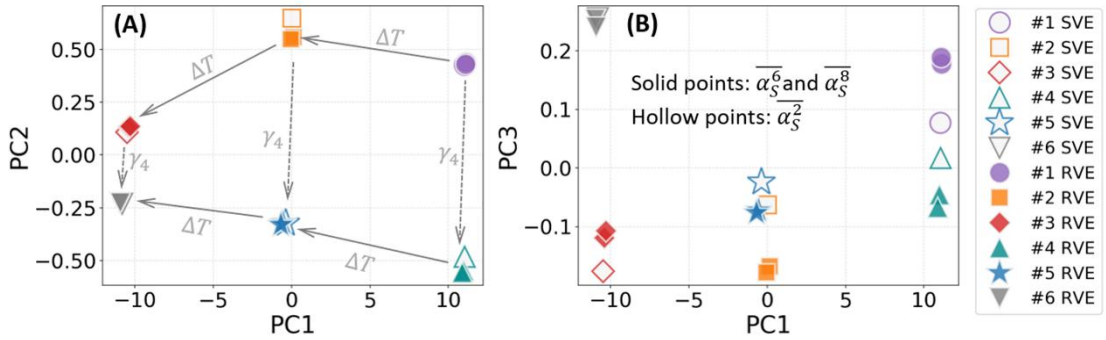


Figure S2. The distribution of the averaged RVE and SVEs in the PCA space: (A) distribution in PC1–PC2 vectors; and (B) distribution in PC1–PC3 vectors.

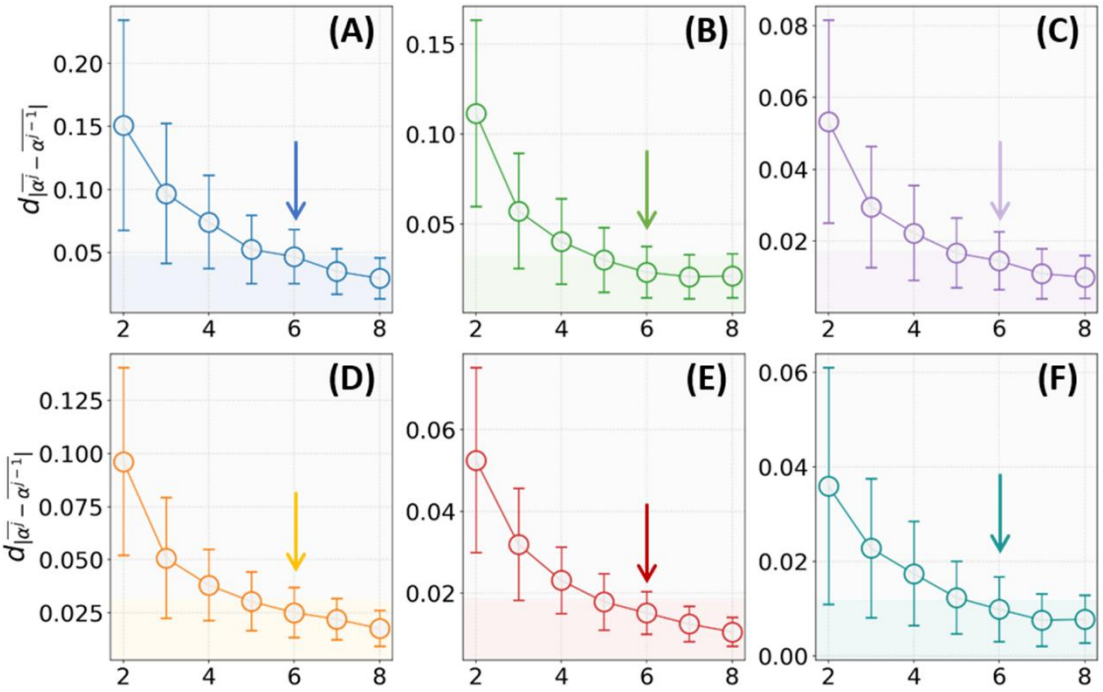


Figure S3. Distance $d_{\left|\overline{\alpha_{s^*}^{j}} - \overline{\alpha_{s^*}^{j-1}}\right|}$ as a function of the number of SVEs by recursively adding for the six simulated samples shown in Figure S1.

