Multilevel simulation of hard-sphere mixtures

Paul Rohrbach

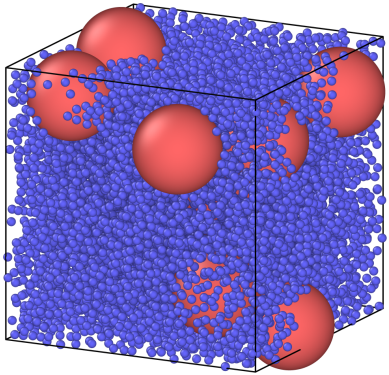
joint work with Hideki Kobayashi, Robert Scheichl, Nigel B. Wilding, and Robert L. Jack

Department of Applied Mathematics and Theoretical Physics
University of Cambridge

3 May 2022

Motivation

- **Example: Size asymmetric hard sphere mixture**
  - Model for colloid-polymer mixtures[1]
- Investigating long length-scales poses challenge for molecular simulation
- Simulations are made possible by coarse-graining[2]
- Multilevel idea[3]: use model hierarchy to speed up **exact simulation**
  - coarse & fast $\longleftrightarrow$ accurate & slow



[1] Roth, R., Evans, R., & Dietrich, S. (2000). Depletion potential in hard-sphere mixtures: Theory and applications. Phys. Rev. E, 62, 5360.

[2] Noid, W. G., et al (2008). The multiscale coarse-graining method. I. A rigorous bridge between atomistic and coarse-grained models. J. Chem. Phys., 128, 244114.

[3] Giles, M. B. (2008). Multilevel Monte Carlo path simulation. Oper. Res., 56, 607-617.

Outline

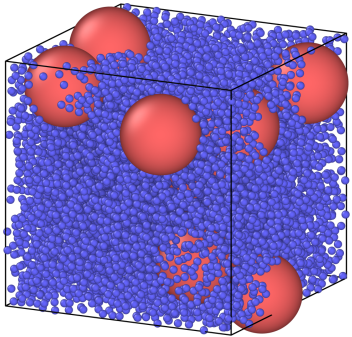
Size-asymmetric hard-sphere mixture

We consider a grand-canonical ensemble of

- big particles $\mathcal{C} = (N; \mathbf{R}_1, \ldots, \mathbf{R}_N)$,
- small particles $\mathcal{F} = (n; \mathbf{r}_1, \ldots, \mathbf{r}_N)$,

with distribution

$$p_{\mathsf{F}}(\mathcal{C}, \mathcal{F}) = \frac{1}{\Xi_{\mathsf{F}}} e^{\mu_B N + \mu_S n - U_{\mathsf{F}}(\mathcal{C}, \mathcal{F})} .$$



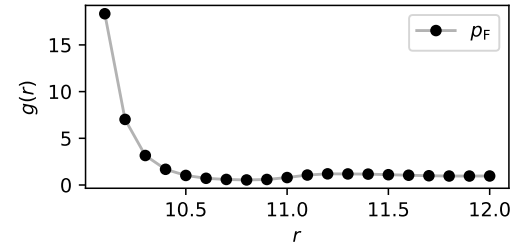
- **Goal:** Compute ensemble average of $A = A(\mathcal{C})$

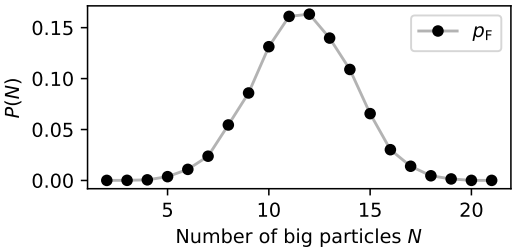
$$\begin{aligned} \langle A \rangle_{\mathsf{F}} &= \int p_{\mathsf{F}}(\mathcal{C}, \mathcal{F}) A(\mathcal{C}) \, \mathrm{d}\mathcal{C} \, \mathrm{d}\mathcal{F} \\ &= \frac{1}{\Xi_{\mathsf{F}}} \sum_{N=0}^{\infty} \sum_{n=0}^{\infty} \frac{1}{n! N! \sigma_B^{3N} \sigma_S^{3n}} \int p_{\mathsf{F}}(\mathcal{C}, \mathcal{F}) A(\mathcal{C}) \, \mathrm{d}\mathbf{R}_1 \cdots \mathrm{d}\mathbf{R}_N \, \mathrm{d}\mathbf{r}_1 \cdots \mathbf{r}_n . \end{aligned}$$

