Machine Learning the Entropy to Estimate Free-Energy Differences without Sampling Transitions

Yamin Ben-Shimon[1,2], Barak Hirshberg[3,4,2], Yohai Bar-Sinai[1,2]

[1]School of Physics, [2]Center for Physics and Chemistry of Living Systems,
[3]School of Chemistry, [4]Center for Computational Molecular and Materials Science, Tel Aviv University, Tel Aviv 6997801, Israel.

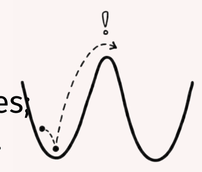
1. Motivation: Free-Energy Differences Without Rare Transition Sampling

- Standard simulation strategies, e.g. Molecular Dynamics (MD), suffer from poor sampling near phase transitions.
- Enhanced-sampling methods can induce transitions but become impractical when barriers are large and good collective variables (CVs) are unknown.
- Many phase transitions are controlled by enthalpy and entropy: $\Delta g = \Delta h - T\Delta s$.
- In simulations, Δh is easy to estimate, but Δs is hard because it is an ensemble property and depends on the partition function.

Key idea: Machine learn the entropy of each phase directly from short, separate equilibrium simulations.

Transition Sampling is Hard

High energy barriers ⇒ rare events.
Require trajectories that cross phases;
enhanced sampling needs good CVs.



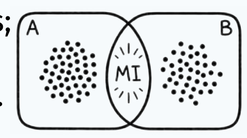
Short Single-Phase Simulation

Sample each phase separately.
Standard equilibrium MD;
no transition path; no CV engineering.



Learn Entropy via Mutual Information

Decompose entropy into MI terms;
train a network to estimate MI;
get Δs and Δg without transitions.



2. Entropy From Multi-Scale Mutual Information

Entropy decomposition. Recursively bisect the volume into two halves. Using

$$S(A,B) = S(A) + S(B) - \mathrm{MI}(A,B),$$

the entropy density becomes a sum of MI contributions across length scales:

$$s(X_0) = s(X_m) - \frac{1}{2^m}\sum_{k=1}^{m}\frac{\mathrm{MI}(X_k)}{V_k},$$

(X_0 : full system, X_k : subvolumes at level k, V_k : their volume).

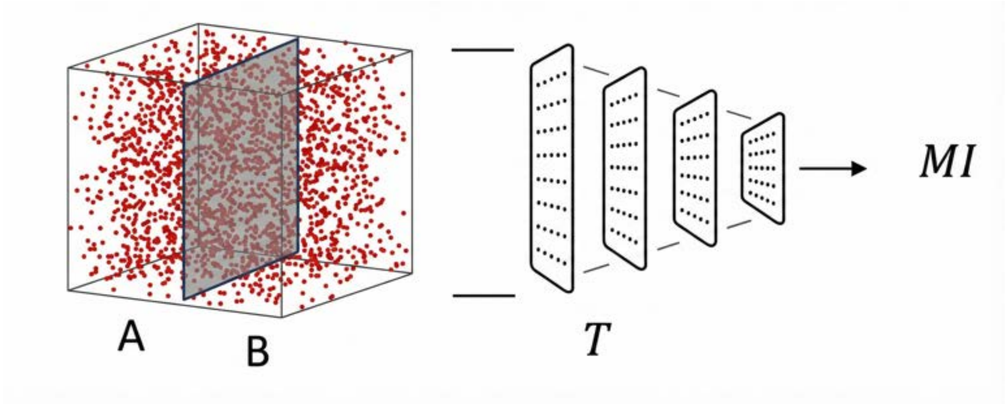
For bulk systems, boundary MI scales with interface area (area-law), giving a thermodynamic-limit correction:

$$s_\infty \approx s(X_0) - 3\frac{\mathrm{MI}(X_0)}{V_0}.$$

Estimating MI with a neural network (MINE). With joint samples P_{XY} and independent pairs $P_X \otimes P_Y$ (shuffled), MI is written as a variational objective over functions T:

$$\mathrm{MI}(X,Y) = \sup_T \left(\mathbb{E}_{P_{XY}}[T] - \log \mathbb{E}_{P_X\otimes P_Y}[e^T]\right),$$