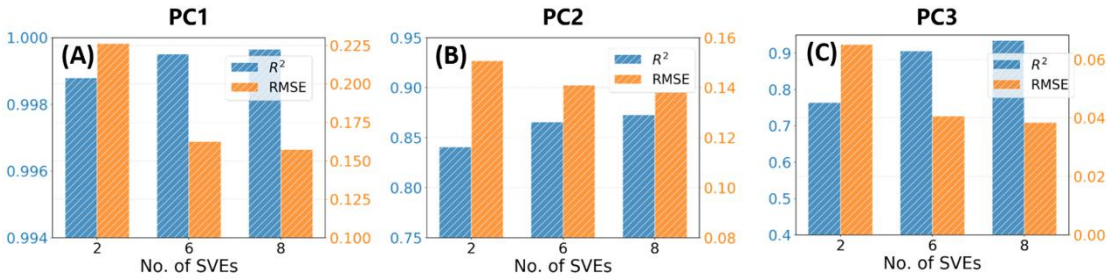


Figure S4. Accuracy of prediction for the first three PCs: (A) PC1; (B) PC2; and (C) PC3.

S2. Application of our method in the ferritic heat-resistant steels dataset (supplementary results)

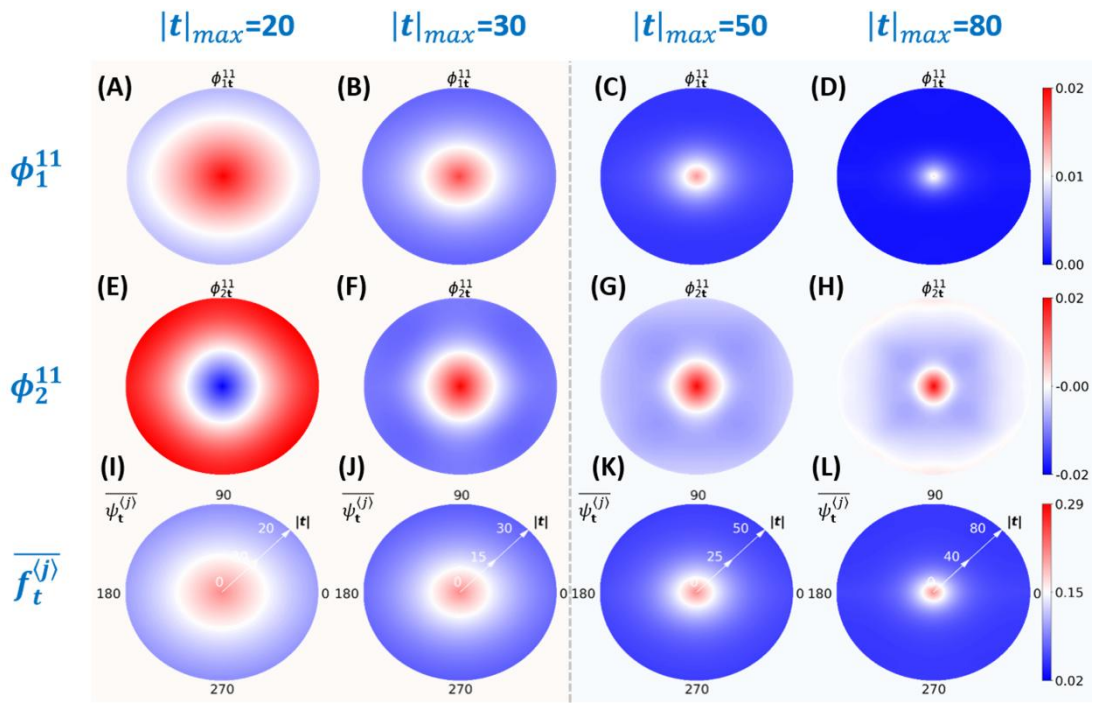


Figure S5. Variation of the first two PC basis and ensemble averaged autocorrelations with different lengths of $|\boldsymbol{t}|_{max}$.

Combining with Equation (3), one can understand the representative meanings of the first two PCs from their basis vectors shown in Figure S5. When $|\boldsymbol{t}|_{max} \geq 50$ pixels, if PC1 increases, the peak value of the autocorrelations that is strongly associated with the volume fraction of austenite phase will also increase. In other words, PC1 represents the volume fraction. As for PC2, it mainly relates to the peak value and the size of the central area, indicating that PC2 determines the volume fraction, average size, and distribution of the austenite phase. When $|\boldsymbol{t}|_{max} < 50$ pixels, PC1 is not only correlated with phase volume fraction but also related to the average size of the phase, and PC2 does not contain the information about phase distribution as it does before.

