Supplementary Information: Gradient-based Active Learning For Intelligent Discovery Of Colloidal Phase Diagrams

Sumeet Vadhavkar, Pradeep Bajracharya, Md Shakil Zaman, Poornima Padmanabhan, and Linwei Wang

1 Implementation of the MD simulation

Colloidal particles are modeled as spheres in an implicit solvent. The simulations are performed in dimensionless units where the mass m of each colloidal particle, its radius a, and the thermal energy in the system given by the product of the Boltzmann constant and absolute temperature k_BT are each set to unity. This allows broad interpretability of our results by rescaling the appropriate mass, length, and energy scales for any experimental system of interest. The interaction between two colloidal particles is given by the Morse potential which represents a steep repulsion due to excluded volume at short separation, followed by attraction over a specified range of separation, and ultimately no interaction at farther separation. Mathematically, the Morse potential is given by,

$$\mathcal{V}(r_{ij}) = V_0 \left[e^{-2\kappa(r_{ij}-2a)} - 2e^{-\kappa(r_{ij}-2a)} \right], \tag{S1}$$

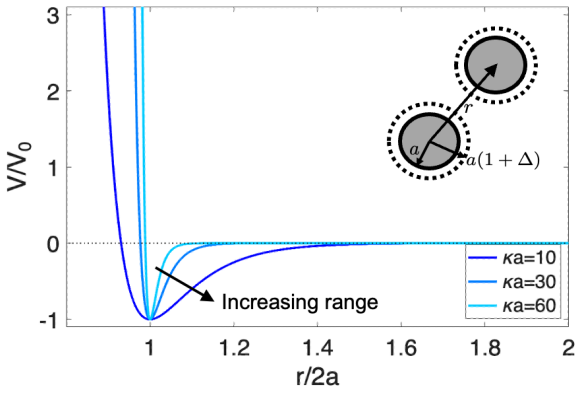


Figure S1: Morse potential profiles, rescaled by V_0, for various ranges tunable by parameter κa. Graphical inset visualizes two colloidal particles separated by r which will experience attractions when the dotted boundaries overlap and repulsions when the solid boundaries overlap.

where r_{ij} is the separation between the two particles labeled i and j, a is the radius of each colloid, V_0 is the absolute value of the lowest point of the potential denoting the

strength of attraction, and κ is an inverse length-scale parameter that sets the effective range of the potential. Its functional form is shown in Figure S1. Equivalently, the non-dimensional parameters of interest are the ratio of the strength of Morse potential to thermal energy, $V^* \equiv V_0/k_BT$ and the product κa. The strength of the Morse potential is allowed to vary between $0.1 \leq V^* \leq 5$, which captures distinct behaviors from very weak bonds to moderately strong bonds; the latter inducing phase separation. The range of attraction in this study is defined as the farthest distance where the potential energy is 1% of the maximum depth, and inversely related to κa. The parameter is varied between $10 \leq \kappa a \leq 60$, translating to an attraction range approximately between 4% and 25% of particle size. We note that although the repulsion strength does vary with κa, all values of κa in this study produce a repulsion that is sufficiently steep to mimic hard repulsion upon contact. The volume fraction of the colloids varied between $0.01 \leq \phi \leq 0.50$.

Molecular dynamics simulations are performed in the NVT ensemble using LAMMPS molecular dynamics simulator, using resources at Research Computing at RIT.[1–3] In addition to the forces felt due to the interparticle potential, the colloidal particles undergo Brownian motion due to the implementation of the Langevin thermostat at a temperature scale of $k_BT = 1.0$. The number of particles simulated was $N = 62,500$ in a cubic simulation box of volume, $V = \dfrac{4\pi a^3 N}{3\phi}$, with periodic boundary conditions to model an infinitely large colloidal solution in the thermodynamic limit. To inhibit crystallization at high volume fractions, the particles are polydisperse with an equal number of particles with radii $0.9a$, $1.0a$, and $1.1a$. The equations of motion are integrated with a dimensionless timestep of 0.001 for $\phi \leq 0.4$ and reduced to 0.0005 for crowded environments when $0.4 < \phi \leq 0.5$. Thermostat parameters are chosen such that one Brownian time is equivalent to 25 timesteps.

62,500 particles are initiated on an face-centered cubic lattice and equilibrated at low attraction strengths $V_0/k_BT = 0.1$ for 40 Brownian times to form a dispersion, following which an instantaneous quench is performed to attain the target interaction strength. Simulations at the target interaction strength were performed for 4000 Brownian times. This timescale is sufficient to allow sufficient phase separation to take place, including non-equilibrium routes to phase separation.

