A Multiscale Generative Model to Understand Disorder in Domain Boundaries

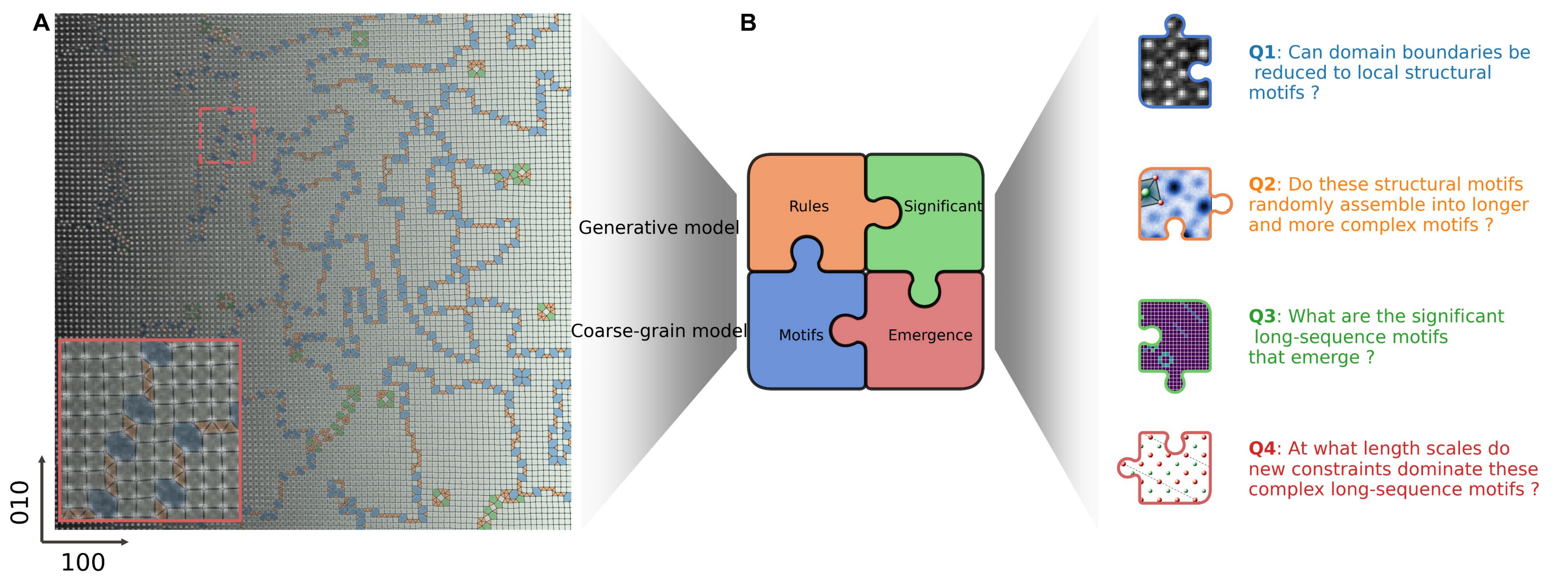
Jiadong Dan[1,2], Moaz Waqar[3], Stephen J. Pennycook[4,5] and N. Duane Loh[1,2]

1. NUS Centre for Bioimaging Sciences, National University of Singapore, Singapore. 2. Dept. of Biological Science, National University of Singapore, Singapore. 3. Dept. of Materials Science and Engineering, National University of Singapore, Singapore. 4. Dept. of Materials Science and Engineering, University of Tennessee, Knoxville, TN, USA. 5. School of Physical Sciences, University of Chinese Academy of Sciences, Beijing, China.

Introduction

A continuing challenge in atomic resolution microscopy is to identify significant structural motifs and their assembly rules in synthesized materials with limited observations[1]. Here we propose and validate a simple and effective hybrid generative model capable of predicting unseen domain boundaries in a potassium sodium niobate film from only a small number of observations, without expensive first-principles calculations or atomistic simulations of domain growth[2]. Our results demonstrate that complicated domain boundary structures can arise from simple interpretable local rules, played out probabilistically. We also found new significant tileable boundary motifs and evidence that our system creates domain boundaries with the highest entropy. More broadly, our work shows that simple yet interpretable machine learning models can help us describe and understand the nature and origin of disorder in complex materials, thus improving functional materials design.