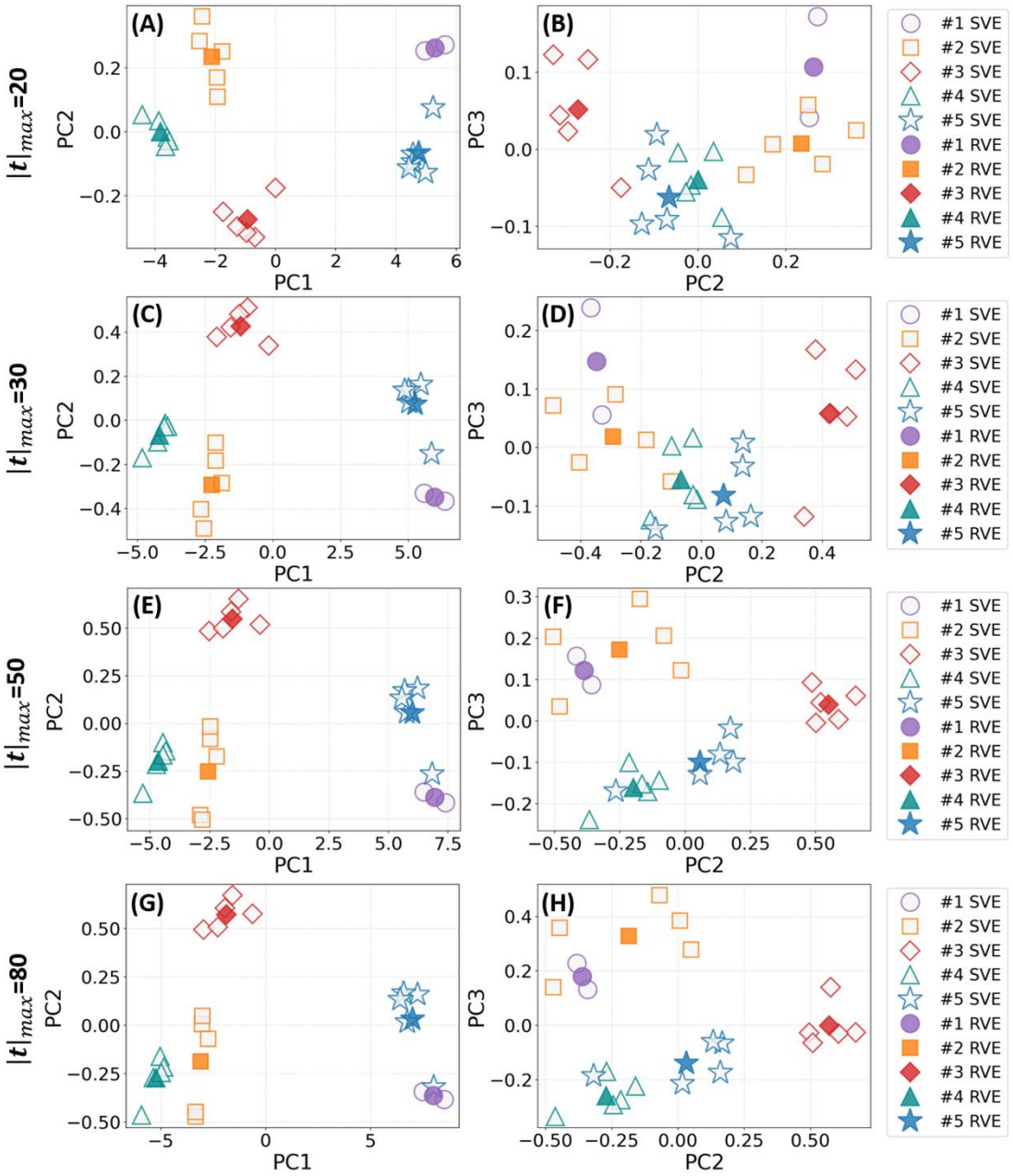


Figure S6. Variation of the distribution of the five experimental alloys in the first three PCs with different lengths of $|t|_{max}$.

