Machine Learning and inverse design of Soft Materials

5th Edwards Symposium - Future / New Directions in Soft Matter 8th of September 2021

Universiteit Utrecht

Marjolein Dijkstra

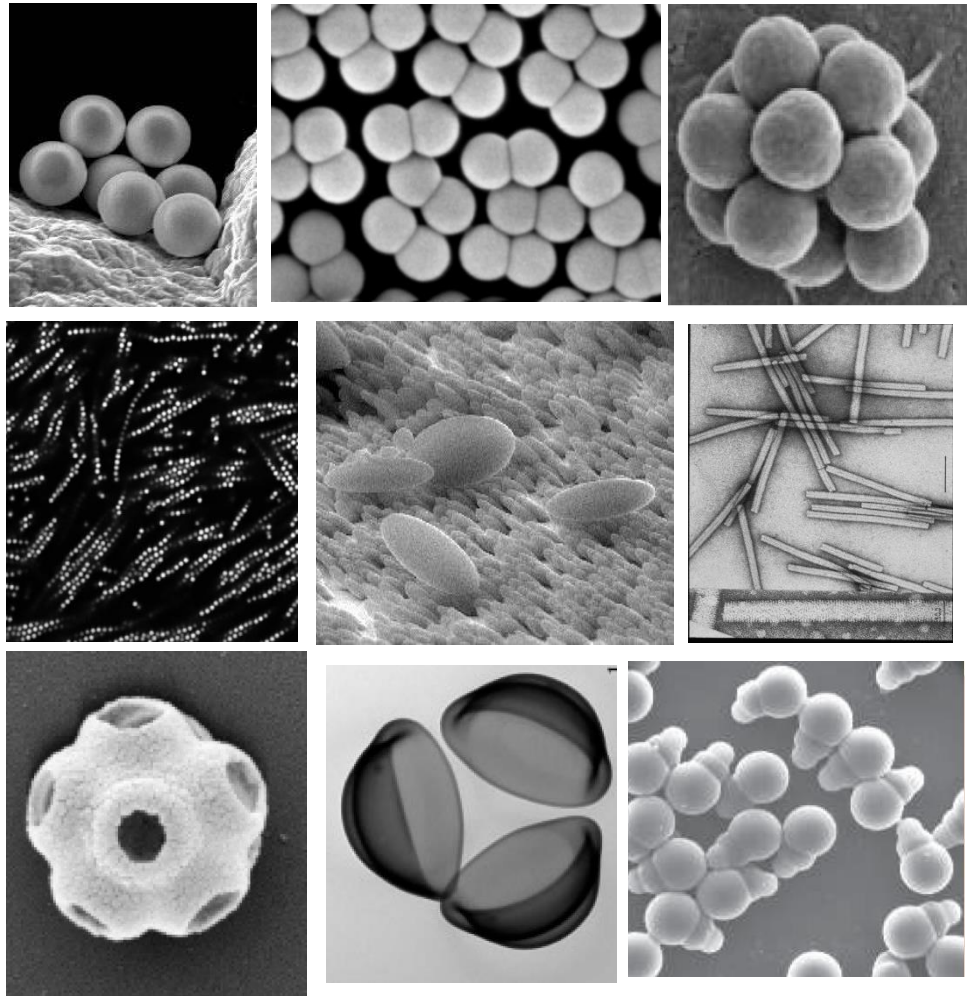
https://colloid.nl/people/marjolein-dijkstra/

Soft condensed matter group
Debye Institute for Nanomaterials Science
Utrecht University

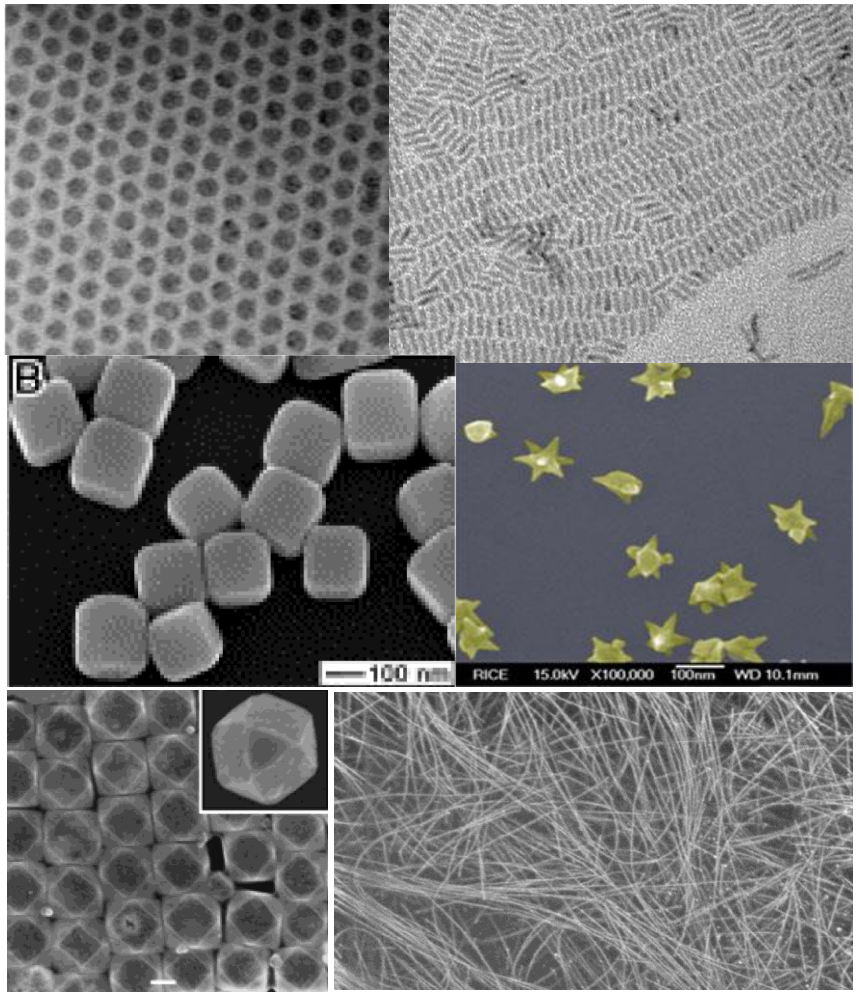
Colloids

Shape, size, physical and material properties, interactions (hard, depletion attractions, charge, dipolar, patchy) can be tuned

Size > 100 nm

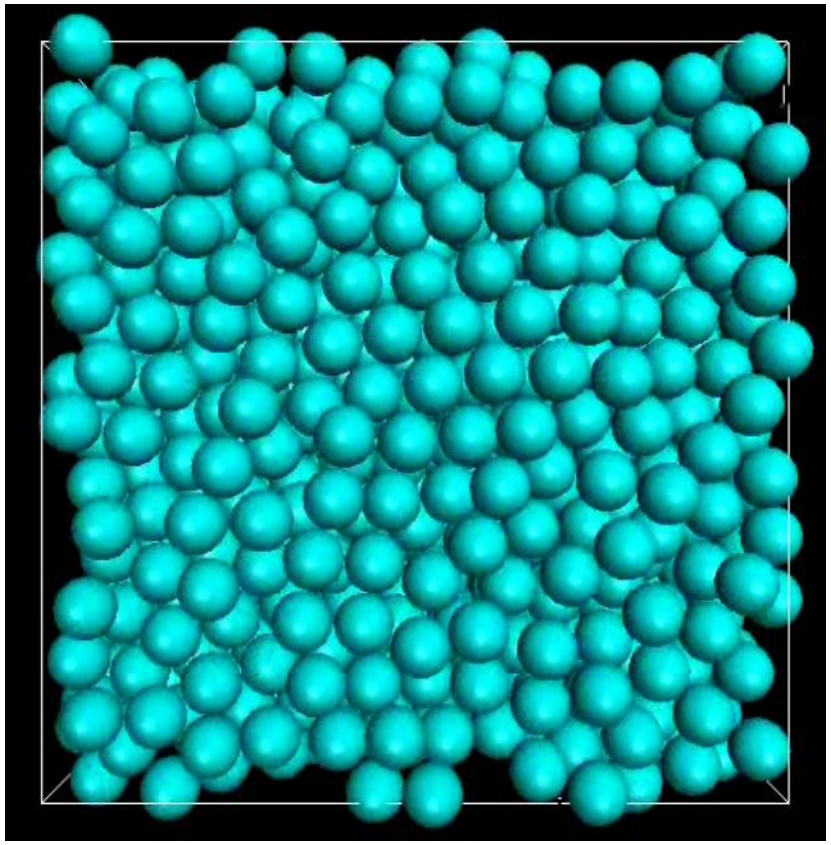


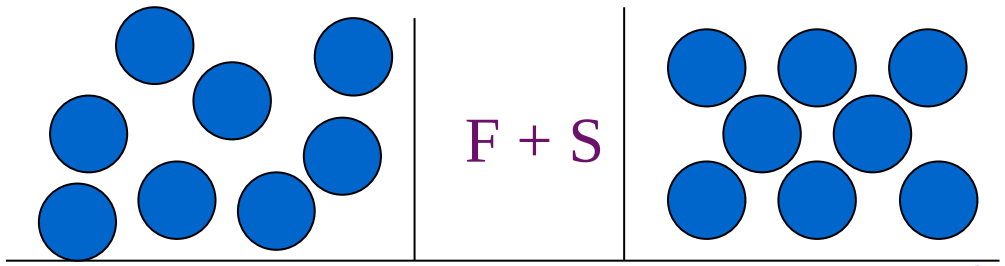
nanoparticles



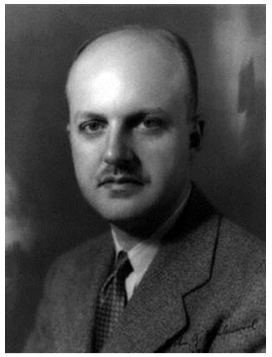
Computer Simulations of hard spheres

In 1957, computer simulations showed spontaneous crystallization of hard spheres at η ~0.5.

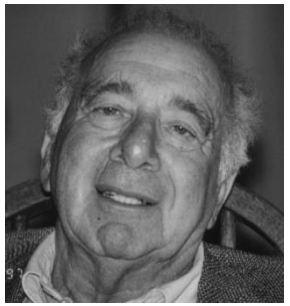




0.494 0.545 $\eta = \frac{\pi\sigma^3}{6}\frac{N}{V}$



John G. Kirkwood



B.J. Alder

Alder and Wainwright, J. Chem. Phys. 27, 1208 (1957)
Wood and Jacobson, J. Chem. Phys. 27, 1207 (1957).

How does the crystal form?