Key Challenges



A Generative model

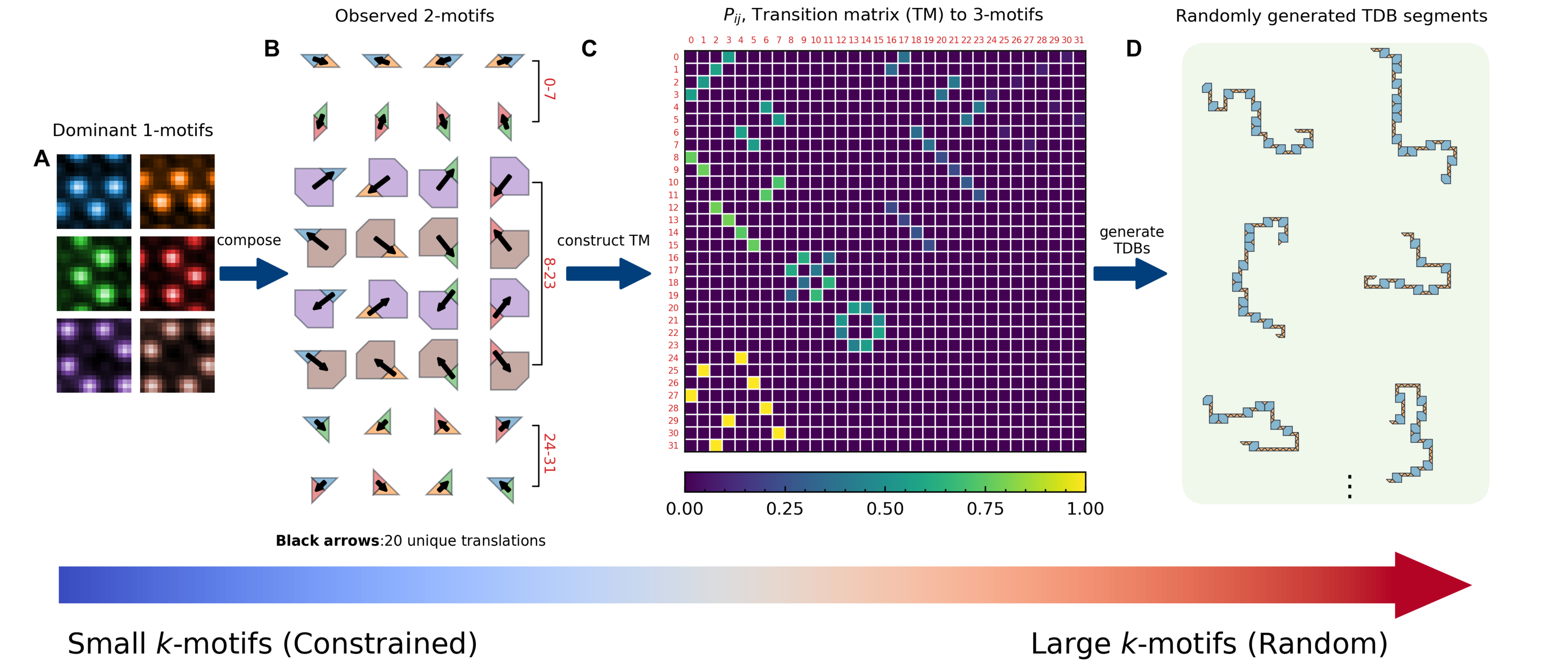


Fig. 1 How the constrained *k*-Motif hierarchy leads to a probabilistic generative model. (**A**) The six dominant 1-motifs identified and averaged from ADF-STEM images. There are 4 orientations for triangular 1-motifs (top) and two orientations for hexagonal 1-motifs (bottom). (**B**) The 32 unique observed 2-motifs that can be formed by the six 1-motifs. Based on Euclidean isometry (rigid transformation), they can be divided into three groups: 0-7, 8-23, and 24-31. (**C**) The transition matrix is estimated from our experimental data, and *Pij* is the probability of 2-motif *i* connecting to 2-motif *j*, where 0≤*i*≤31 and 0≤*j*≤31. (**D**) A collection of synthetic domain boundary samples generated from our model. From left to right, it shows a hierarchical structure of forming random complex domain boundaries from low-level constrained structural motifs.