S3. Application of our method in the Ni-Fe-based superalloy dataset

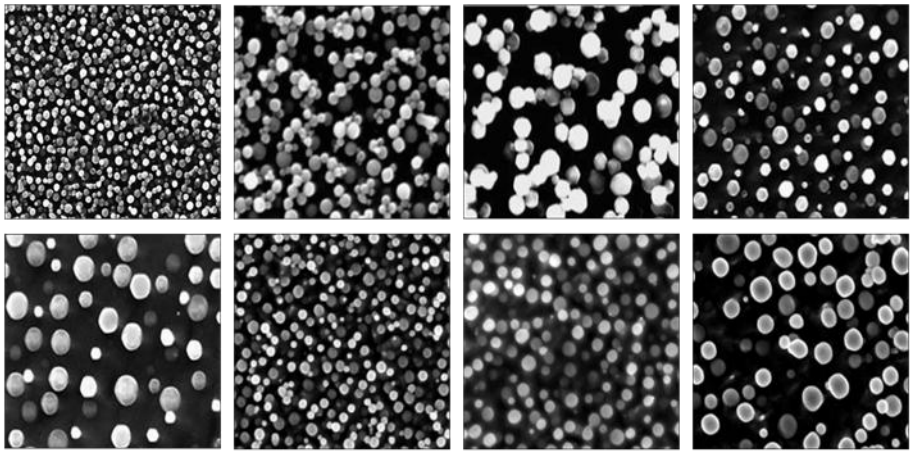


Figure S7. Microstructures extracted from [12].

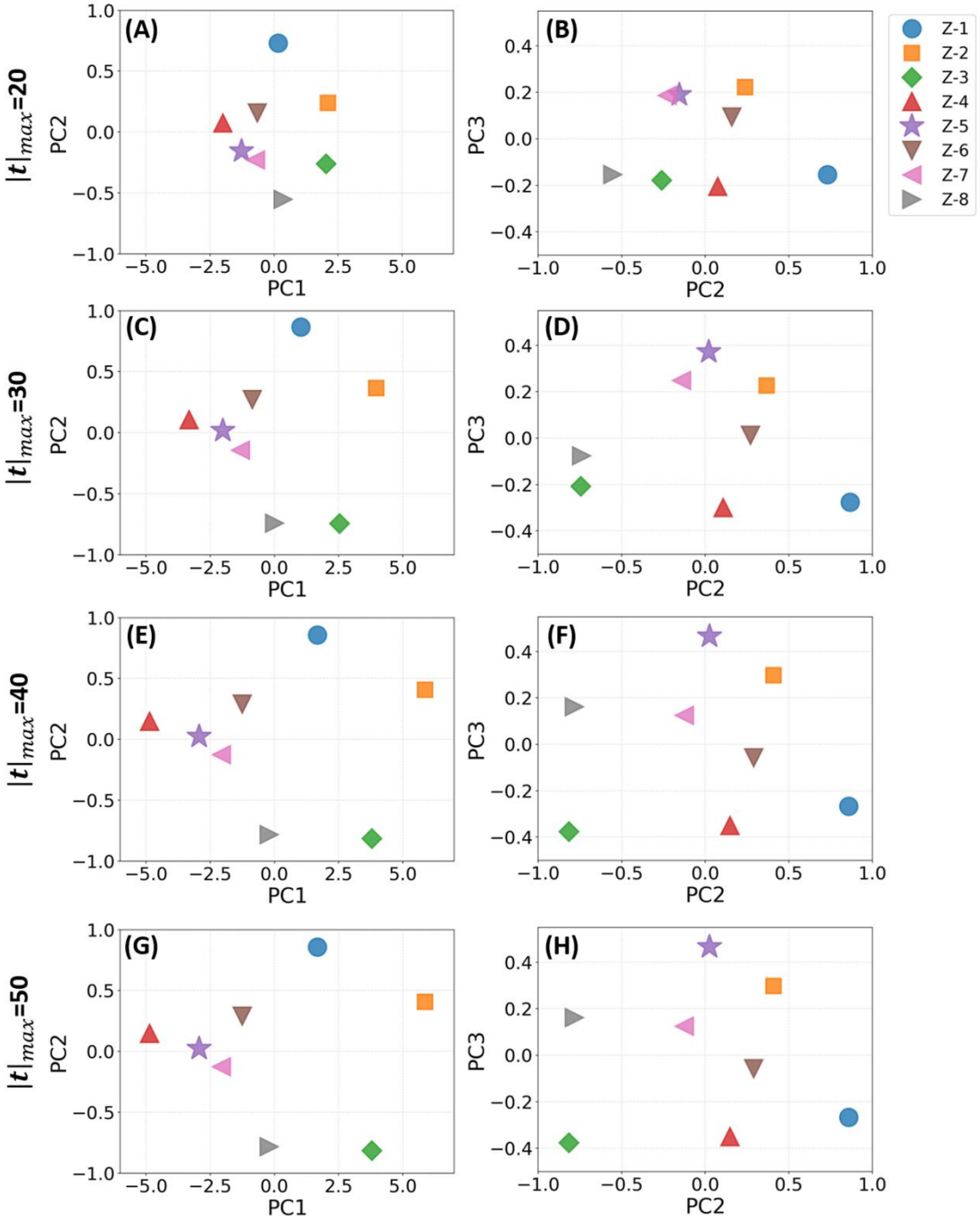


Figure S8. Variation of the distribution of the eight validated Ni-Fe-based superalloys in the first three PCs with different lengths of $|\boldsymbol{t}|$.