Model problem

- **Example parameters**
  - particle diameter $\sigma_B = 10, \quad \sigma_S = 1$
  - small particle volume fraction $\eta_S = 0.2$
  - system size $L = 31$

- **Quantities of interest**
  - Big particle pair correlation function $g(r)$
  - Distribution of the number of big particles $P(N)$





Coarse-grained model

- Integrating out small particles gives intractable marginal

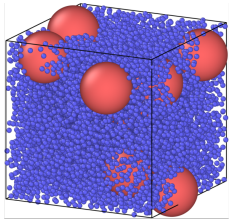
$$p_{\mathsf{M}}(\mathcal{C}) = \int p_{\mathsf{F}}(\mathcal{C}, \mathcal{F})\, \mathrm{d}\mathcal{F} = \frac{1}{\Xi_{\mathsf{M}}} e^{\mu_B N + U_{\mathsf{M}}(\mathcal{C})}.$$

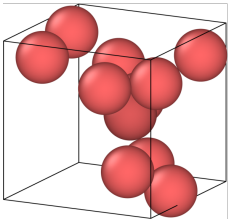
- **Coarse-grained approximation**

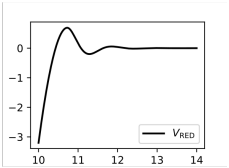
$$p_{\mathsf{C}}(\mathcal{C}) = \frac{1}{\Xi_{\mathsf{C}}} e^{\mu_B N - U_{\mathsf{C}}(\mathcal{C})}$$

where

$$U_{\mathsf{C}}(\mathcal{C}) = N\Delta\mu + \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} V_2(|\,\mathbf{R}_i - \mathbf{R}_j\,|).$$

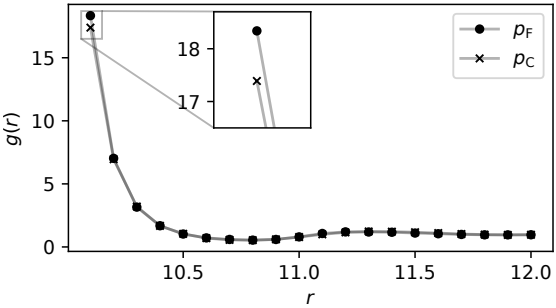


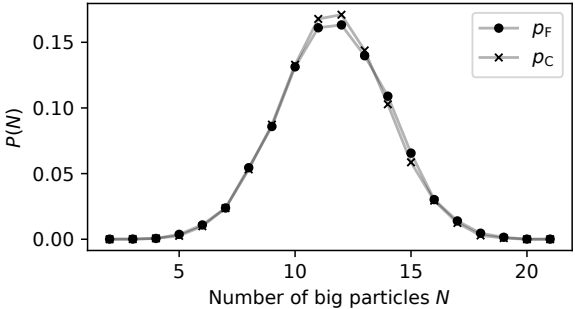




Coarse-grained model – Accuracy

- Coarse-grained model: RED potential[4]
  → accurate, but not exact





[4] Roth, R., Evans, R., & Dietrich, S. (2000). Depletion potential in hard-sphere mixtures: Theory and applications. Phys. Rev. E, 62, 5360.

Connecting coarse and fine

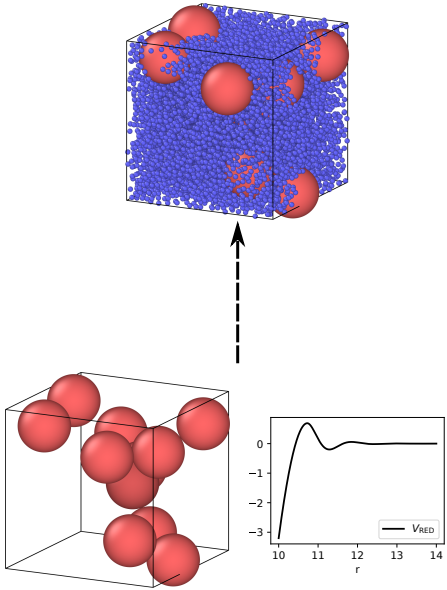
- Importance sampling:

$$\langle A \rangle_{\mathsf{F}} = \langle w(\mathcal{C}) A(\mathcal{C}) \rangle_{\mathsf{C}}$$

with importance weight

$$w(\mathcal{C}) = \frac{p_{\mathsf{M}}(\mathcal{C})}{p_{\mathsf{C}}(\mathcal{C})}.$$

- Importance sampling still works if
  - $w(\mathcal{C}) \propto p_{\mathsf{M}}(\mathcal{C})/p_{\mathsf{C}}(\mathcal{C})$,
  - $w(\mathcal{C})$ is an **unbiased estimator** of the unnormalised weight.