A Random nucleation model

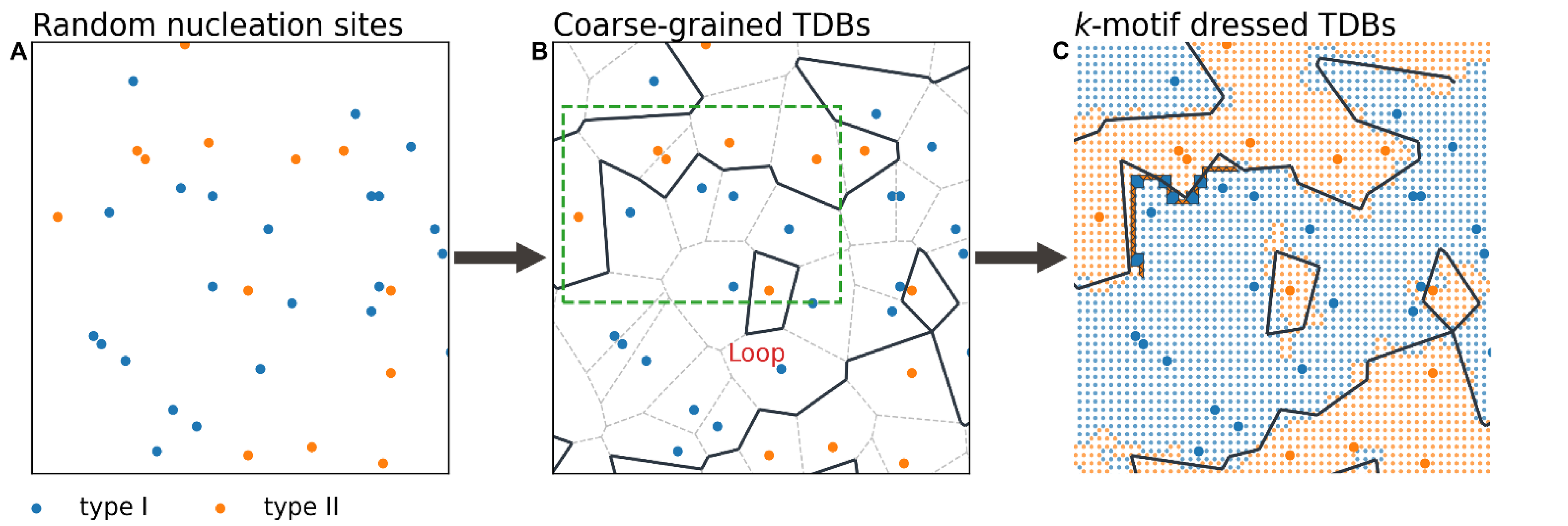


Fig. 2 Generating coarse-grained domain boundaries using a random nucleation model. (**A**) Nucleation sites of two types of translation domains (*i.e.*, I, II) are randomly distributed across the substrate. Blue dots belong to the lattice from type I and orange dots are from type II. (**B**) Voronoi tessellation of nucleation sites in panel (A); solid edges are the coarse-grained TDBs between different domain types while dashed edges are those between the same type. (**C**) We dress the coarse-grained TDBs in panel (B) with a possible arrangement of I, II translation domains. Here the edges in the latter are replaced by *k*-motifs generated from the Markovian Transition matrix.