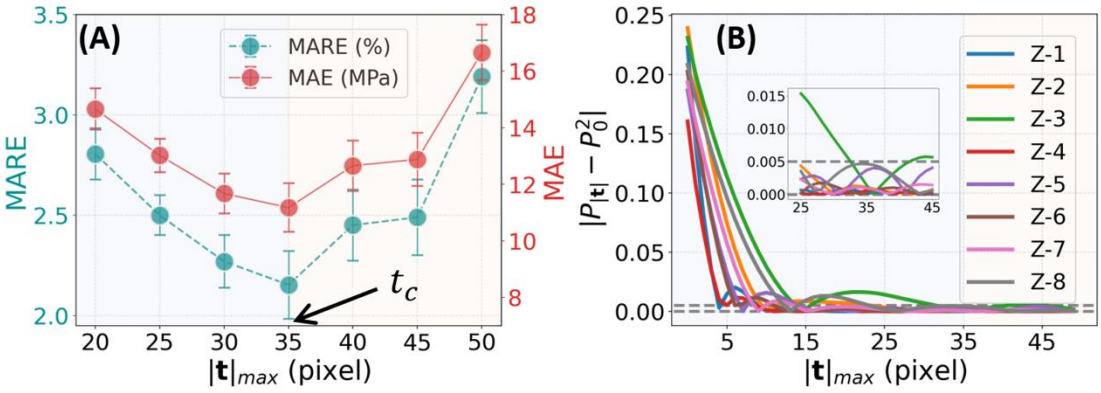


Figure S9. Variation of the predicted accuracy (MAE and MARE) of the yield strength σ and the tolerance factor ε with different length of $|\boldsymbol{t}|_{max}$ for the eight validated Ni-Fe-based superalloys: (A) MAE and MARE; and (B) ε. MAE, represent mean absolute error; MARE, mean absolute relative error.