The radial distribution function $g(r)$ measures the local structure around a typical colloidal particle, and is defined as the number density within a thin shell of width Δr around the particle, divided by the bulk number density and computed with a bin radius $\Delta r = 0.04a$. The average number of nearest neighbors for each particle N_{nn} can be found by integrating the radial distribution function multiplied by the bulk volume fraction from the center of the particle to the maximum range of the interaction

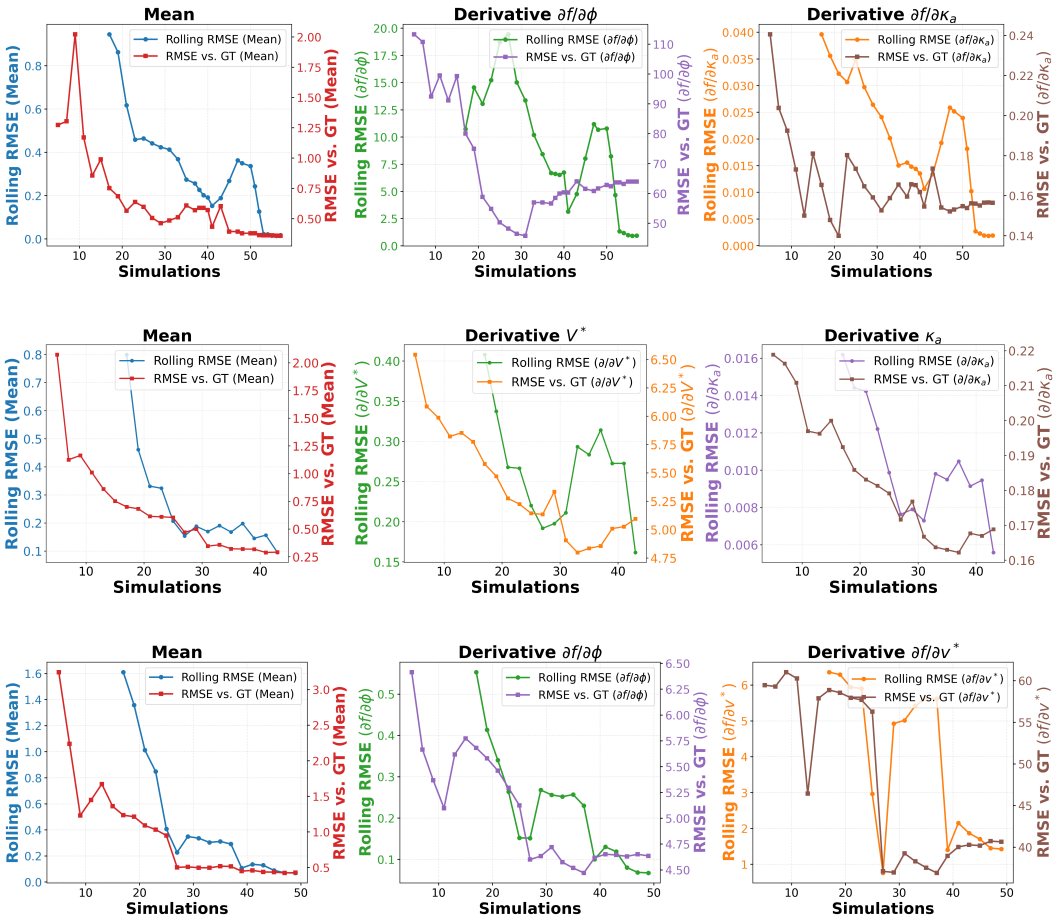


Figure S2: **Convergence diagnostics of rolling RMSE across mean and gradient dimensions.** Top to bottom: $(V^*, \kappa a)$ slice at $V^* = 4$, $(V^*, \kappa a)$ slice at $\phi = 0.05$, and (V^*, ϕ) slice at $\kappa a = 50$. Each panel shows the evolution of the rolling RMSE of the GP mean and its derivatives as active learning progresses.

potential in spherical coordinates,

$$N_{nn} = \int_0^{1.15a} 24\phi g(r) r^2 dr \tag{S2}$$

The integral is performed using trapezoid integration. The N_{nn} can take continuous values from 0 to 12. The transition from dilute to condensed phases are generally considered to occur beyond $N_{nn} \geq 6$. The function being learned is the value of N_{nn} at each point in the 1-dimensional or 2-dimensional phase diagram.

2 Stopping Criterion Analysis

Figure S2 summarizes the behavior of the rolling root mean square error (RMSE) for the Gaussian process (GP) mean and its first-order derivatives as active learning progresses.