Model Validations

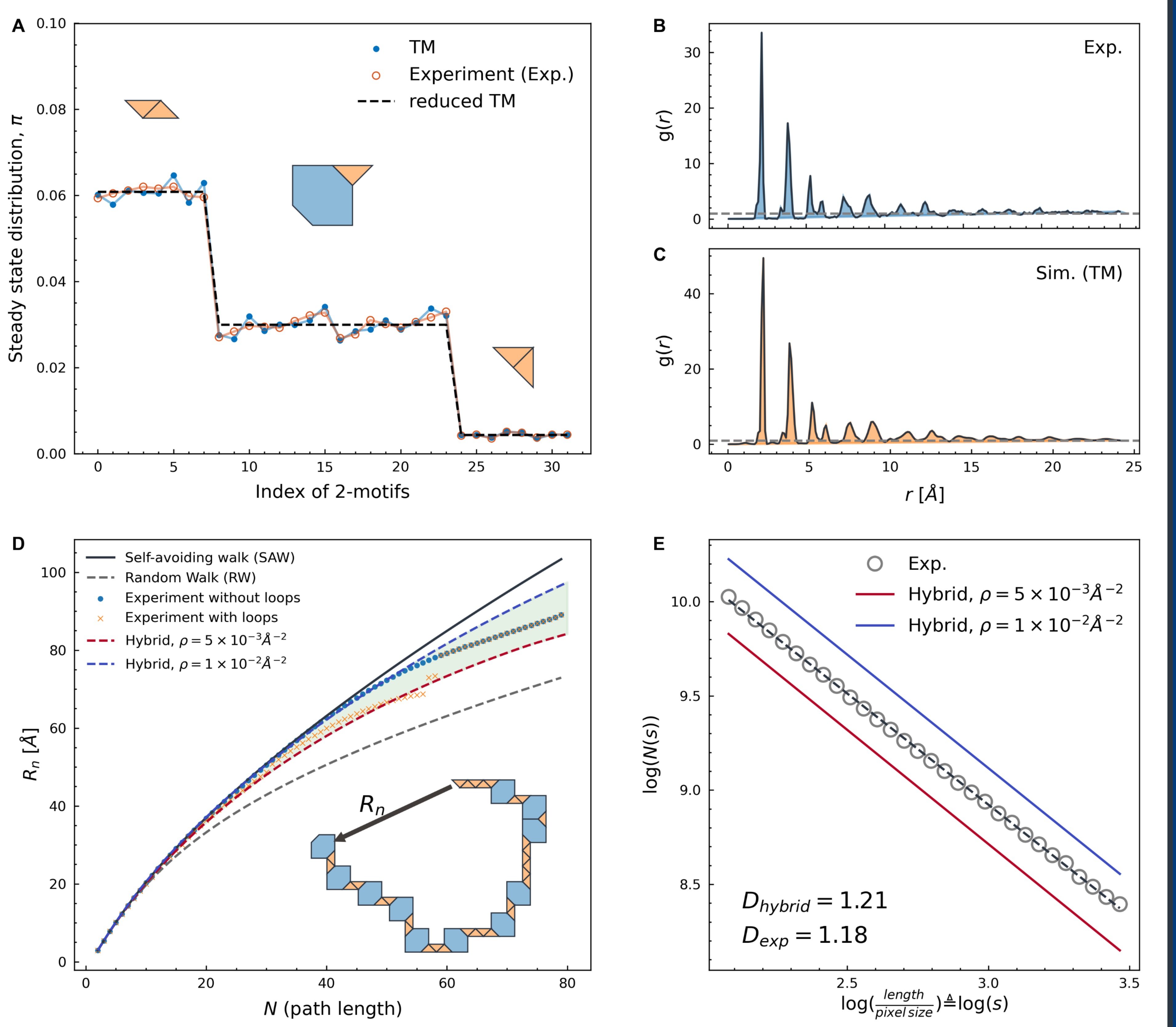


Fig. 3 Model evaluation and validation. (**A**) The frequency distribution of thirty-two 2-motif states. This distribution shows three plateaus, where above each plateau we show its characteristic 2-motif (up to isometries). (**B**) The radial distribution function of all 1-motifs from experiments. (**C**) The radial distribution function of all 1-motifs from synthetic samples drawn using our hybrid model. (**D**) Comparison of end-to-end distance, R_n (exemplified by inset), as a function of motif path length N in different cases. (**E**) Fractal dimension of the domain boundary assemblies from the experiments (gray circles) and synthetic samples (blue and red lines). Both experiments and simulations show a power-law scaling with a fractal dimension of ~1.2.

Conclusions

We show how from a handful of simple rules, arbitrarily complex domain boundaries can emerge in potassium-sodium niobate piezoelectric films. We capture these rules with an interpretable, probabilistic, generative model. Our model is in stark contrast with opaque and complex generative machine models (e.g., Generative Adversarial Networks, GANs) that require a huge amount of data to train, or models that require expensive first-principles calculations. Interpretable, generative models like ours can make sense of this randomness beyond what is practically observable, and have the potential to catalyze new directions for a common language to understand complex disordered materials[3,4].