S4. Analysis of the traditional evaluation metric

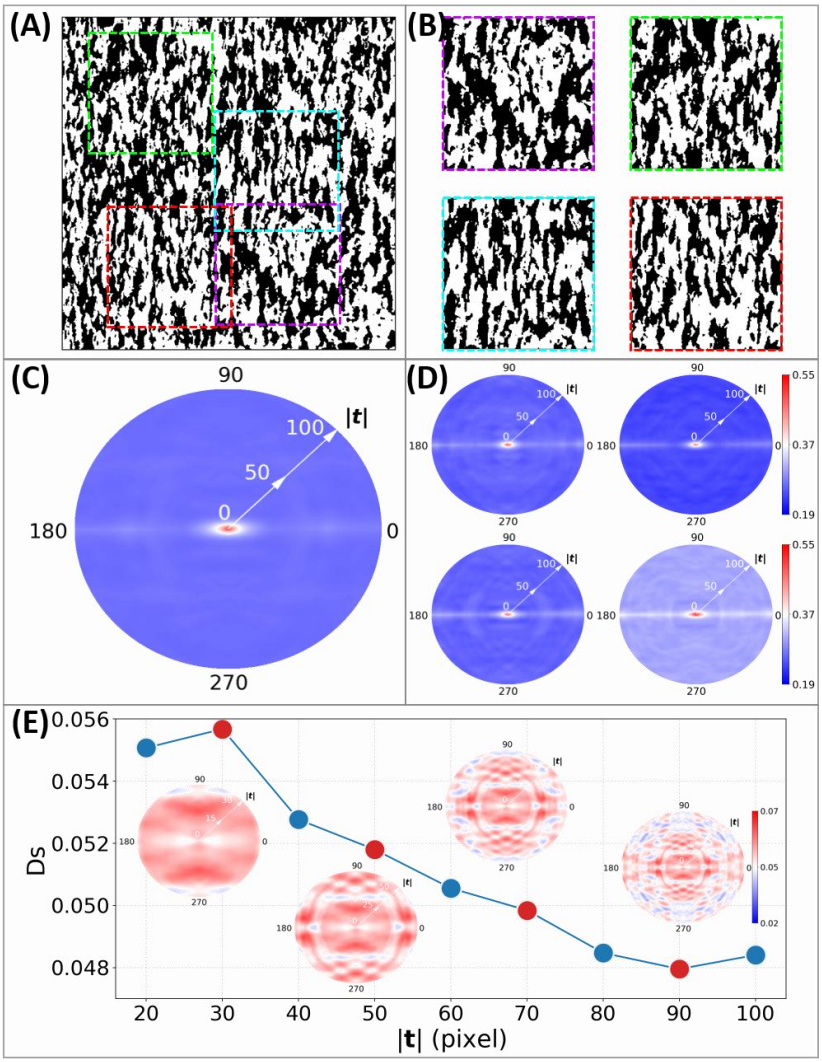


Figure S10. Ensemble averaged autocorrelations and scalar error D_s between the statistics of sub-domains and the averaged one. (A,C) Digital microstructure (550×550 pixels) and associated autocorrelations from [13]. (B,D) Sub-domains randomly extracted from (A) and associated autocorrelations. (E) D_s vs $|\boldsymbol{t}|$. The autocorrelations are all plotted with the same color scale to highlight the differences between each other. The embedded subgraphs in (E) visualize the absolute error between the first member in (D) and the ensemble average; they are plotted with the same color scale and can be distinguished by different lengths of $|\boldsymbol{t}|$.