Nucleation

Stage 1

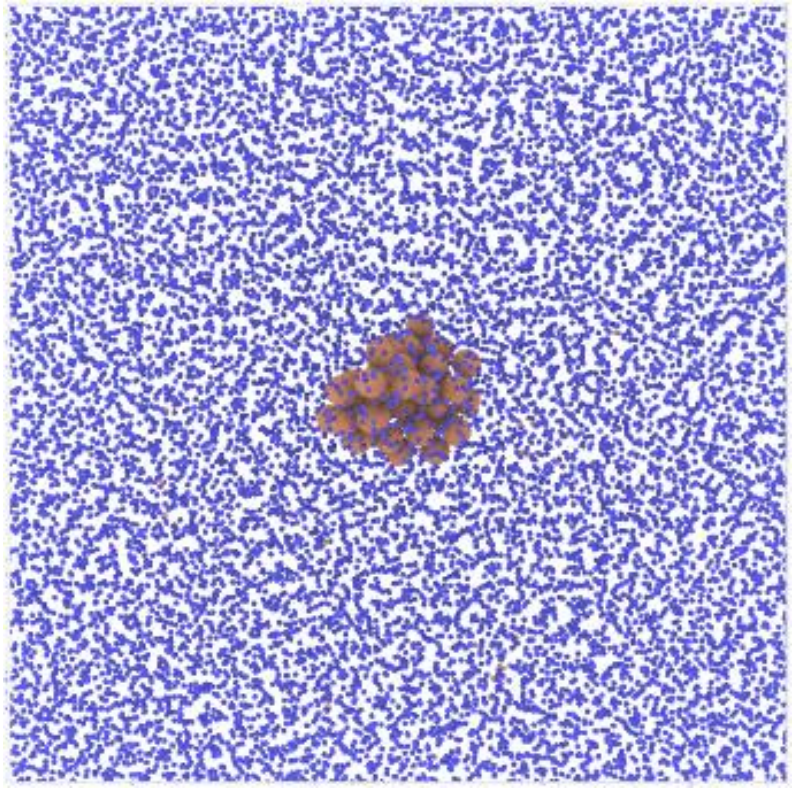
Due to thermal fluctuations **small nuclei of the stable phase** appear and disappear in the parent phase.

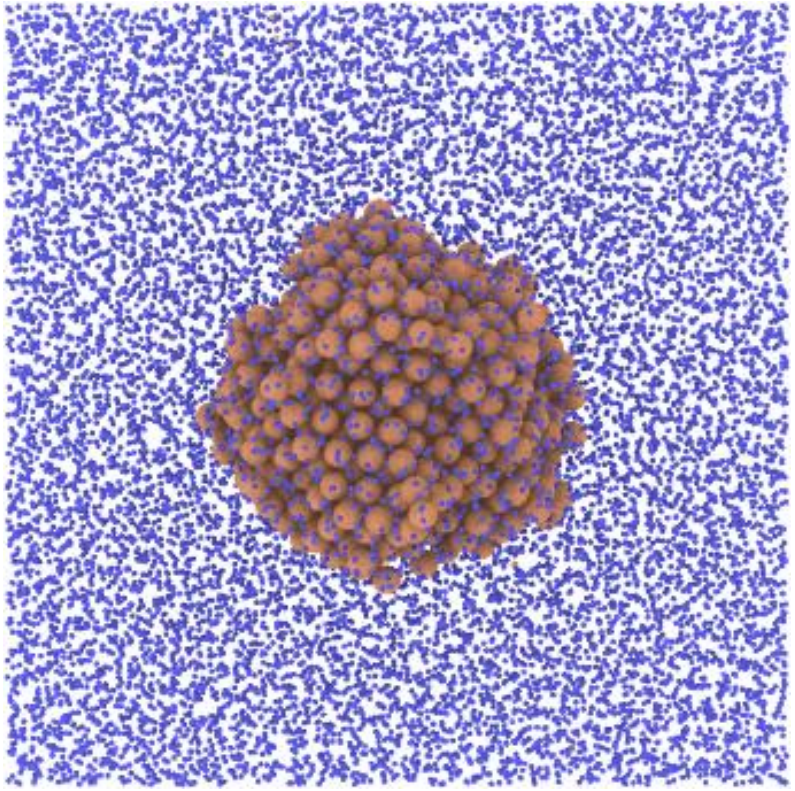
Stage 2

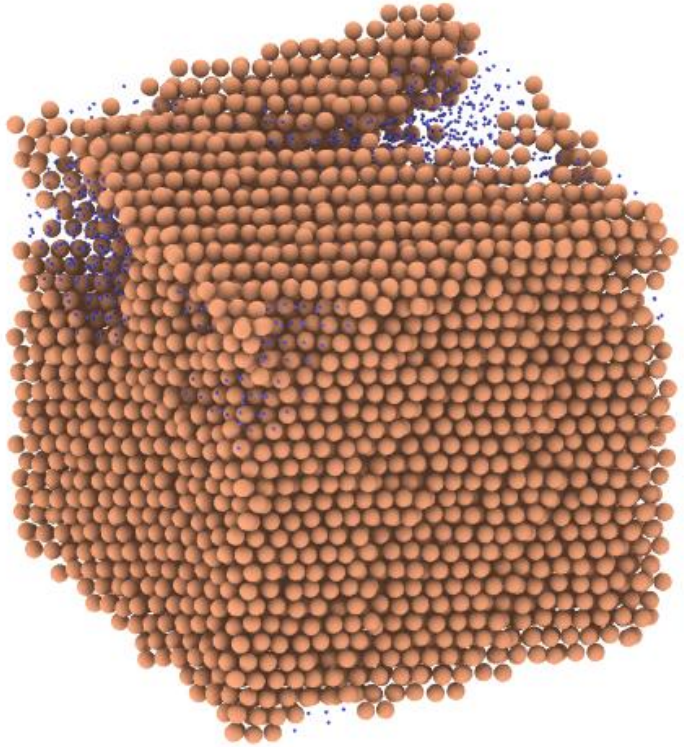
If an embryo has a size larger than the **critical size**, it will keep on growing…

Stage 3

… until it spans the **whole system**!

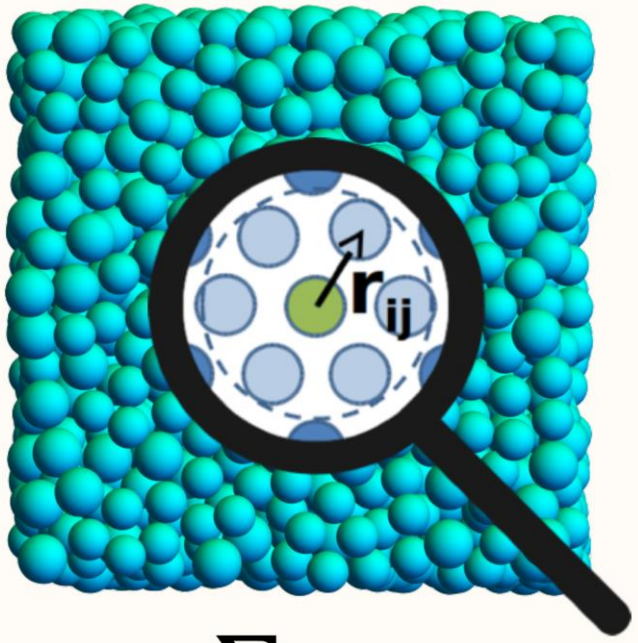






Fluid particles shown in blue and ten times smaller

Bond Order parameters



$$\rho_i = \sum_{j \in N_b(i)} \delta(\mathbf{r} - \mathbf{r}_{ij})$$

Local environment**: all particles within a certain distance are called **neighbours

$$q_{lm}(i) = \frac{1}{N_b(i)} \sum_{j=1}^{N_b(i)} Y_l^m(\mathbf{r}_{ij})$$

$$\bar{q}_{lm}(i) = \frac{1}{1 + N_b(i)} \left[q_{lm}(i) + \sum_{j=1}^{N_b(i)} q_{lm}(j) \right]$$

$$\bar{q}_l(i) = \sqrt{\frac{4\pi}{2l+1} \sum_{m=-l}^{l} |\bar{q}_{lm}(i)|^2}$$

Steinhardt, Nelson, Ronchetti, PRB 28, 784 (1983)
Lechner, Dellago, JCP 129, 114707 (2008)

q_6 (and q_4) often used for crystallization

How does the crystal nucleate?