Acknowledgement

N.D.L acknowledges funding support from the National Research Foundation (grant number NRF-CRP16-2015-05), and the NUS Early Career award (A-0004744-00-00). This project is supported by the Eric and Wendy Schmidt AI in Science Postdoctoral Fellowship, a Schmidt Futures program.

References

1. Dan, J. et al., *Science Advances* **8** (2022), doi: 10.1126/sciadv.abk1005
2. Dan, J. et al, arXiv preprint arXiv:2305.18325 (2023).
3. Huang, P.Y. et al. *Science* 342.6155 (2013): 224-227.
4. Toh, C-T. et al. *Nature* 577.7789 (2020): 199-203.

Supplementary Materials →

A Multiscale Generative Model to Understand Disorder in Domain Boundaries

Jiadong Dan[1,2], Moaz Waqar[3], Stephen J. Pennycook[4,5] and N. Duane Loh[1,2]

1. NUS Centre for Bioimaging Sciences, National University of Singapore, Singapore. 2. Dept. of Biological Science, National University of Singapore, Singapore. 3. Dept. of Materials Science and Engineering, National University of Singapore, Singapore. 4. Dept. of Materials Science and Engineering, University of Tennessee, Knoxville, TN, USA. 5. School of Physical Sciences, University of Chinese Academy of Sciences, Beijing, China.