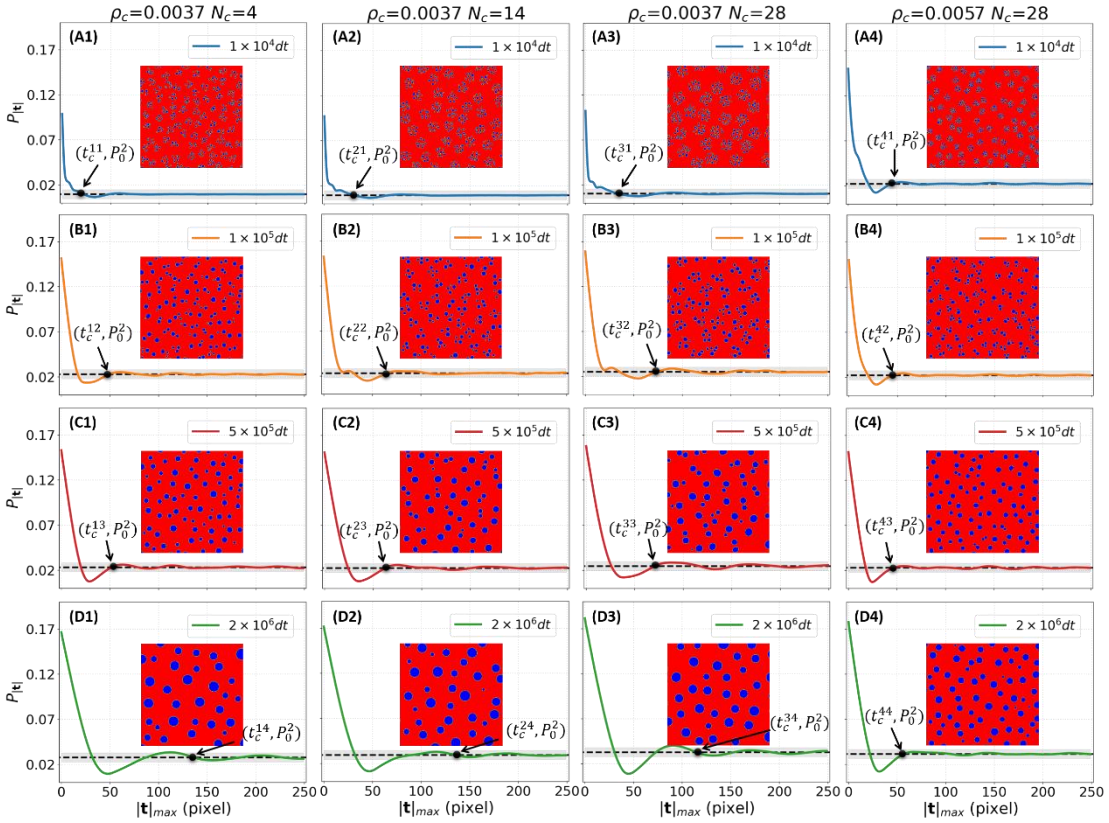


Figure S11. Variation of the pair correlation function $P_{|t|}$ with different length of $|t|_{max}$ for the evolving microstructures with four groups of parameters at: (A1–A4) 1×10^4dt; (B1–B4) 1×10^5dt; (C1–C4) 5×10^5dt; and (D1–D4) 2×10^6dt. The corresponding microstructures are embedded into each plot. A black arrow indicates the t_c for each microstructure, which is the minimal $|t|_{max}$ when $P_{|t|}$ value fluctuates around P_0^2 with a negligible level (i.e., gray shadow area representing the variation of $P_0^2 \pm 0.005$). A black dotted horizontal line labels the location of P_0^2. The superscripts i and j in t_c^{ij} represent the specific combination of parameters (N_c and ρ_c) and evolving time, respectively.

The evolving microstructures shown in Figure S11 are from Figure 3 in our previous work[14]. They were simulated by using different initial parameters: particle number N_c and particle density ρ_c in a particle cluster. The specific explanation and value of each parameter can be found in Section 2.2 and Table 1 in [14]. The original size of each simulated microstructure was resized to 500×500 pixel. Two-point statistics method was used to calculate the autocorrelations of the spherical particles

in the microstructures; the pair correlation function ($P_{|\boldsymbol{t}|}$) was then obtained from the autocorrelations, as shown in Figure S11. For each microstructure, coherence length t_c was equal to the minimal $|\boldsymbol{t}|_{max}$ when $P_{|\boldsymbol{t}|}$ value fluctuates around P_0^2 with a negligible level (i.e., gray shadow area representing the variation of $P_0^2 \pm 0.005$) and highlighted by the black arrow shown in Figure S11.

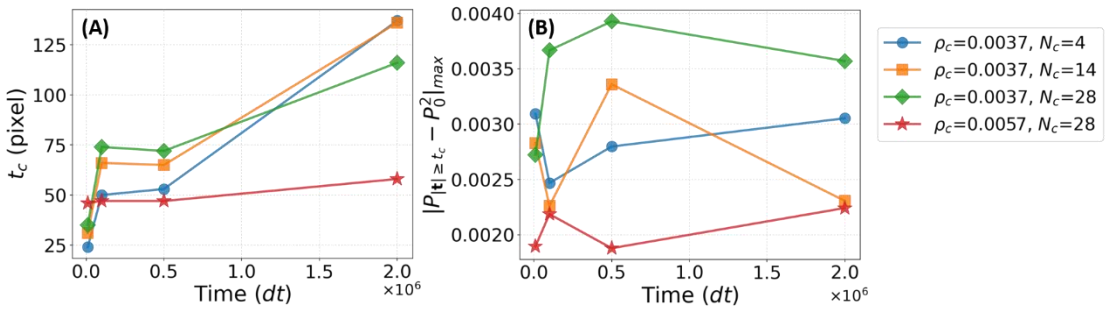


Figure S12. Coherence length t_c and $\left|P_{|\boldsymbol{t}|\geq t_c} - P_0^2\right|_{\max}$ *vs.* evolving time for the microstructures with four groups of parameters: (A) t_c *vs.* time; and (B) $\left|P_{|\boldsymbol{t}|\geq t_c} - P_0^2\right|_{\max}$ *vs.* time. t_c was determined according to the points labeled by black arrows shown in Figure S11, while $\left|P_{|\boldsymbol{t}|\geq t_c} - P_0^2\right|_{\max}$ is the maximum $\left|P_{|\boldsymbol{t}|} - P_0^2\right|$ when $|\boldsymbol{t}|$ is larger than t_c.

Figure S12A shows the variation of t_c with the increase of evolving time. For the first three combinations of ρ_c and N_c, we can see that t_c increases rapidly (Stage 1), then changes slightly (Stage 2), and finally reaches the maximum after a long-time evolution (Stage 3). At Stage 1, the solute diffusion mode changes from short-range diffusion to long-range diffusion, leading to the characteristic length t_c of the microstructures increasing rapidly. At Stage 2, the velocity of long-range diffusion is much smaller than that of the remaining short-range diffusion in each microstructure, thus t_c varies slightly. After a long time of evolving, sufficient long-range diffusion results in obvious coarsening of the particles, thus the large changes in t_c. In terms of the last combination of ρ_c and N_c, t_c barely changes during the first two stages because the short-range solute diffusion within each particle cluster plays a major role in grain growth, which can be demonstrated from the smaller average particle size and cluster-like particle distribution in the

microstructures compared with that in the microstructures with other three groups of combinations of ρ_c and N_c. After a long time of evolving, coarsening particles caused by long-range solute diffusion appear and the t_c of each microstructure increases slowly (Stage 3). Interestingly, even if t_c changes with different degrees over time, $\left|P_{|\boldsymbol{t}|\geq t_c} - P_0^2\right|$ is still at a low level (<0.005), indicating $\left|P_{|\boldsymbol{t}|\geq t_c} - P_0^2\right|$ is independent of evolving time.