You may conclude that nucleation proceeds via BCC

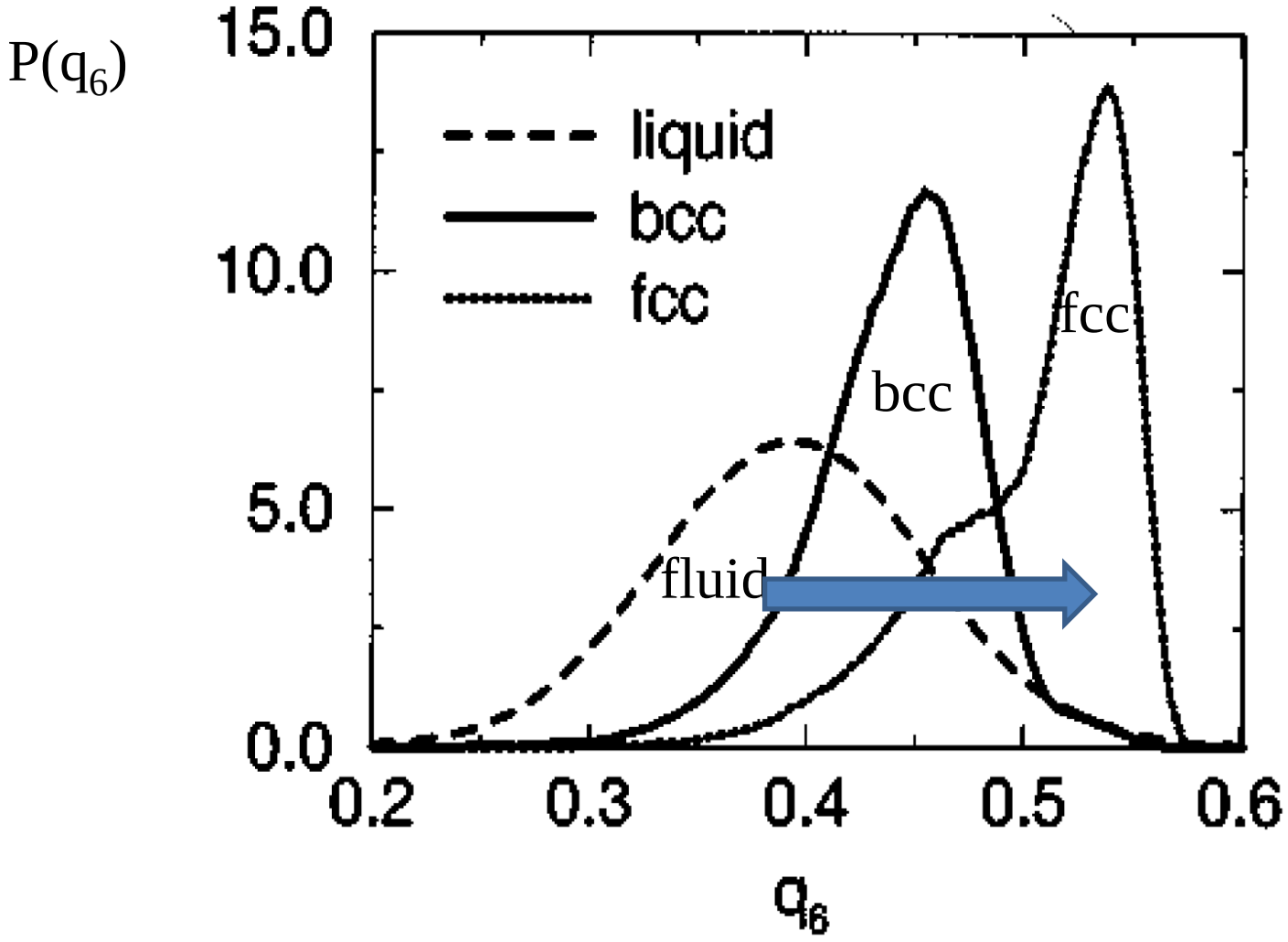
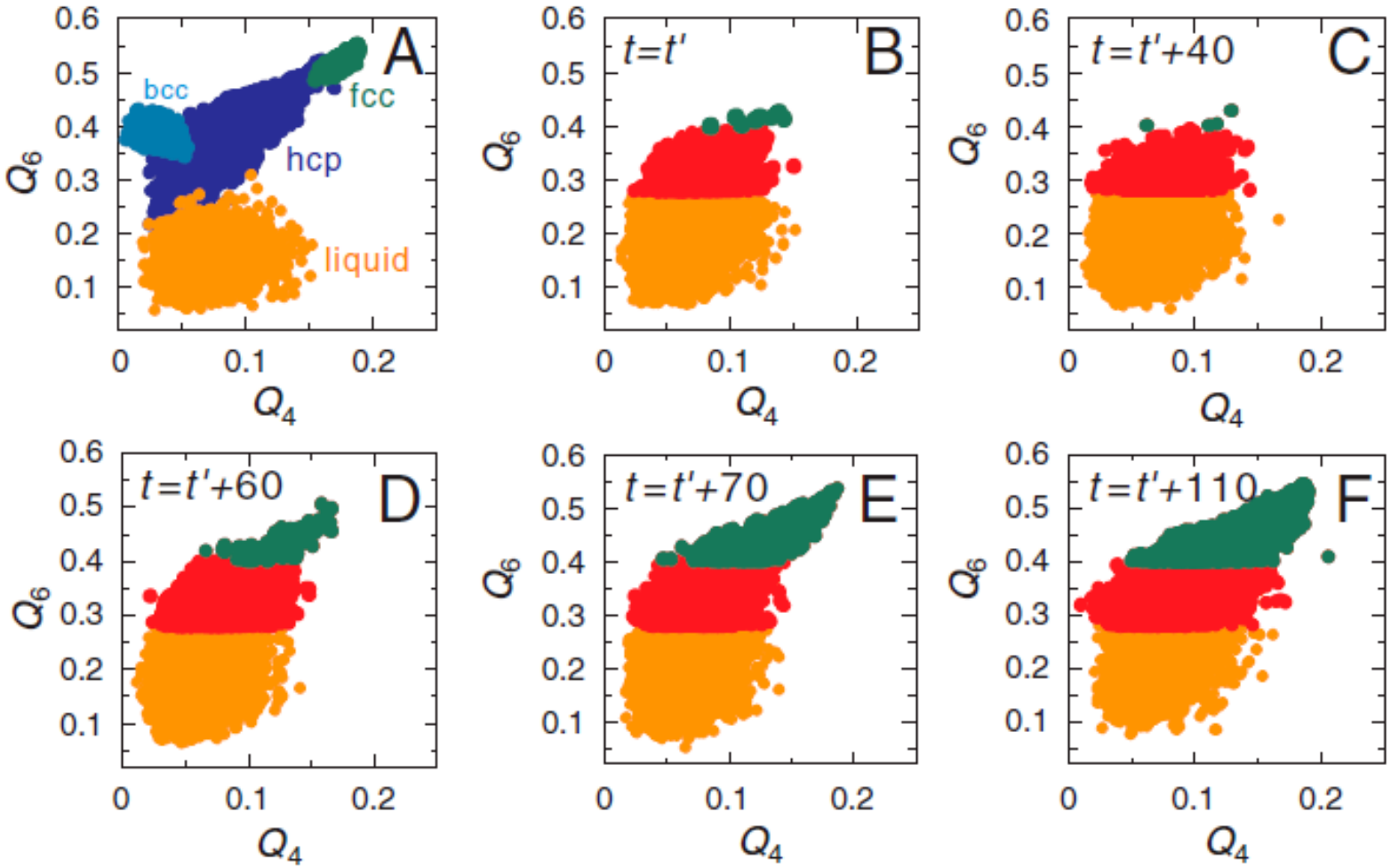


Figure from Ten Wolde, Ruiz-Montero, Frenkel, PRL 75 (1997)

How does the crystal nucleate?

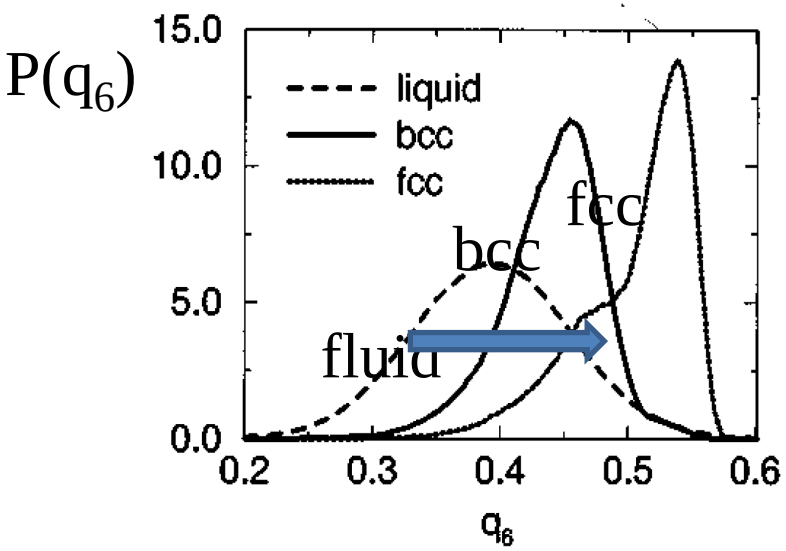
Kawasaki and Tanaka, PNAS 107 (2010) conclude that

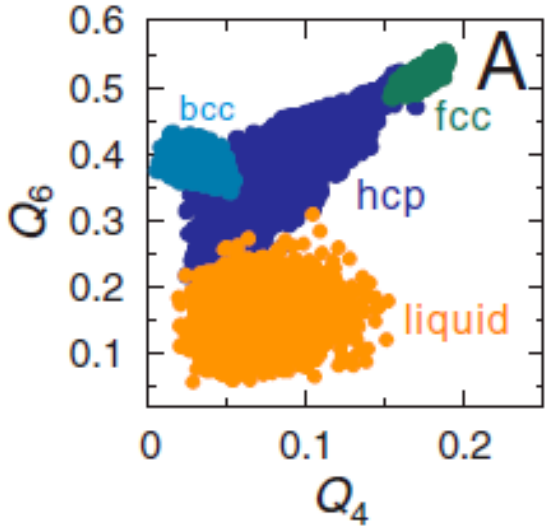
Fluid shows medium-range hcp-like order from which nucleation of hard spheres occurs

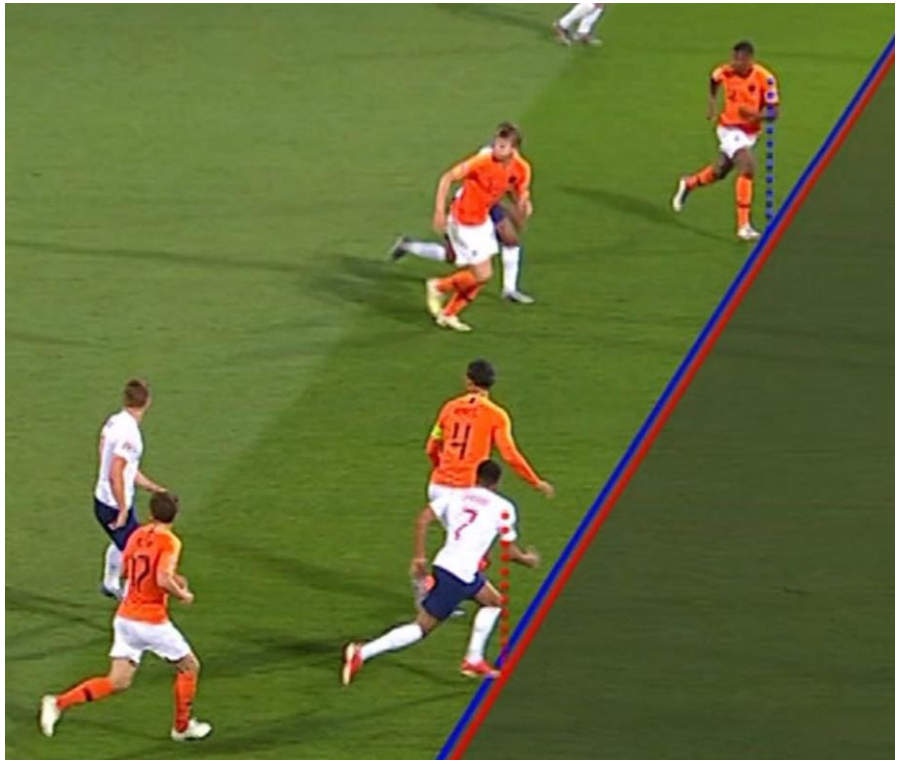


Nucleation proceeds via BCC/HCP???

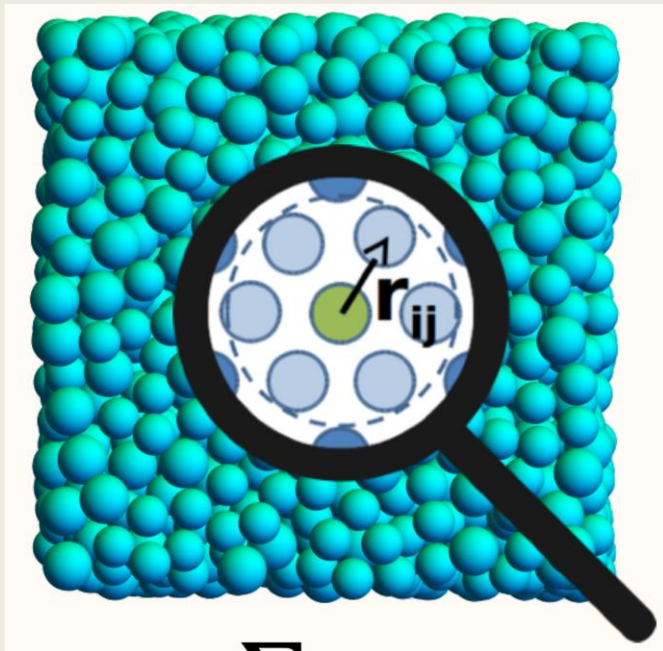
Offside???