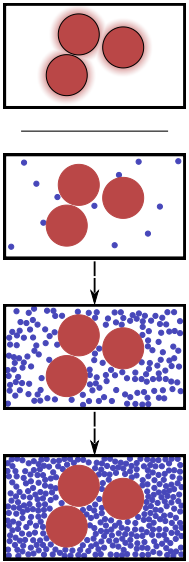
Computing importance weights

- The marginal density is a partition function

$$p_{\mathsf{M}}(\mathcal{C}) \propto \int e^{\mu_B N + \mu_S n - U_{\mathsf{F}}(\mathcal{C},\mathcal{F})} \, \mathrm{d}\mathcal{F}$$
$$= e^{\mu_B N} \Xi[\mathcal{C}, \mu_S].$$

- For small $\mu_0 \ll \mu_S$, evaluate $\Xi[\mathcal{C}, \mu_0]$ directly
- Unbiased estimate of $\Xi[\mathcal{C}, \mu_S]/\Xi[\mathcal{C}, \mu_0]$ via Annealed Importance Sampling[5]

$$\Rightarrow \quad \hat{w}(\mathcal{C}) = \frac{\hat{\Xi}[\mathcal{C}, \mu_S]}{e^{-U_{\mathsf{C}}(\mathcal{C})}} \quad \text{such that} \quad \langle \hat{w}(\mathcal{C}) \rangle \propto p_{\mathsf{M}}(\mathcal{C}) / p_{\mathsf{C}}(\mathcal{C})$$



[5] Neal, R.M. (2001). Annealed importance sampling. Stat. Comput. , 11(2), 125-139.

Two-level Monte Carlo

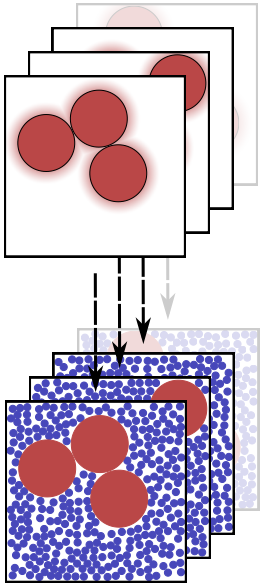
For coarse samples $\mathcal{C}_1, \ldots, \mathcal{C}_M \sim p_{\mathsf{C}}$, we have

- $\hat{A}_{\mathsf{C}} = \frac{1}{M} \sum_{j=1}^{M} A(\mathcal{C}_j) \approx \langle A \rangle_{\mathsf{C}}$,
- $\hat{A}_{\mathsf{F}} = \frac{1}{M} \sum_{j=1}^{M} \hat{w}(\mathcal{C}_j) A(\mathcal{C}_j) \approx \langle A \rangle_{\mathsf{F}}$,

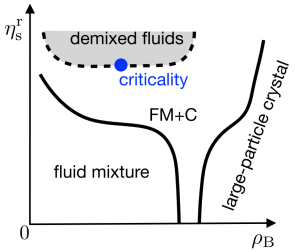
where the weights are normalised to $\frac{1}{M} \sum_{j=1}^{M} \hat{w}(\mathcal{C}_j) = 1$.

We can reduce the variance by computing a **difference estimate**

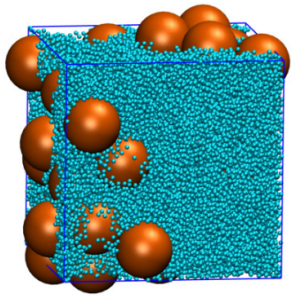
- $\hat{\Delta} = \frac{1}{M} \sum_{j=1}^{M} \left(\hat{w}(\mathcal{C}_j) - 1\right) A(\mathcal{C}_j) \approx \langle A \rangle_{\mathsf{F}} - \langle A \rangle_{\mathsf{C}}$,
- $\hat{A}_{\mathsf{F},\Delta} = \hat{A}_{\mathsf{C}} + \hat{\Delta}$.