Reference

[1] J. Eiken, B. Böttger, I. Steinbach, Multiphase-field approach for multicomponent alloys with extrapolation scheme for numerical application, Phys. Rev. E. 73 (2006) 066122. https://doi.org/10.1103/PhysRevE.73.066122.

[2] S.G. Kim, A phase-field model with antitrapping current for multicomponent alloys with arbitrary thermodynamic properties, Acta Mater. 55 (2007) 4391–4399. https://doi.org/10.1016/j.actamat.2007.04.004.

[3] S.R. Niezgoda, D.T. Fullwood, S.R. Kalidindi, Delineation of the space of 2-point correlations in a composite material system, Acta Mater. 56 (2008) 5285–5292. https://doi.org/10.1016/j.actamat.2008.07.005.

[4] P. Steinmetz, Y.C. Yabansu, J. Hötzer, M. Jainta, B. Nestler, S.R. Kalidindi, Analytics for microstructure datasets produced by phase-field simulations, Acta Mater. 103 (2016) 192–203. https://doi.org/10.1016/j.actamat.2015.09.047.

[5] Y.C. Yabansu, P. Steinmetz, J. Hötzer, S.R. Kalidindi, B. Nestler, Extraction of reduced-order process-structure linkages from phase-field simulations, Acta Mater. 124 (2017) 182–194. https://doi.org/10.1016/j.actamat.2016.10.071.

[6] A. Choudhury, Y.C. Yabansu, S.R. Kalidindi, A. Dennstedt, Quantification and classification of microstructures in ternary eutectic alloys using 2-point spatial correlations and principal component analyses, Acta Mater. 110 (2016) 131–141. https://doi.org/10.1016/j.actamat.2016.03.010.

[7] I. Steinbach, Effect of interface anisotropy on spacing selection in constrained dendrite growth, Acta Mater. 56 (2008) 4965–4971. https://doi.org/10.1016/j.actamat.2008.06.009.

[8] J. Li, Z. Wang, Y. Wang, J. Wang, Phase-field study of competitive dendritic growth of converging grains during directional solidification, Acta Mater. 60 (2012) 1478–1493. https://doi.org/10.1016/j.actamat.2011.11.037.

[9] C. Guo, J. Li, H. Yu, Z. Wang, X. Lin, J. Wang, Branching-induced grain boundary evolution during directional solidification of columnar dendritic grains, Acta Mater. 136 (2017) 148–163. https://doi.org/10.1016/j.actamat.2017.07.002.

[10] Y. Wang, J. Li, H. Yang, J. Ma, D. Cui, Y. Gao, Z. Wang, J. Wang, The formation mechanism of special globular surface grain during the solidification of laser surface remelted near β titanium alloys, Comput. Mater. Sci. 191 (2021) 110353. https://doi.org/10.1016/j.commatsci.2021.110353.

[11] A. Farzadi, M. Do-Quang, S. Serajzadeh, A.H. Kokabi, G. Amberg, Phase-field simulation of weld solidification microstructure in an Al–Cu alloy, Model. Simul. Mater. Sci. Eng. 16 (2008) 065005. https://doi.org/10.1088/0965-0393/16/6/065005.

[12] Z. Zhong, Y. Gu, Y. Yuan, Microstructural stability and mechanical properties of a newly developed Ni–Fe-base superalloy, Mater. Sci. Eng. A. 622 (2015) 101–107. https://doi.org/10.1016/j.msea.2014.11.010.

[13] S.R. Niezgoda, D.M. Turner, D.T. Fullwood, S.R. Kalidindi, Optimized structure based representative volume element sets reflecting the ensemble-averaged 2-point statistics, Acta Mater. 58 (2010) 4432–4445. https://doi.org/10.1016/j.actamat.2010.04.041.

[14] Y. Wang, J. Li, L. Zhang, Z. Wang, J. Wang, Phase-field study on the effect of initial particle aggregation on the transient coarsening behaviors, Model. Simul. Mater. Sci. Eng. 28 (2020) 075007. https://doi.org/10.1088/1361-651X/abaff5.