Bond Order parameters



$$\rho_i = \sum_{j \in N_b(i)} \delta(\mathbf{r} - \mathbf{r}_{ij})$$

Local environment: *all particles within a certain distance are called neighbours*

$$q_{lm}(i) = \frac{1}{N_b(i)} \sum_{j=1}^{N_b(i)} Y_l^m(\mathbf{r}_{ij})$$

$$\bar{q}_{lm}(i) = \frac{1}{1 + N_b(i)} \left[q_{lm}(i) + \sum_{j=1}^{N_b(i)} q_{lm}(j) \right]$$

$$\bar{q}_l(i) = \sqrt{\frac{4\pi}{2l+1} \sum_{m=-l}^{l} |\bar{q}_{lm}(i)|^2}$$

Steinhardt, Nelson, Ronchetti, PRB 28, 784 (1983)

Description of the local environment

$$\{\mathbf{Q}(i) = (\bar{q}_1(i), \bar{q}_2(i), \bar{q}_3(i), \ldots, \bar{q}_d(i))\}$$

How to find reliable order parameters?

Generate **configurations** $\{\mathbf{r}^N\}$ using simulations

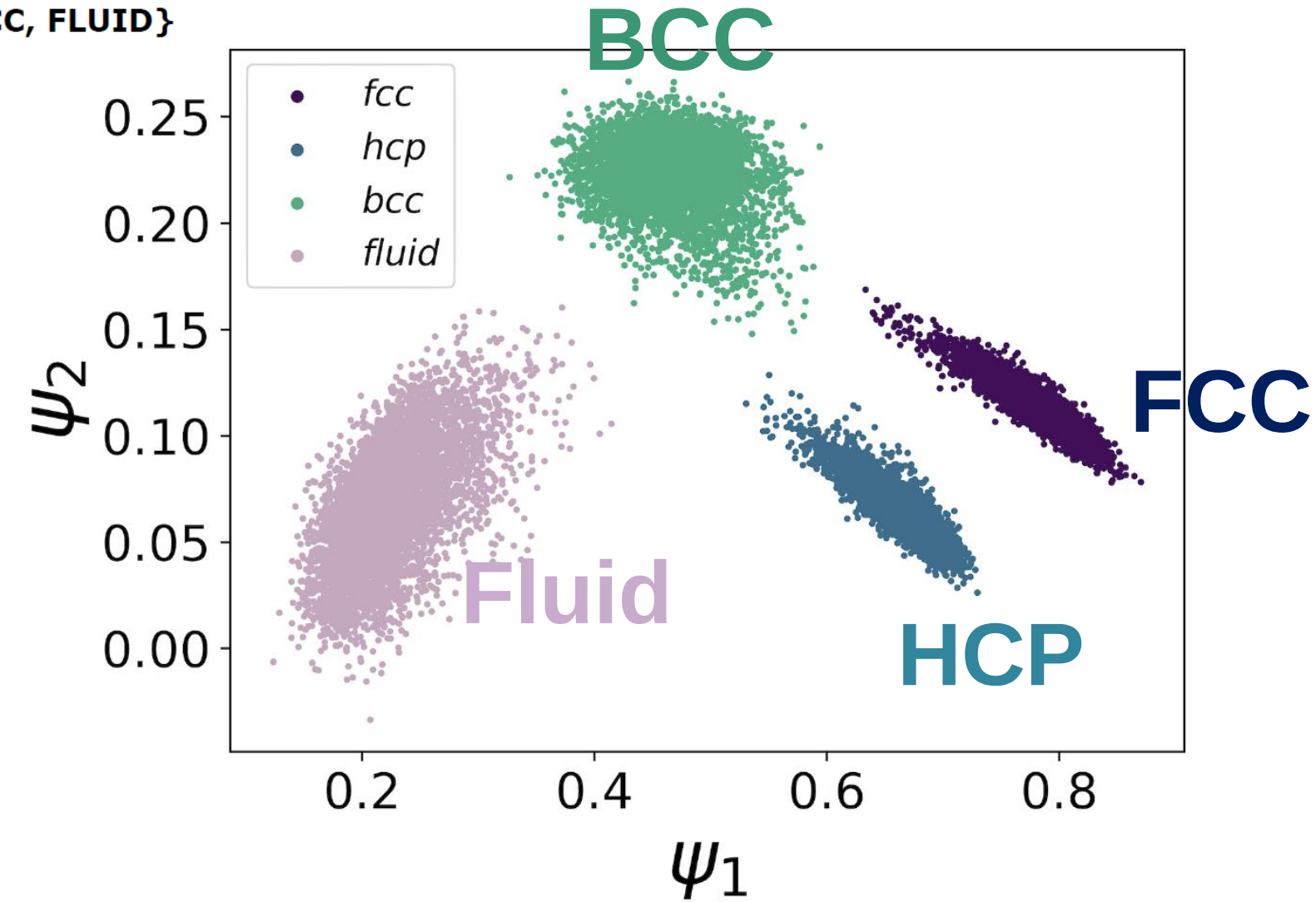
For every particle *i* in every configuration describe the **local environment** in terms of

$$\overline{\mathbf{Q}}(i) = \{\overline{q}_1(i), \overline{q}_2(i), \overline{q}_3(i), \cdots, \overline{q}_{12}(i)\}$$

Use **Principal Component Analysis** to find the relevant order parameters

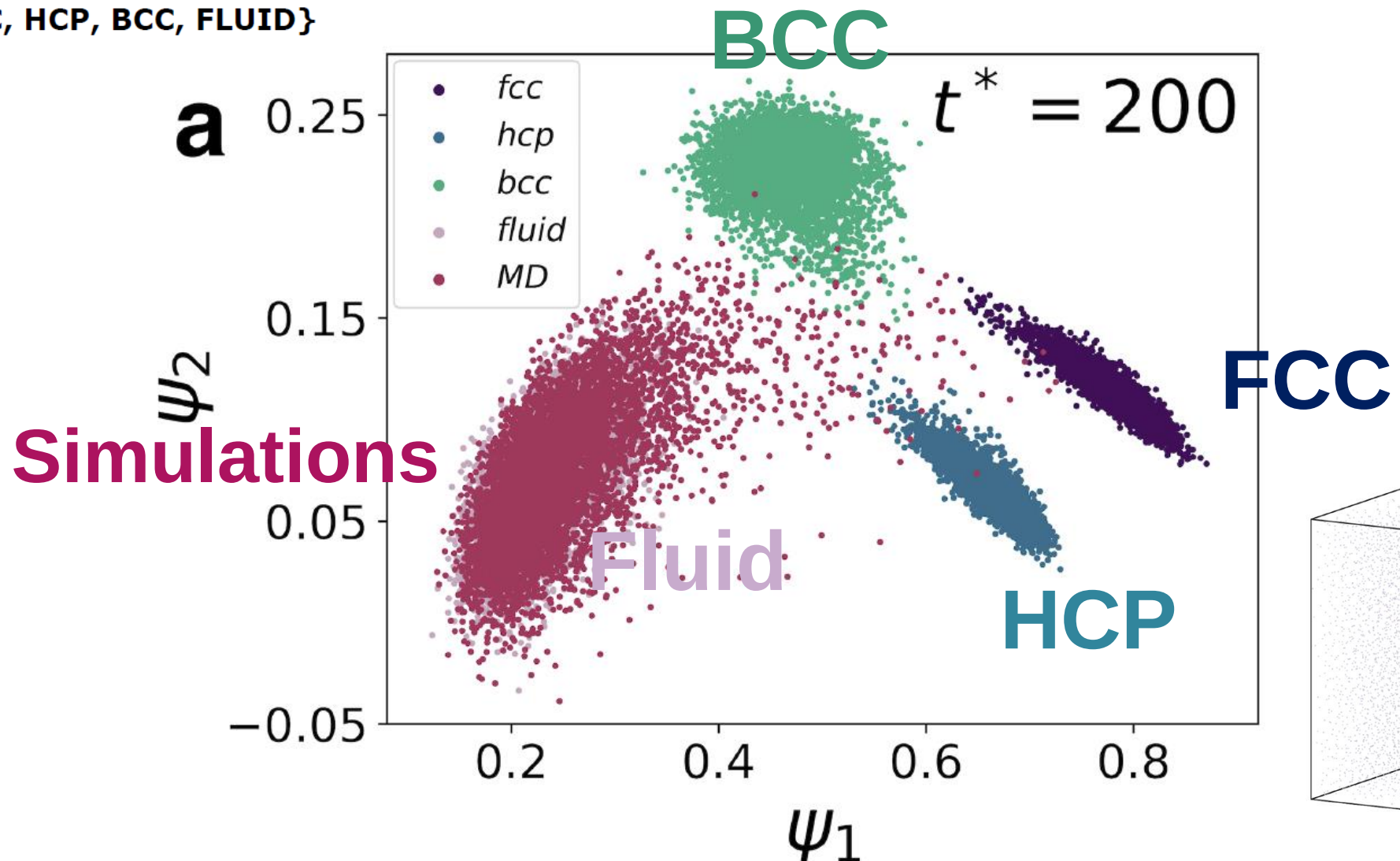
Principal Component Analysis:
finding order parameters