Significant tileable long motifs at domain boundaries

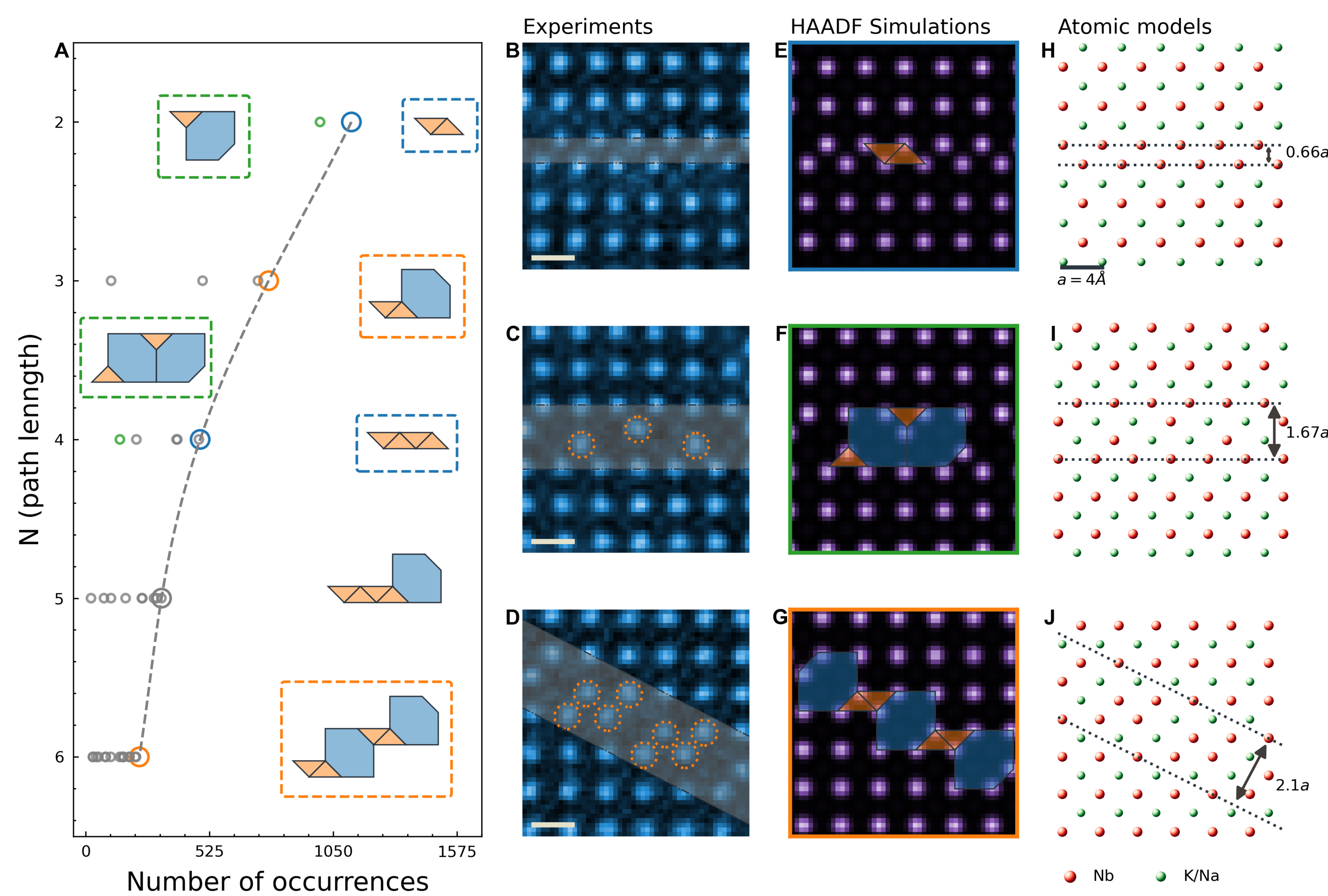


Fig. 4 Significant tileable long sequences of *k*-motifs. (**A**) Number of occurrences of tileable *k*-motifs that are statistically significant in Fig. 1A. Besides a simple chain of triangle-only motifs (blue dashed boxes), we observed a significant number of previously neglected longer sequences that contain hexagonal motifs (green or orange dashed boxes). (**B-D**) HAADF-STEM images showing single examples of the most frequently occurring, tileable, longer sequence motifs. (**E-F**) Multislice simulations of structures that match those in b-d respectively. (**H-J**) Atomic models of regions in (B-D) respectively.