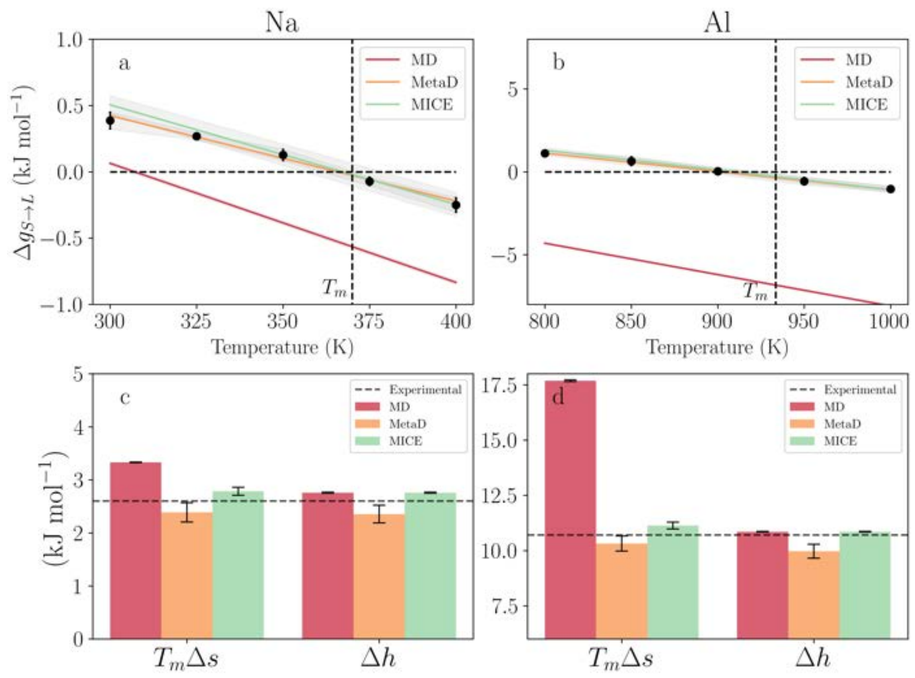
where T is parameterized by a 3D CNN trained on voxel grids of sub-volumes.



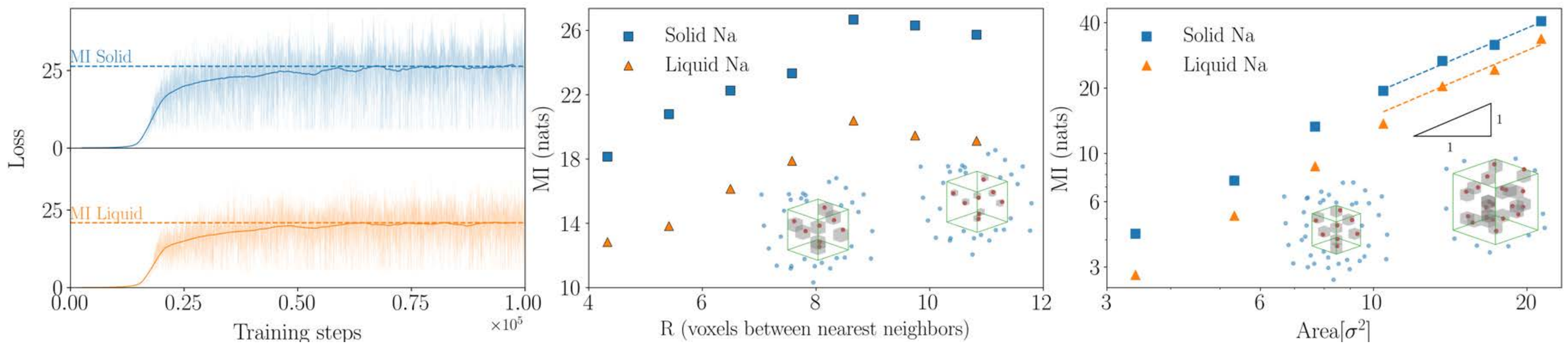
3. Practical Workflow

1. Run unbiased MD for each phase separately (solid and liquid).
2. Sample random atom configurations; represent each as a binary 3D voxel grid.
3. For each scale k, train MINE to estimate $\mathrm{MI}(X_k)$ using original samples for joint and artificially stitched samples for marginals.
4. Combine MI terms to obtain absolute entropy density s for each phase; then compute $\Delta g(T) = \Delta h - T\,\Delta s$ and find T_m from $\Delta g(T_m) = 0$.

4. Accurate Melting Temperatures From Single-Phase Data



5. MI training / resolution / area-law trends (Na)



Left: *MINE optimization for solid (blue) and liquid (orange) Na near T_m. Bold curves are running averages (5000 steps) that highlight convergence; dashed lines indicate the final MI estimates averaged over the last 10^4 training steps.*

Middle: *Estimated MI for a fixed physical subvolume as voxel resolution is increased. The estimate reaches a plateau at high resolution, indicating that the discretization is fine enough to capture the relevant local structure.*

Right: *MI between two halves of a cubic subsystem as a function of the interface area. The near-linear scaling is consistent with an area law and motivates the thermodynamic-limit correction used to estimate bulk entropies.*

6. Takeaways

- Pair-entropy (s_2) as an estimate for entropy from the unbiased **MD** data is not sufficient and can yield large errors in T_m.
- Well-Tempered Metadynamics (**MetaD**) can be accurate but requires CVs and explicit phase transitions.
- Machine-Learning Iterative Calculation of Entropy (**MICE**) provides a direct entropy estimate from MI across scales.
- Trained network on Na transfers to Al without hyperparameter tuning.

Future work: improved molecular representations as graphs and equivariant models, and lower-bias MI estimators could further improve accuracy.

Reference: Y. Ben-Shimon, B. Hirshberg, Y. Bar-Sinai (2025). *Machine Learning the Entropy to Estimate Free Energy Differences without Sampling Transitions*. arXiv:2510.24930.
Related: Nir et al., PNAS (2020) (MICE); Belghazi et al., ICML (2018) (MINE).

מדע בונה חברה
The Raymond and Beverly Sackler Faculty of Exact Sciences
Tel Aviv University
ISEF אייסף
ISRAEL SCHOLARSHIP EDUCATION FOUNDATION
הקרן הבינלאומית לחינוך ע"ר