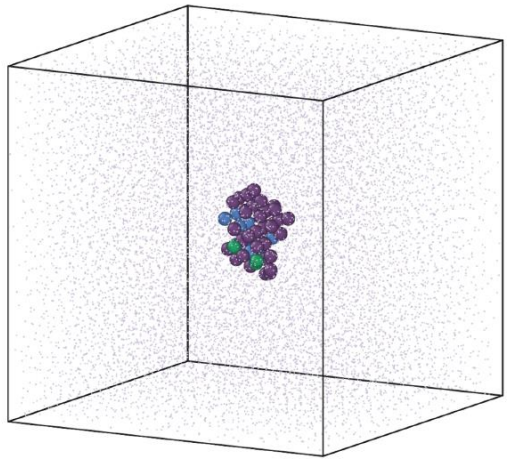
Dataset = {FCC, HCP, BCC, FLUID}



Principal Component Analysis:
finding order parameters

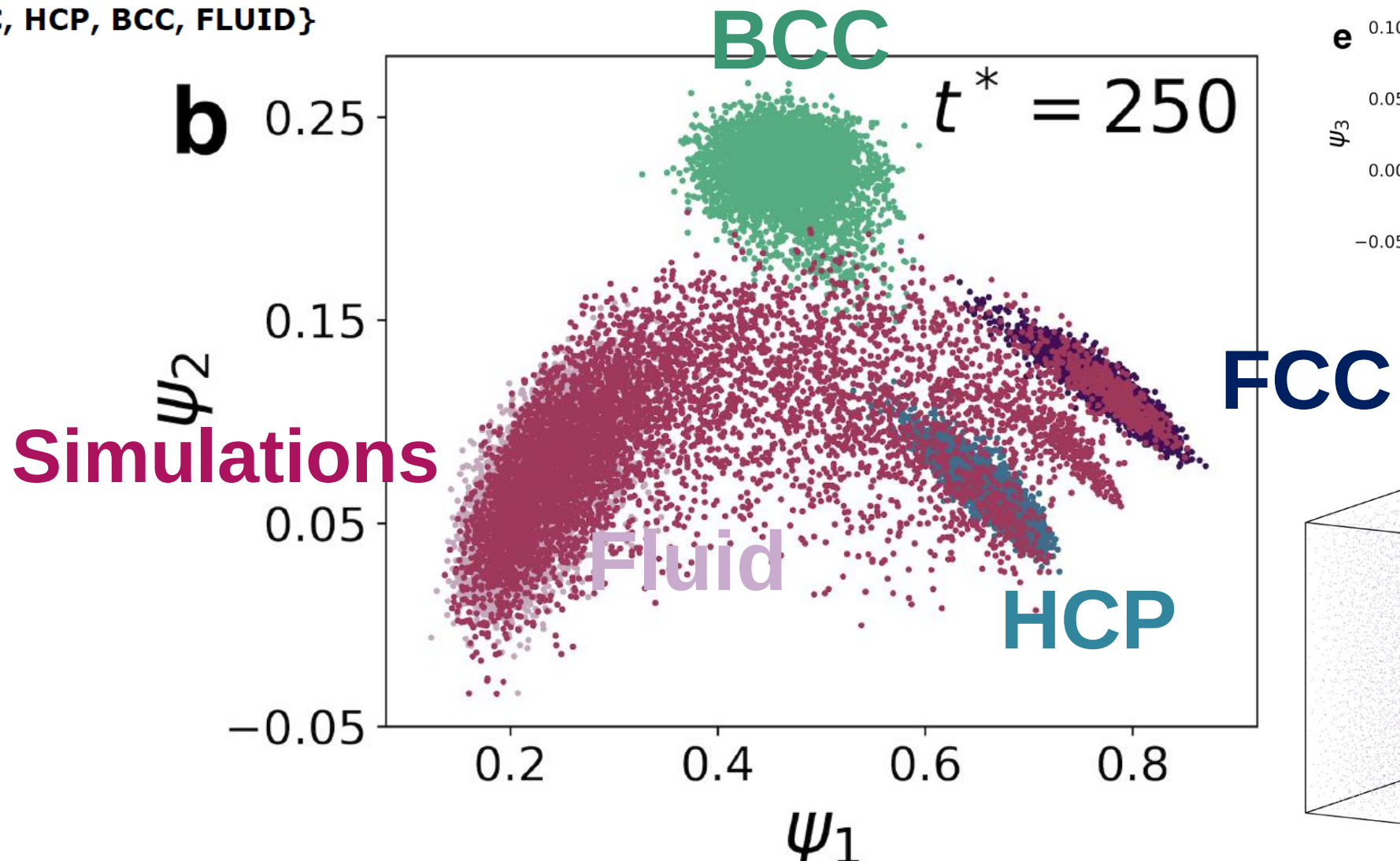
Dataset = {FCC, HCP, BCC, FLUID}

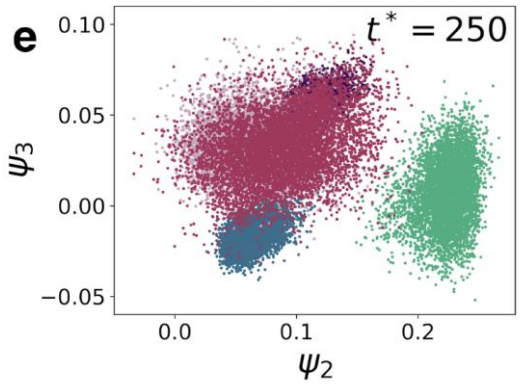


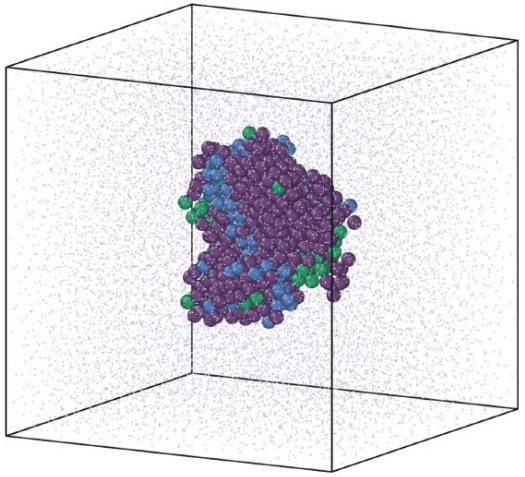


Principal Component Analysis:
finding order parameters

Dataset = {FCC, HCP, BCC, FLUID}

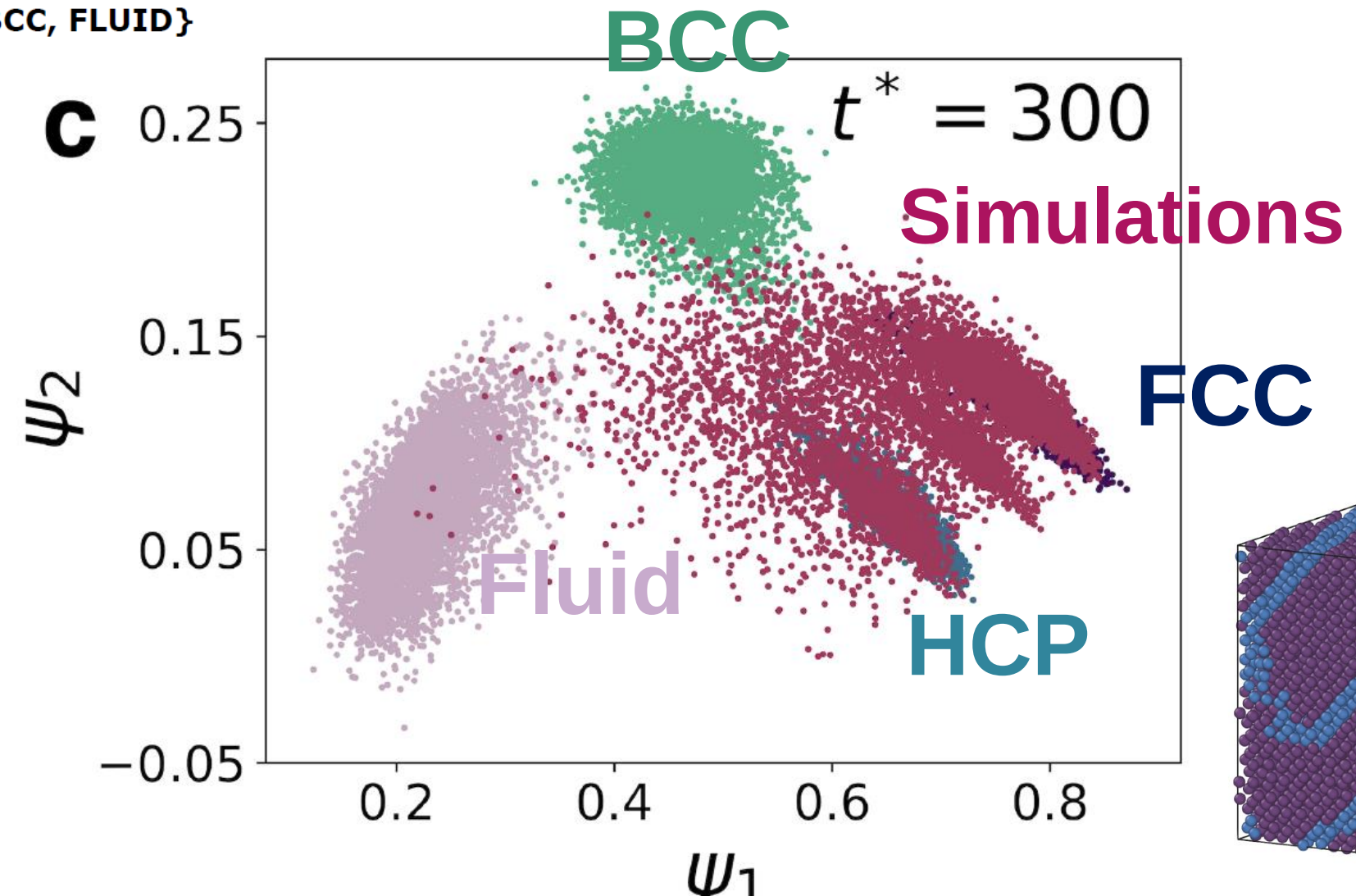


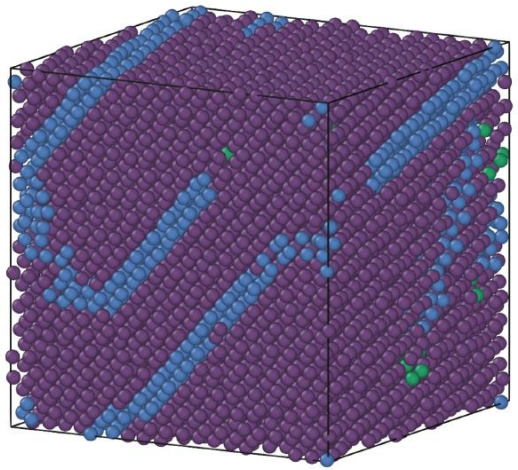




Principal Component Analysis:
finding order parameters

Dataset = {FCC, HCP, BCC, FLUID}