The system maximizes the configurational entropy of domain boundaries

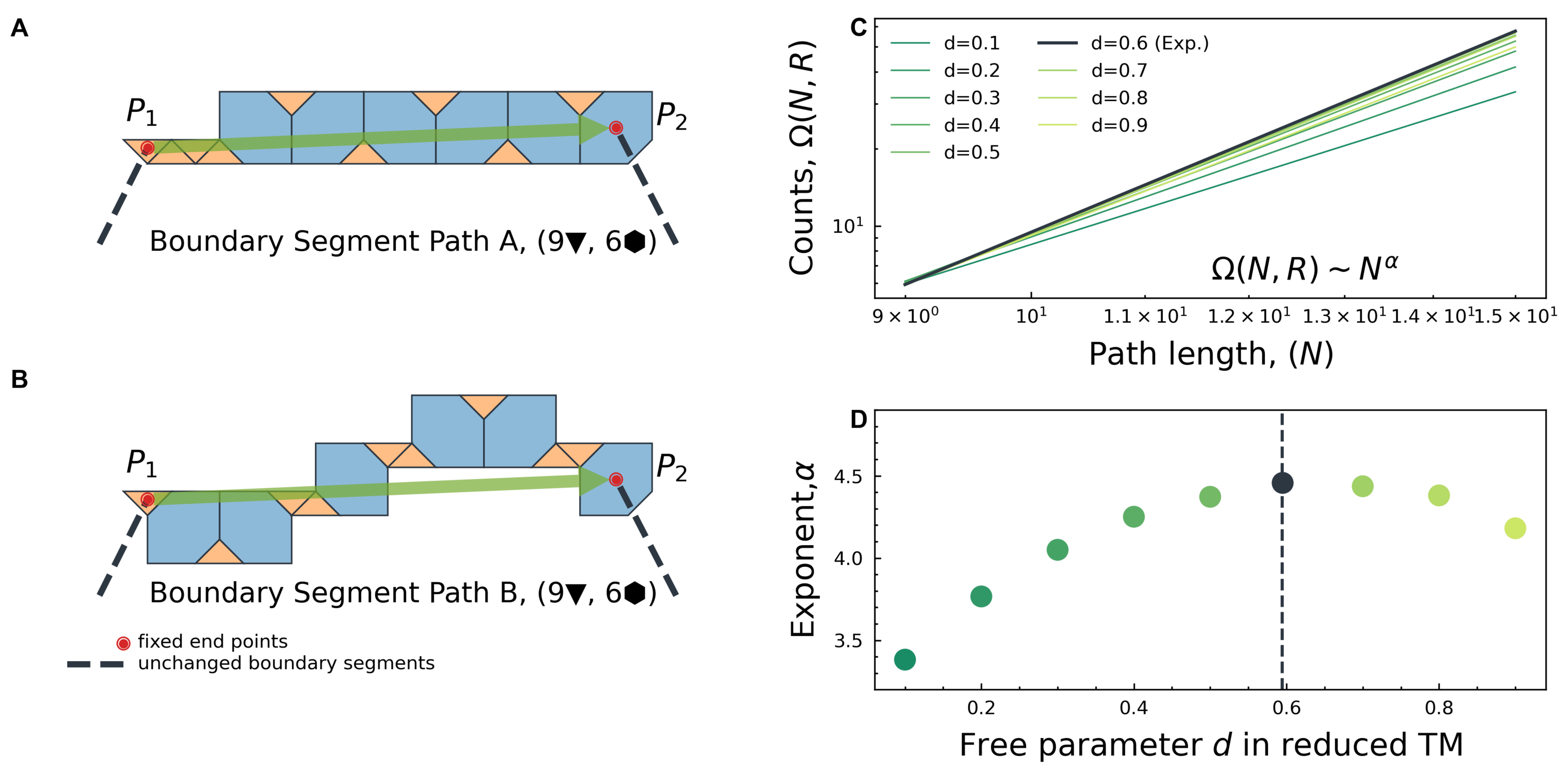


Fig. 5 Estimating configurational entropy of domain boundaries. (**A, B**) Two boundary segments that begin and end at equivalent fixed-end positions (red circles). Both domain boundaries have 9 triangular motifs and 6 hexagonal motifs. (**C**) The average number of configurational microstates of boundary segments, $\Omega(N,R)$, increases with the segments' path length as a power law, shown for different values of the free parameter d in the reduced transition matrix. (**D**) The system maximizes the power law exponent α when the free parameter $d \approx 0.6$ (dark gray circle), which was observed experimentally (gray dash line).