Applications and limitations



- Critical point of demixing[6]
  - system size $L = 4\sigma_B$
  - $q = 11/1, \quad \eta_S \approx 0.301$
- Computation possible but substantial effort required
  - Weight computations fully parallel
- Performance determined by
  - Accuracy of coarse level $D_{\chi^2}(p_{\mathsf{M}} \,\|\, p_{\mathsf{C}})$
  - Variance of weights $\hat{w}(\mathcal{C})$

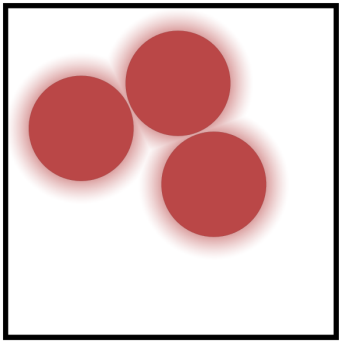


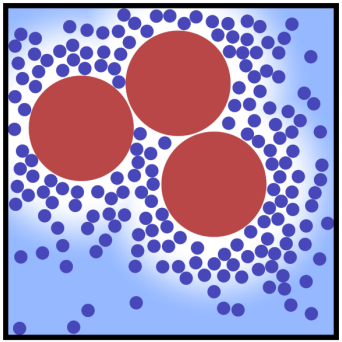
[6] Kobayashi, H., PBR, Scheichl, R., Wilding, N. B., & Jack, R. L. (2021). Critical point for demixing of binary hard spheres. Phys. Rev. E, 104(4), 044603.

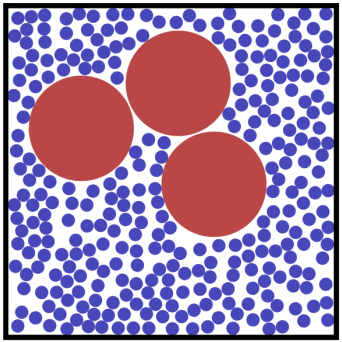
Section 2

Three- and higher level methods

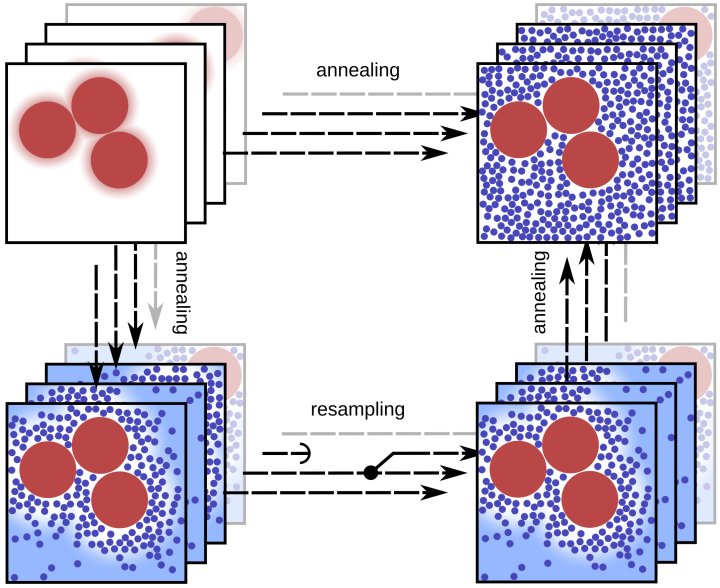
Three-level method





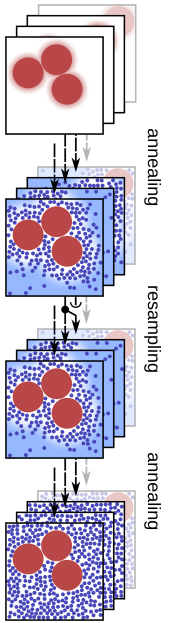


Three-level method II



Algorithm

- Random-weight sequential Monte Carlo
  - Repeat importance weighting – resampling steps for further levels
- Estimators
  - $\hat{A}^{\mathsf{3L}}_{\mathsf{F}} = \frac{1}{M}\sum_{j=1}^{M} \hat{w}_{\mathsf{F}}(\mathcal{C}^2_j, \mathcal{F}^2_j) A(\mathcal{C}^2_j) \approx \langle A \rangle_{\mathsf{F}}$
  - $\hat{A}^{\mathsf{3L}}_{\mathsf{F},\Delta} = \hat{A}_{\mathsf{C}} + \hat{\Delta}^{\mathsf{3L}}_{\mathsf{I}} + \hat{\Delta}^{\mathsf{3L}}_{\mathsf{F}} \approx \langle A \rangle_{\mathsf{F}}$.
  - CLT for final-level estimators as $M \to \infty$ by adapting SMC results[7]
- Potential benefits of resampling
  - Focus on relevant configurations
  - Tapering of population
  - Improved control variates for difference estimate