Across all three planes, we observe a consistent pattern: once the rolling RMSE of the mean prediction falls below the threshold of $\epsilon = 0.2$, the deviation from the ground truth is also small, and the corresponding RMSE values for the derivatives in both coordinate directions likewise remain low. This joint stabilization indicates that the GP has formed a reliable local approximation of the underlying phase landscape, and that additional acquisitions might yield diminishing improvements. These observations support the use of the rolling-mean criterion as a practical means of determining when active learning has converged to a satisfactory solution.

3 Results for Random Acquisition

Figure S3 shows the progression of the GP mean and contour predictions when simulation points are selected purely at random. The $(\phi, \kappa a)$ slice exhibits reasonably satisfactory reconstruction: as more random samples are drawn, the GP gradually recovers the general shape of the phase boundary and approaches the ground-truth contour. However, this behavior does not generalize across the other two directions. In both the $(V^*, \kappa a)$slice and the (V^*, ϕ) slice, random acquisition fails to place sufficient samples within the narrow, steep transition region. As a result, the GP routinely misidentifies the location of the boundary, with large deviations that persist even after tens of additional simulations. In the case of $(V^*, \kappa a)$ slice, the method completely misses the top left boundary, and in the (V^*, ϕ) it's able to capture the horizontal boundary bot only after 31 simulations, as compared to 19 of the gradient-based method.

These observations highlight a key limitation of random acquisition: without gradient awareness, sampling density does not naturally concentrate near the physically meaningful transition manifold. In contrast, our gradient-based active learning strategy consistently targets high-sensitivity regions, ensuring reliable and much faster boundary identification across all slices.

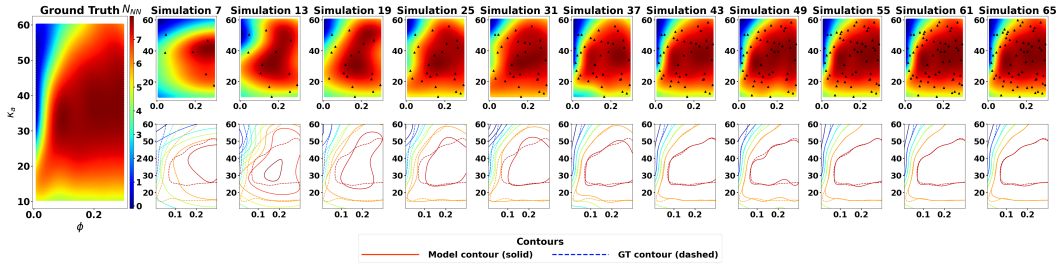
4 Comparison of Target-based acquisition baseline to our gradient-based acquisition.

We additionally evaluate a benchmark acquisition strategy that directly leverages domain knowledge of the phase boundary. Because the dilute–condensed transition is thought to occur at the midway point between dilute and condensed phases, i.e., $N_{nn} = 6$, this baseline preferentially samples locations whose GP predictions lie near this value. To achieve this, we design an uncertainty-weighted objective

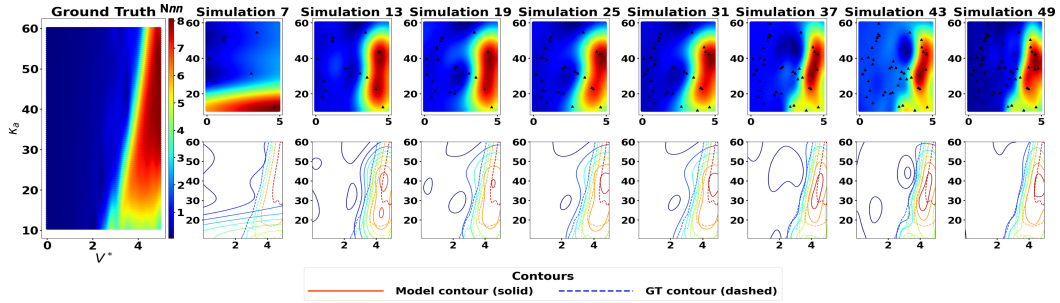
$$\alpha(\mathbf{x}) = -|\mu(\mathbf{x}) - 6| + \beta * \sigma(\mathbf{x}), \tag{S3}$$

where β is a multiplying factor that allows for exploration. The value of $\beta = 1$ is used for comparison with other acquisition methods. This favors points close to the known transition while still promoting exploratory sampling through the predictive