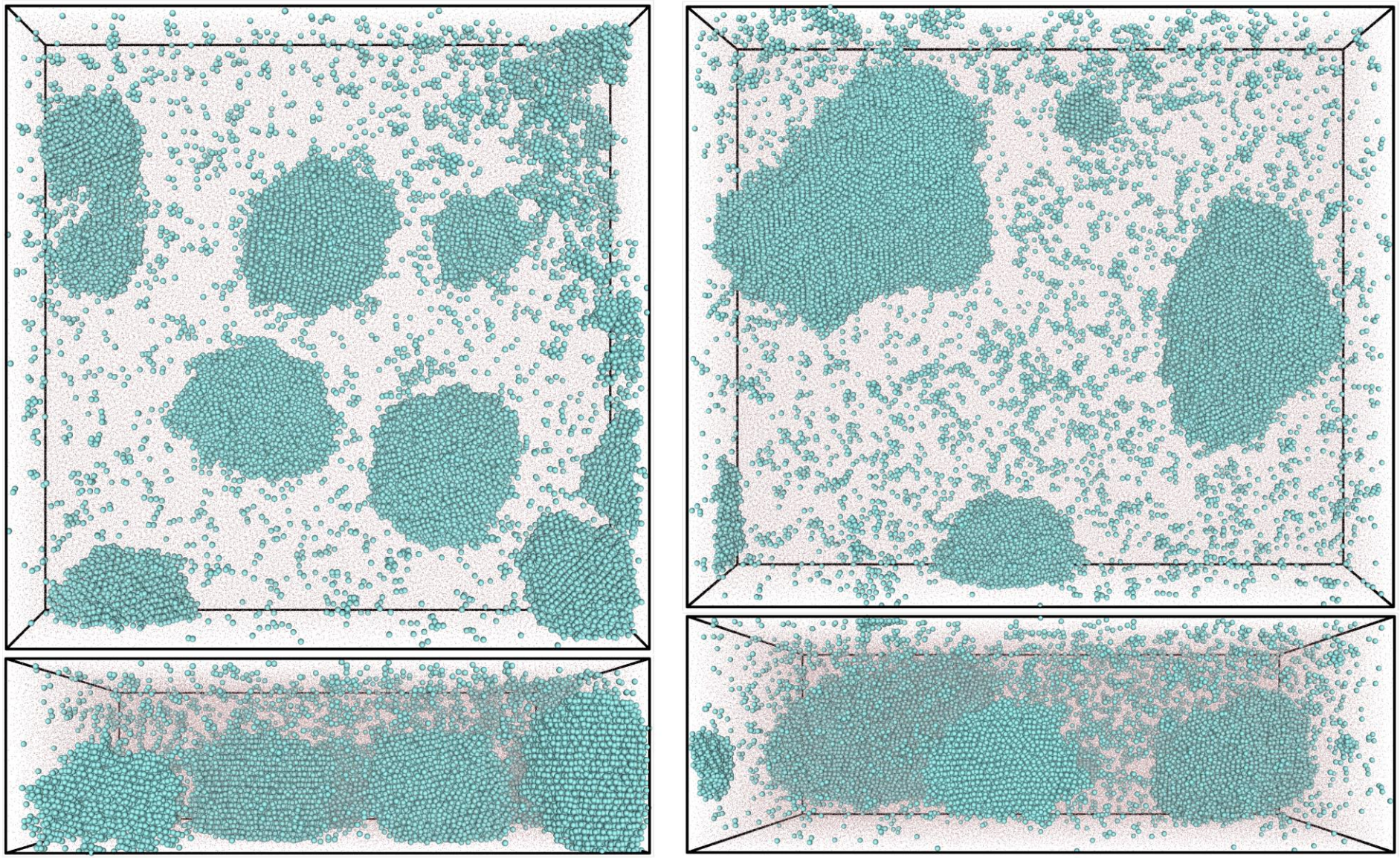
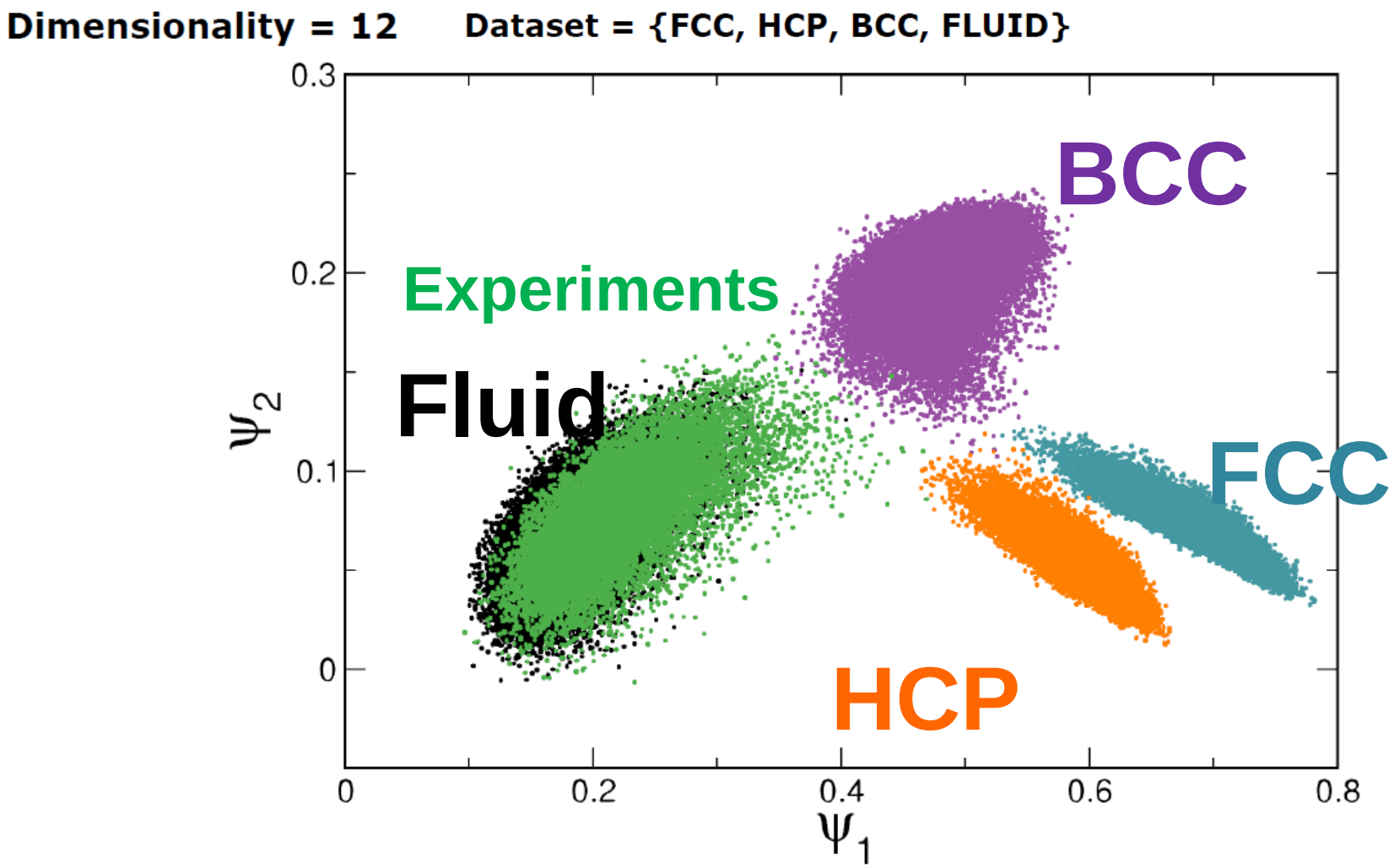


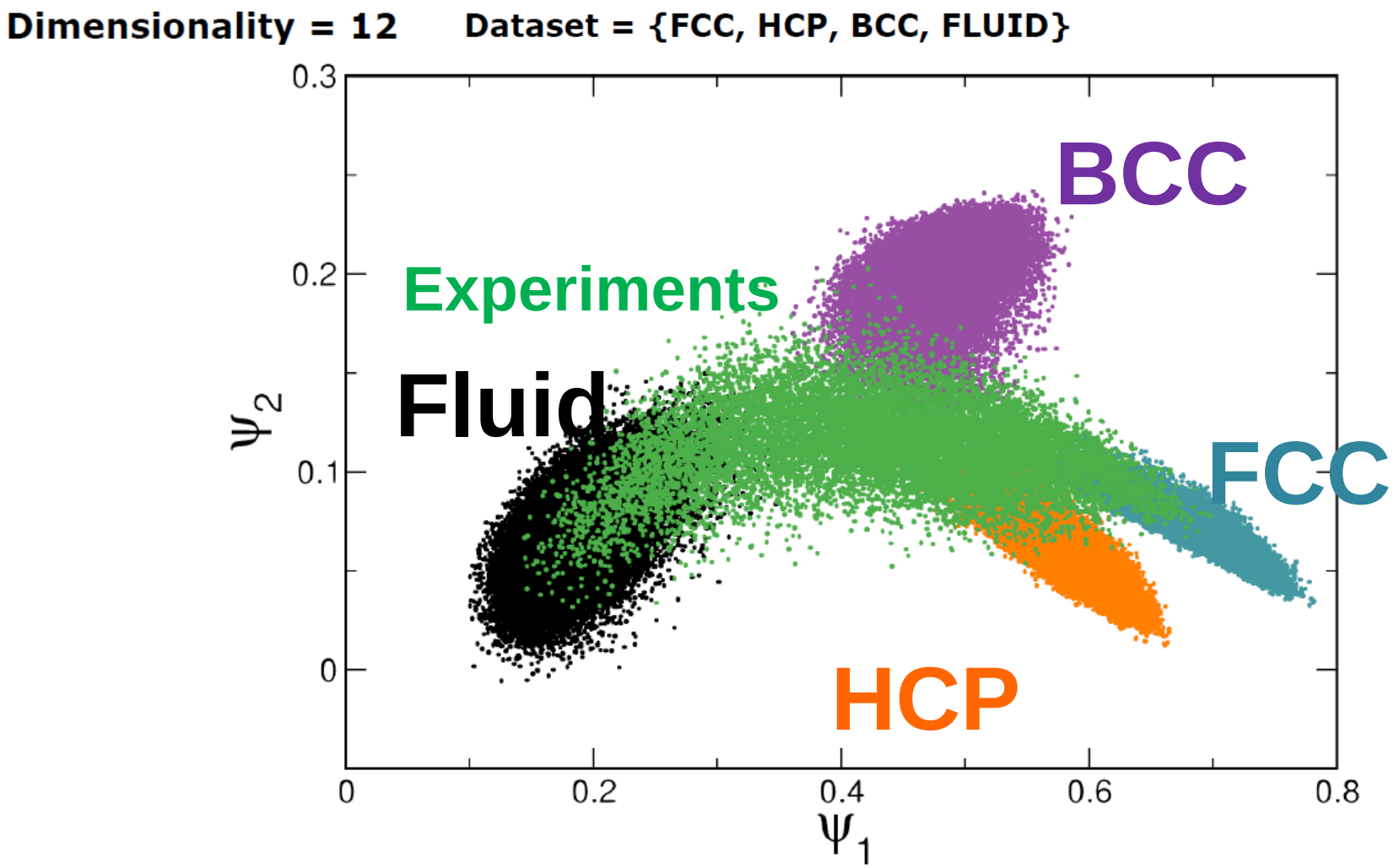
Figure 5.3: Computer reconstruction of the nuclei formed in hard sphere-like nucleation at $\phi = 0.542$. Computer rendering of the coordinates obtained from a confocal z-stack of a crystallizing PMMA dispersion after arrest, from the top (a) and side (b). Using bond order parameters the particles were considered crystalline (light blue) or fluid (pink). The size of the fluid particles are reduced to 1/10th of the original size, to increase the visibility of the structure.