[7] Cappé, O., et al (2005). Inference in Hidden Markov Models. Springer-Verlag.

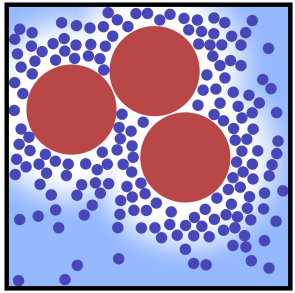
Intermediate level

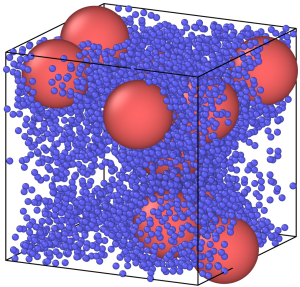
For fixed coarse configuation $\mathcal{C}$:

1. Suppress small particle insertion

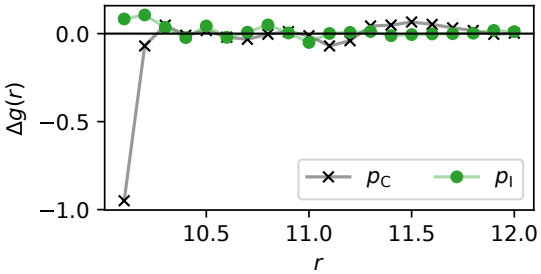
$$E(\mathbf{r}) = E(\mathrm{dist}(\mathbf{r}, \mathcal{C}))$$

2. Estimate restricted small particle partition function
3. Approximate missing factor
   → square-gradient approximation





Accuracy:



Numerical example

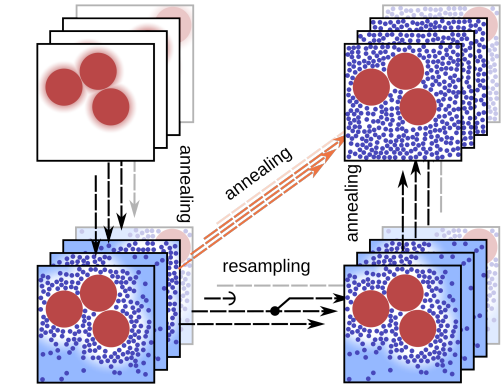
Compare two- and three level method

- Intermediate level is halfway point
- Same annealing schedule
- What is the effect of resampling?

- Quantity of interest: coordination number

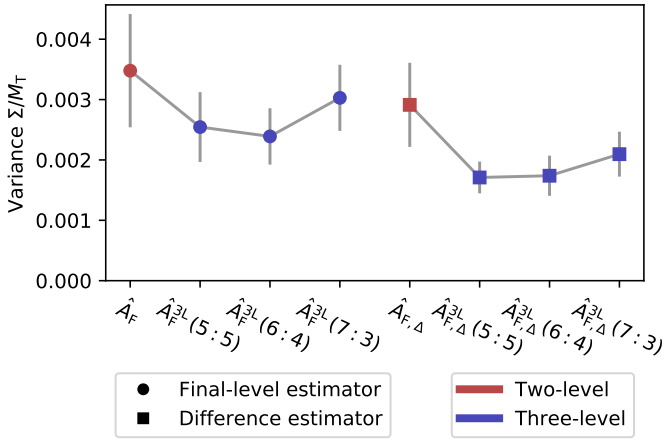
$$N_c = 4\pi \int_{10}^{10.73} r^2 g(r) \rho \, \mathrm{d}r$$

$$\langle N_c \rangle_{\mathsf{F}} \approx 1.61, \quad \langle N_c \rangle_{\mathsf{C}} \approx 1.56.$$