Random Baseline: (ϕ, κ) **Slice,** $V^* = 4$



Random Baseline: (V^*, κ) **Slice,** $\phi = 0.05$



Random Baseline: (V^*, ϕ) **Slice,** $\kappa a = 50$

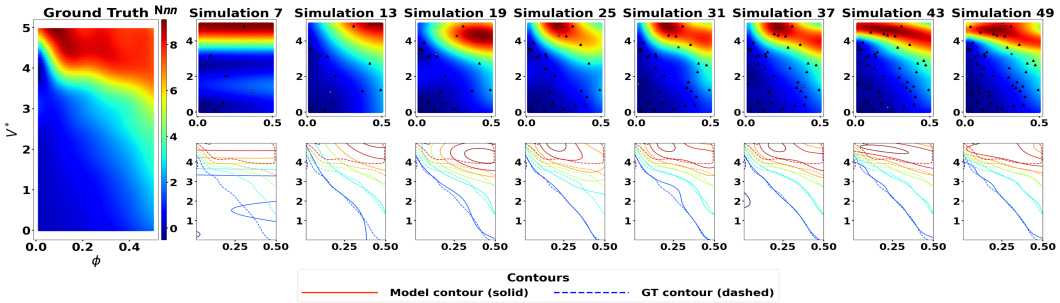


Figure S3: Random-sampling baseline for the three two-dimensional slices. For each slice, the leftmost panel shows the ground-truth phase diagram. Subsequent columns show GP means (top rows) and their contour comparisons with the ground-truth $N_{nn} = 6$ contour (bottom rows) as more random points are sampled. Unlike the gradient-aware active learning strategy, random sampling produces slow and spatially inconsistent refinement of the phase boundary, requiring substantially more simulations to approximate the correct contour.

uncertainty $\sigma(\mathbf{x})$. Since this method is informed by the ground-truth target characterizing the phase boundary, it provides an upper benchmark for what a boundary-aware sampling strategy can achieve.