Experiments: Ernest van der Wee, Anna Nikolaenkova, Alfons van Blaaderen, Arnout Imhof, Patrick Baesjou

Principal Component Analysis:
finding order parameters



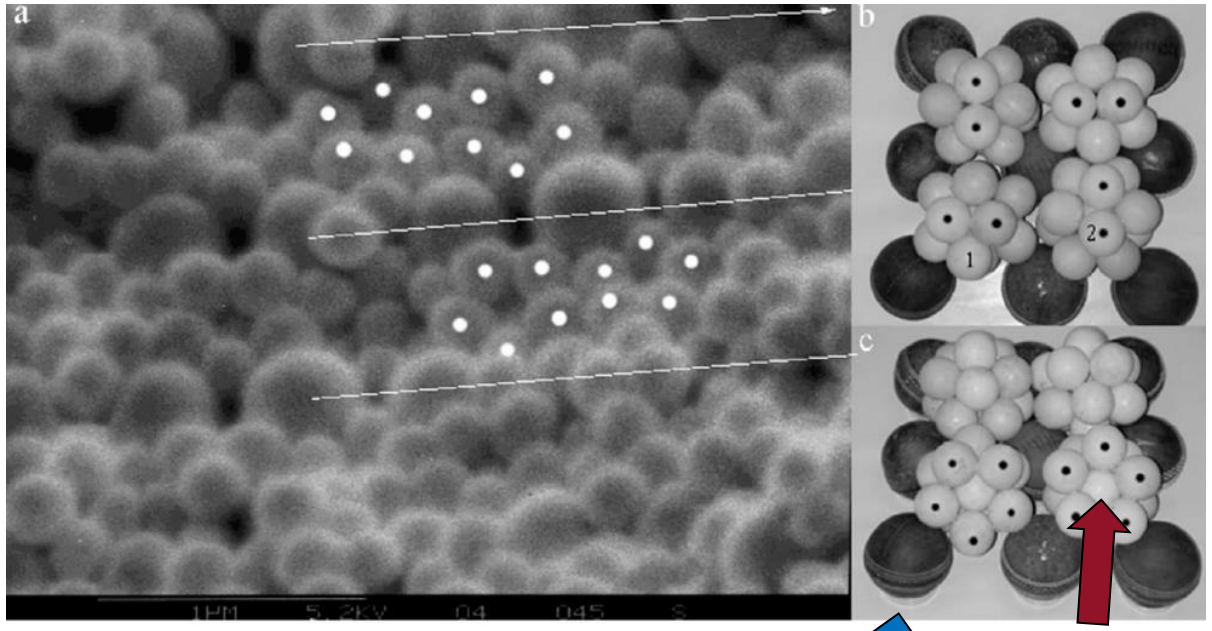
Principal Component Analysis:
finding order parameters



How does a binary crystal nucleate?

AB_{13} crystal for binary hard spheres with size ratio ~0.5

PMMA colloidal hard spheres
100-400 nm

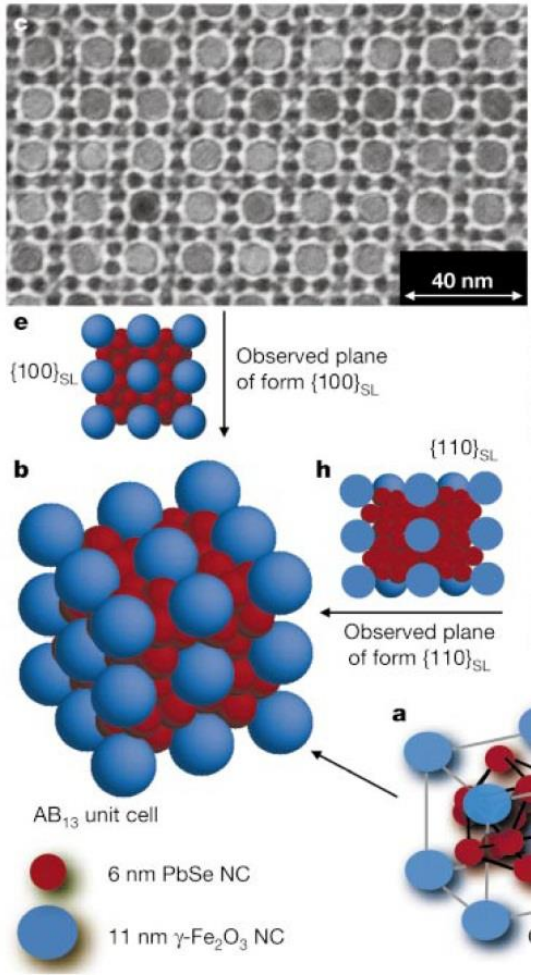


Schofield, Pusey, Radcliffe, Phys. Rev. E 72, 031407 (2005)

Simple cubic of large spheres but with a large lattice spacing

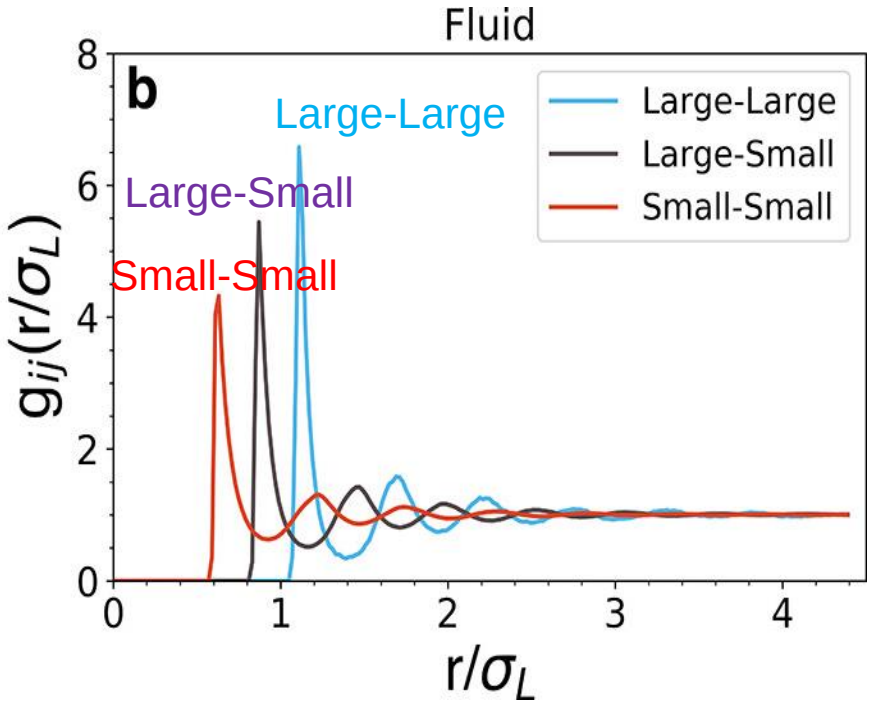
Icosahedral clusters of 13 small spheres

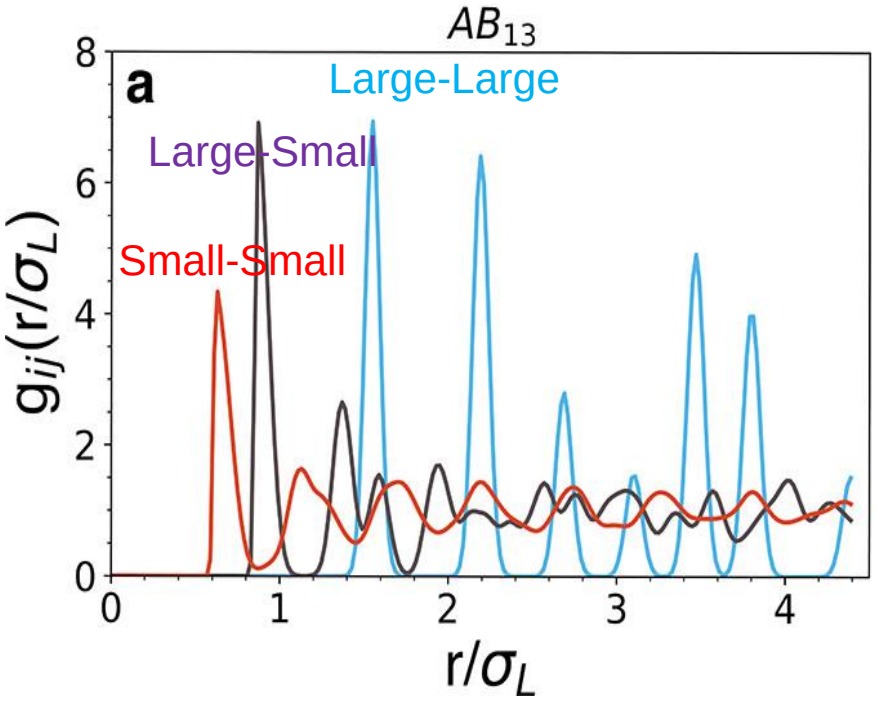
6 nm PbSe semiconductor quantumdots and 11 nm Fe_2O_3 magnetic nanocrystals



Redl, Cho, Murray, O'Brien, Nature 423, 968 (2003)