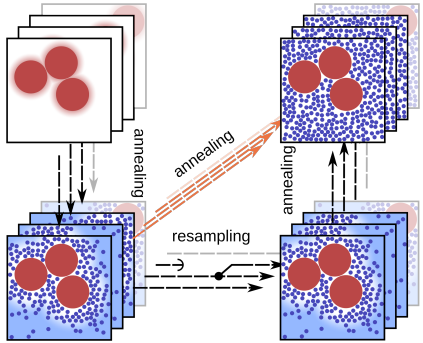
Numerical example II

- Population size $M = 500$
- 60 independent evaluations
- Tapering rates:
  - $M_\mathsf{I} = 600, \quad M_\mathsf{F} = 400$
  - $M_\mathsf{I} = 700, \quad M_\mathsf{F} = 300$



What is the effect of resampling?

- Final-level estimate only
- Simplifying assumptions
  - $\mathcal{C}_1, \ldots, \mathcal{C}_M \sim p_{\mathsf{C}}$ i.i.d.
  - $\langle A \rangle_{\mathsf{F}} = 0$
  - Equivalent annealing schedule



Two-level method

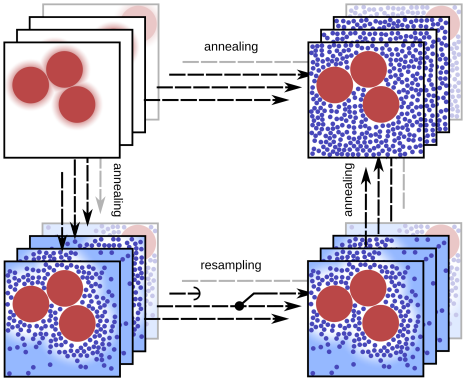
- $\sqrt{M}\hat{A}_{\mathsf{F}} \xrightarrow{\mathcal{D}} \mathcal{N}(\langle A \rangle_{\mathsf{F}}, \Sigma_{\mathsf{F}})$
- $\Sigma_{\mathsf{F}} = \mathbb{E}_{p_{\mathsf{C}}}\left[\left(\hat{w}_{\mathsf{I}}(\mathcal{C}^1)\hat{w}_{\mathsf{F}}(\mathcal{C}^1, \mathcal{F}^2)A(\mathcal{C}^1)\right)^2\right]$

Three-level method

- $\sqrt{M}\hat{A}_{\mathsf{F}}^{\mathsf{3L}} \xrightarrow{\mathcal{D}} \mathcal{N}(\langle A \rangle_{\mathsf{F}}, \Sigma_{\mathsf{F}}^{\mathsf{3L}})$
- $\Sigma_{\mathsf{F}}^{\mathsf{3L}} = \mathbb{E}_{p_{\mathsf{C}}}\left[\left(\hat{w}_{\mathsf{I}}(\mathcal{C}^1)\,\mathbb{E}\left[\hat{w}_{\mathsf{F}}(\mathcal{C}^1, \mathcal{F}^2)\right]A(\mathcal{C}^1)\right)^2\right] + \mathbb{E}_{p_{\mathsf{I}}}\left[\left(\hat{w}_{\mathsf{F}}(\mathcal{C}^2, \mathcal{F}^2)A(\mathcal{C}^2)\right)^2\right]$
- Tapering can balance two variance contributions

Summary

- Three-level method makes efficient use of intermediate information
- Application to hard-sphere mixture limited by intermediate level
  - Fitting $2 + 3$-body potential expensive but more accurate
- Further applications to models with natural mutli-scale description
  - Multiblob polymer models



Questions?

Bibliography:

- Kobayashi, H., Rohrbach, P.B., Scheichl, R., Wilding, N.B., & Jack, R.L. (2019). Correction of coarse-graining errors by a two-level method: Application to the Asakura-Oosawa model. The Journal of Chemical Physics, 151, 144108.
- Kobayashi, H., Rohrbach, P.B., Scheichl, R., Wilding, N.B., & Jack, R.L. (2021). Critical point for demixing of binary hard spheres. Physical Review E, 104(4), 044603.
- Rohrbach, P.B., Kobayashi, H., Scheichl, R., Wilding, N.B., & Jack, R.L. (2022). Multilevel simulation of hard-sphere mixtures. arXiv:22xx.xxxxx (soon!).

LEVERHULME TRUST

This work was supported by the Leverhulme Trust through research project grant number RPG–2017–203.