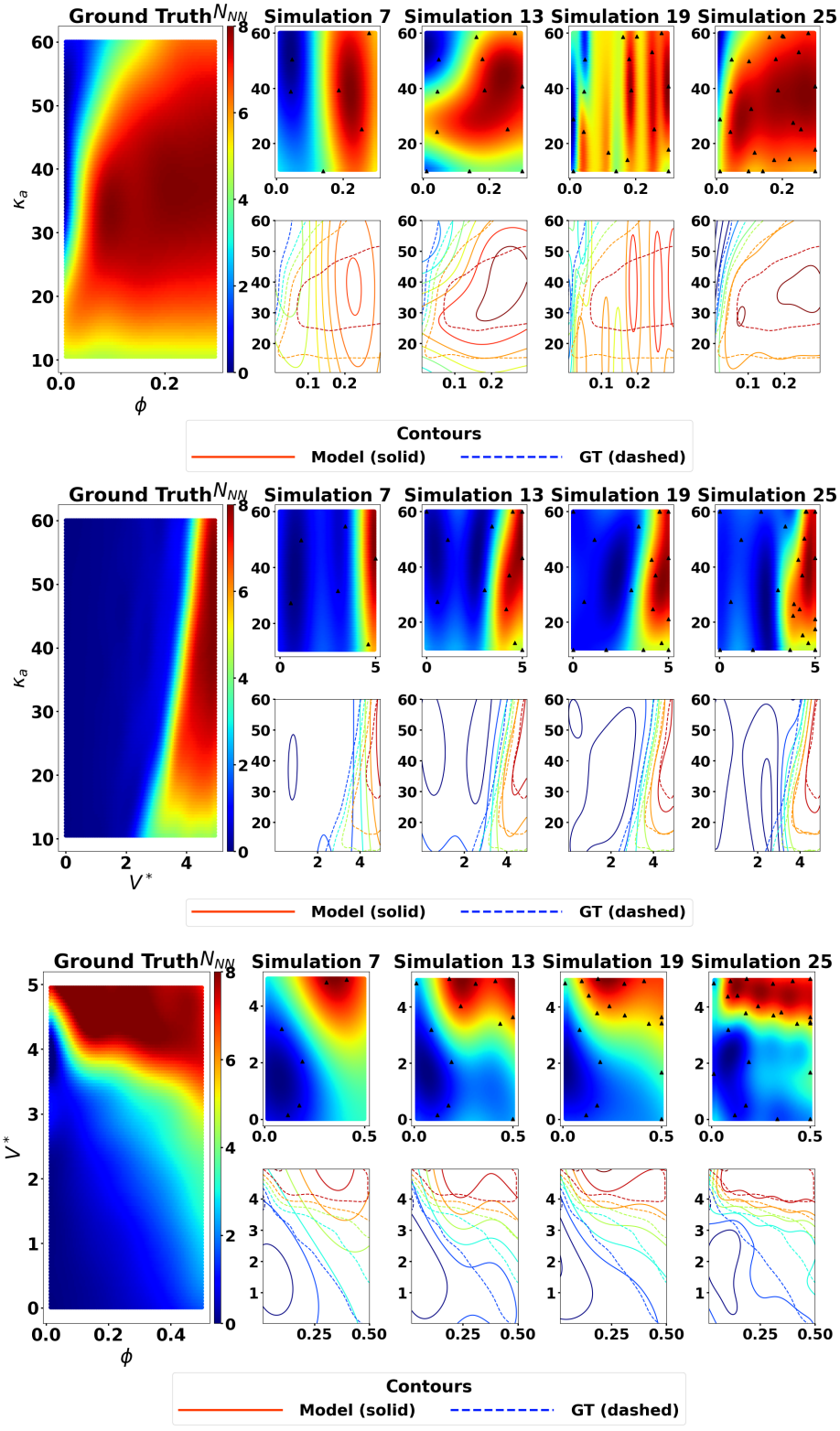


Figure S4: Progression of GP mean and contour estimates under target-based acquisition for fixed $V^* = 4$ (top panel), fixed $\phi = 0.05$ (middle panel), and $\kappa a = 50$ (bottom panel).

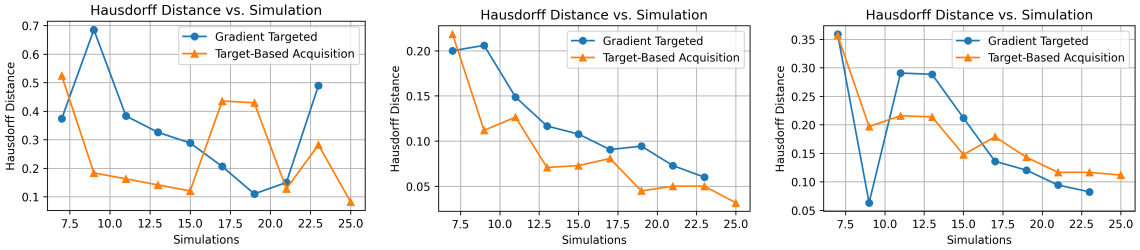


Figure S5: **Hausdorff-distance convergence for gradient-based and target-based acquisition.** Comparison of boundary reconstruction error as a function of the number of simulations for three two-dimensional slices: (left) $(\phi, \kappa a)$ at fixed $V^* = 4$, and (middle) $(V^*, \kappa a)$ at fixed $\phi = 0.05$, and (right) (V^*, ϕ) at fixed $\kappa a = 50$,. The target-based acquisition (orange) exploits prior knowledge that the phase boundary occurs at $N_{nn} = 6$, while the gradient-based acquisition (blue) relies solely on local sensitivity information. Despite lacking access to the true transition value, the gradient-based method approaches the performance of the target-informed baseline across all slices.

Figure S4 illustrates the evolution of the Gaussian Process (GP) mean and the corresponding phase-boundary contours under the target-based acquisition strategy. Across all three two-dimensional slices, the GP rapidly concentrates samples along the known transition manifold, leading to a steady sharpening of the predicted phase boundary as additional simulations are acquired. As shown in Fig. S5, our gradient-targeted method achieves Hausdorff-distance convergence that is *remarkably close* to this target-informed baseline, despite having no prior knowledge of the transition value. Both methods rapidly reduce the Hausdorff distance within the first ∼15–20 simulations, with the target-based acquisition maintaining a slight but expected advantage due to its access to the correct phase-boundary target. The small gap between the two curves highlights that the proposed gradient-based acquisition can autonomously identify and prioritize high-sensitivity regions of the phase diagram, thereby achieving performance close to a strategy explicitly guided by domain knowledge. This demonstrates the practical strength of our approach, especially in settings where such target values are unknown or experimentally expensive to determine.

References

[1] S. Plimpton. "Fast parallel algorithms for short-range molecular dynamics". In: *J. Comput. Phys.* 117.1 (1995), pp. 1–19.

[2] Rochester Institute of Technology. *Research Computing Services*. 2025. DOI: 10.34788/0S3G-QD15. URL: https://www.rit.edu/researchcomputing/.

[3] A. P. Thompson et al. “LAMMPS - a flexible simulation tool for particle-based materials modeling at the atomic, meso, and continuum scales”. In: *Comp. Phys. Comm.* 271 (2022), p. 108171. DOI: 10.1016/j.cpc.2021.108171.