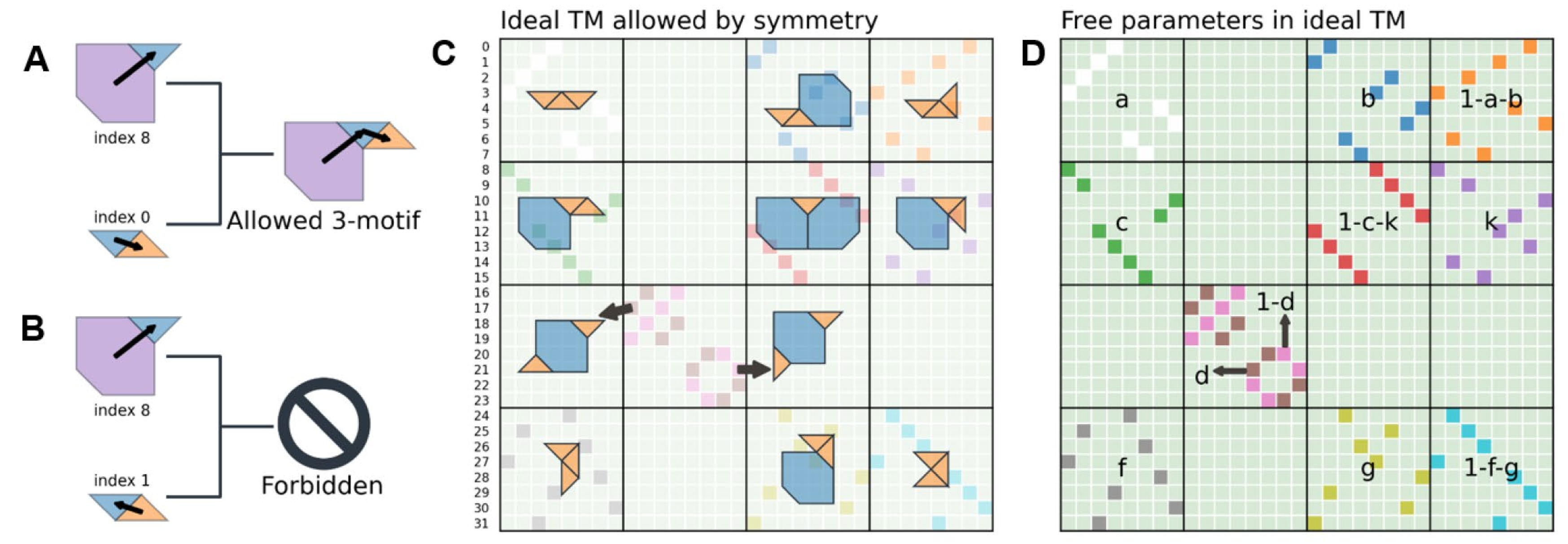


Fig. 6 The ideal generative model reduced by the isometries of 3-motifs. The transition matrix that encodes the local rules is sparse since composing 2-motifs is not arbitrary. For example, panel **(A)** shows 2-motifs indexed by 8 and 0 can combine to form an allowed 3-motif, while panel **(B)** illustrates the combination of 2-motifs indexed by 8 and 1 is prohibited. In total, there are 88 permissible 3-motifs, which means the transition matrix has 88 valid inputs. These 88 permissible 3-motifs can be further classified into 11 classes because rotation and reflection are isometries. **(C)** 11 Unique color blocks in the ideal TM mark the resultant 3-motifs overlaid on top. **(D)** In terms of model parametrization, the ideal system is fully determined by {a, b, c, k, d, f, g}.

Reduced by isometries, the steady-state distribution of 2-motifs should take the following form:

$$\pi = [x, x, \cdots, x, y, y, \cdots, y, z, z, \cdots, z].$$

There are 8 consecutive x, 16 consecutive y, and 8 consecutive z, therefore π is a row vector of a length 32. If the transition probability is P, then the following relation holds,

$$\boldsymbol{P\pi} = \boldsymbol{\pi}.$$

We can obtain a system of linear equations as follow,

$$\begin{gathered} ax + cy + fz = x, \\ bx + (1-c-k)y + gz = y, \\ 8x + 16y + 8z = 1, \\ y(1-d) + yd = y, \\ (1-a-b)x + ky + (1-f-g)z = z. \end{gathered}$$

By solving this linear equation system, we can obtain x, y and z.

$$x = \frac{cf + cg + fk}{8(-ac - 2ag - ak - bc + 2bf + cf + cg + c + fk + 2g + k)}$$
$$y = \frac{-ag + bf + g}{8(-ac - 2ag - ak - bc + 2bf + cf + cg + c + fk + 2g + k)}$$
$$z = \frac{-ac - ak - bc + c + k}{8(-ac - 2ag - ak - bc + 2bf + cf + cg + c + fk + 2g + k)}$$

From our observations, k=0, g=0, and f=1, they can further simplify to forms as follow,

$$x = \frac{c}{16b+16c-8ac-8bc},$$
$$y = \frac{b}{16b+16c-8ac-8bc},$$
$$z = \frac{c(1-a-b)}{16b+16c-8ac-8bc}.$$

Conclusions

We show how from a handful of simple rules, arbitrarily complex domain boundaries can emerge in potassium-sodium niobate piezoelectric films. We capture these rules with an interpretable, probabilistic, generative model. Our model is in stark contrast with opaque and complex generative machine models (e.g., Generative Adversarial Networks, GANs) that require a huge amount of data to train, or models that require expensive first-principles calculations. Interpretable, generative models like ours can make sense of this randomness beyond what is practically observable, and have the potential to catalyze new directions for a common language to understand complex disordered materials[3,4].

Acknowledgement

N.D.L acknowledges funding support from the National Research Foundation (grant number NRF-CRP16-2015-05), and the NUS Early Career award (A-0004744-00-00). This project is supported by the Eric and Wendy Schmidt AI in Science Postdoctoral Fellowship, a Schmidt Futures program.

References

1. Dan, J. et al., *Science Advances* **8** (2022), doi: 10.1126/sciadv.abk1005
2. Dan, J. et al, arXiv preprint arXiv:2305.18325 (2023).
3. Huang, P.Y. et al. *Science* 342.6155 (2013): 224-227.
4. Toh, C-T. et al. *Nature* 577.7789 (2020): 199-203.

Supplementary Materials →