Structure of the fluid and AB_{13} crystal



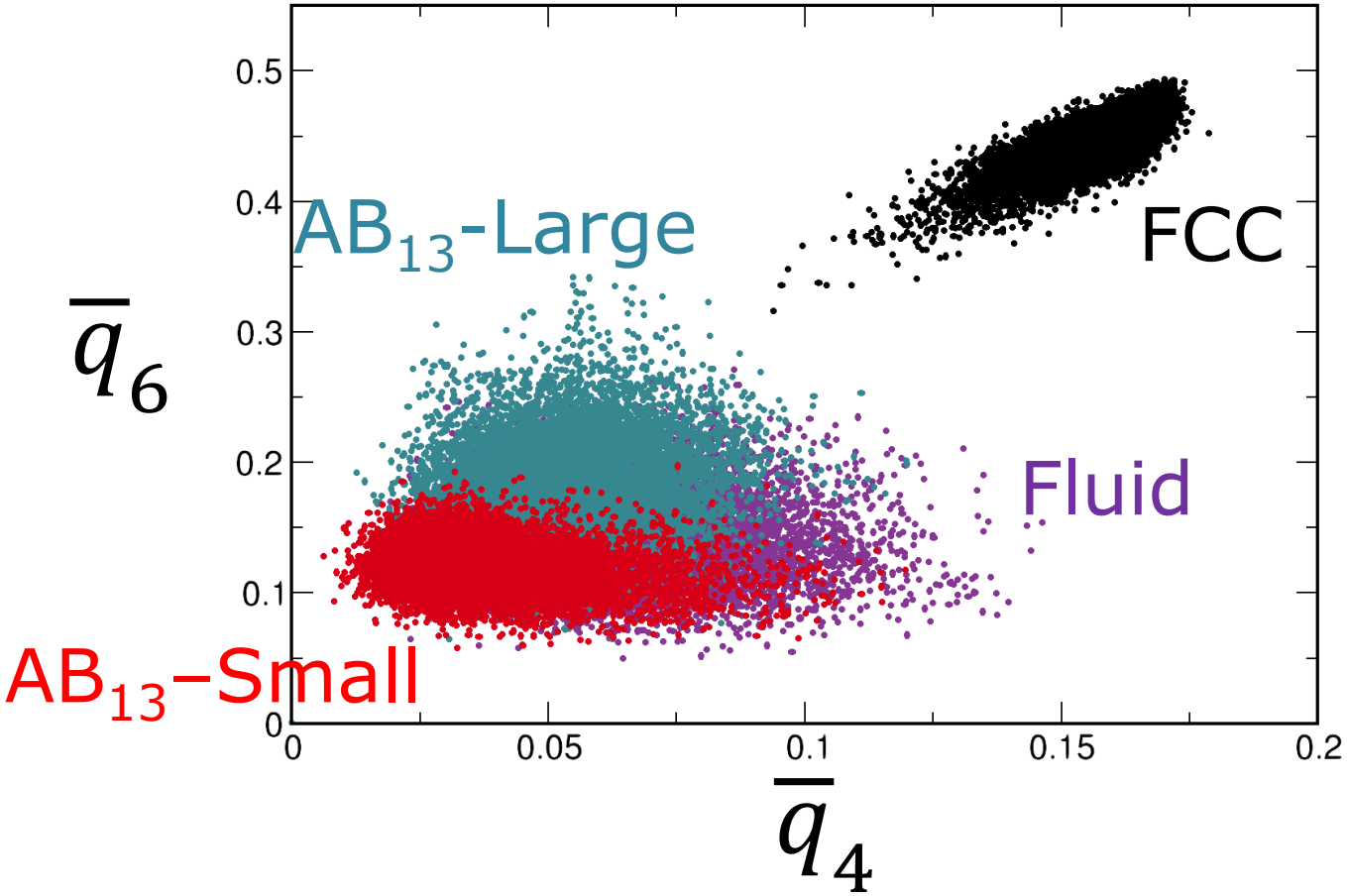


Structure of the <u>small spheres</u> very similar in the fluid and solid phase

<u>Large spheres</u> at an extreme large distance in the solid phase compared to the fluid phase

Bond order parameters

Local environment defined by all particles within a certain distance irrespective of species



Large and small particles of the AB_{13} crystal indistinguishable from the fluid phase

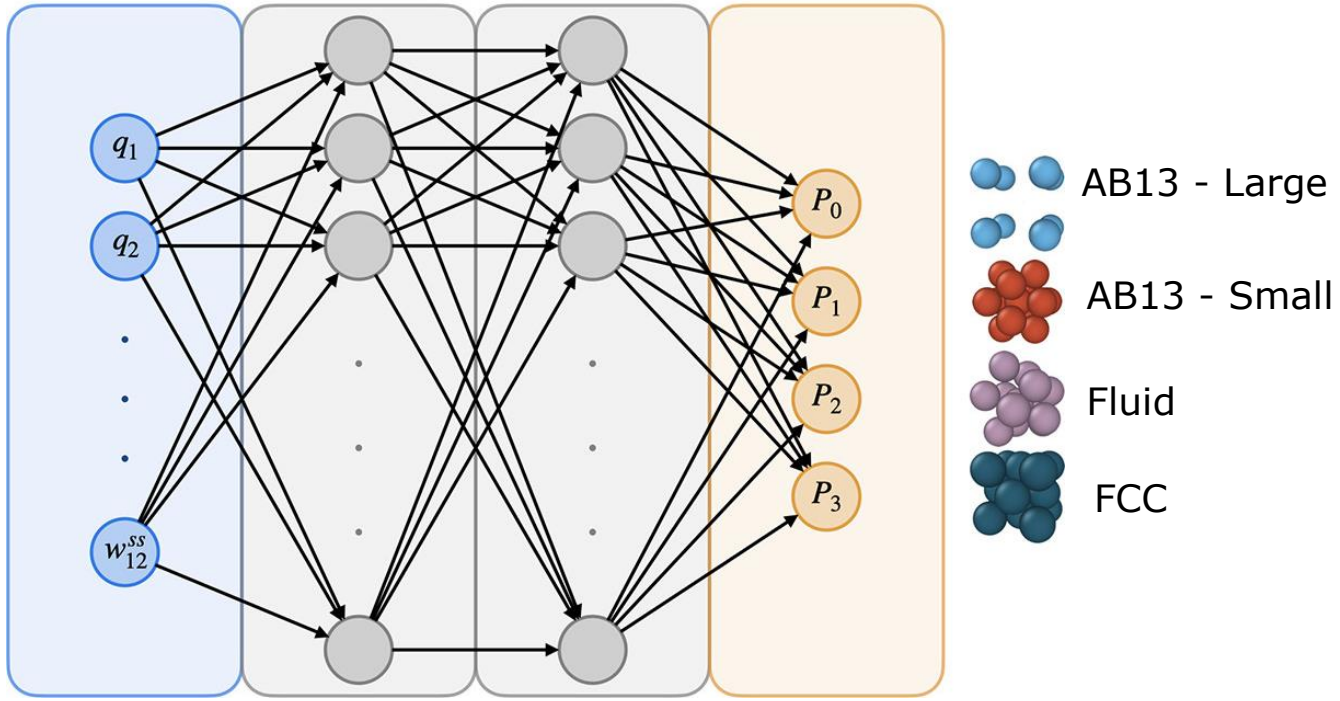
Can we solve this using Neural Networks?

Neural Network

Input: 36 nodes **2 Hidden: 72 nodes** **Output: 4 nodes**

$$(\{q_l(i)\}, \{w_{l'}(i)\}, \{q_l^{ss}(i)\}, \{w_{l'}^{ss}(i)\})$$

$l \in [1,12]$
l' only even values

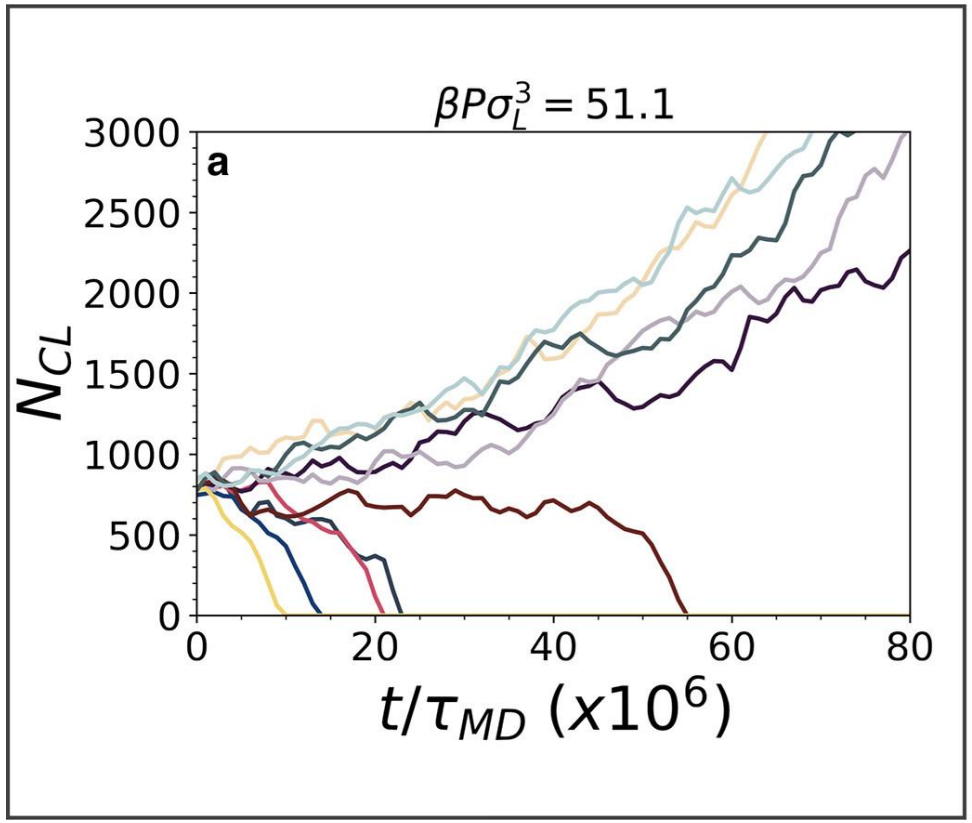


Weights and biases are **optimised** using **stochastic gradient descent**

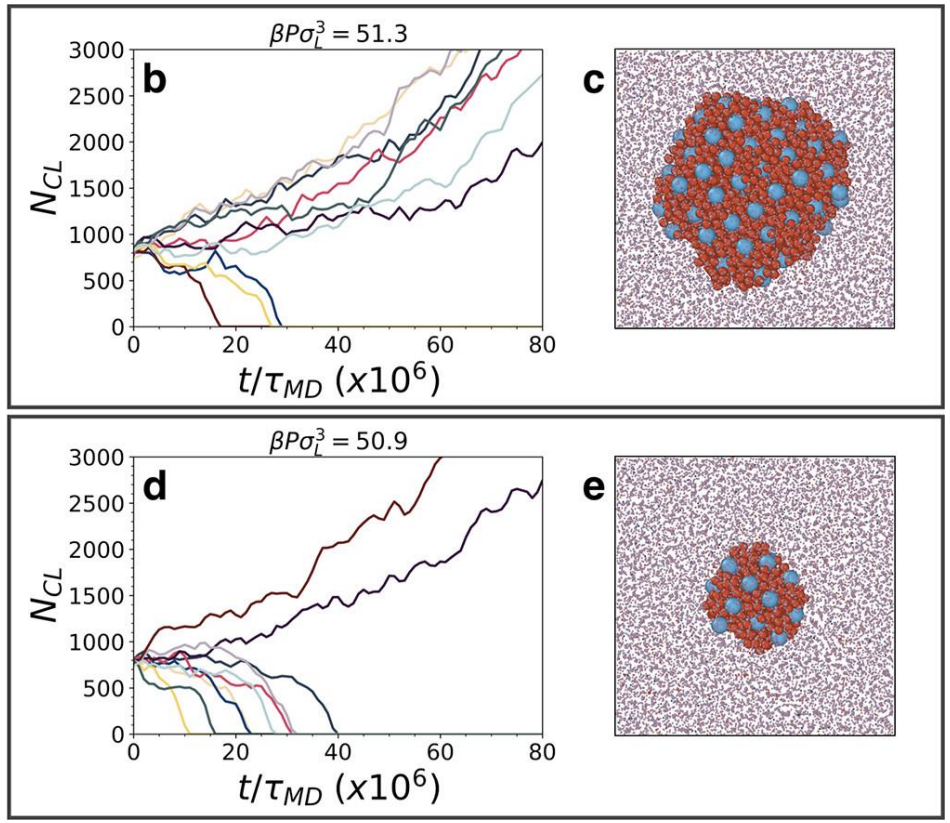
At the end of the training we reach an accuracy of 97.8-100%!

Nucleation of AB_{13}

Critical pressure



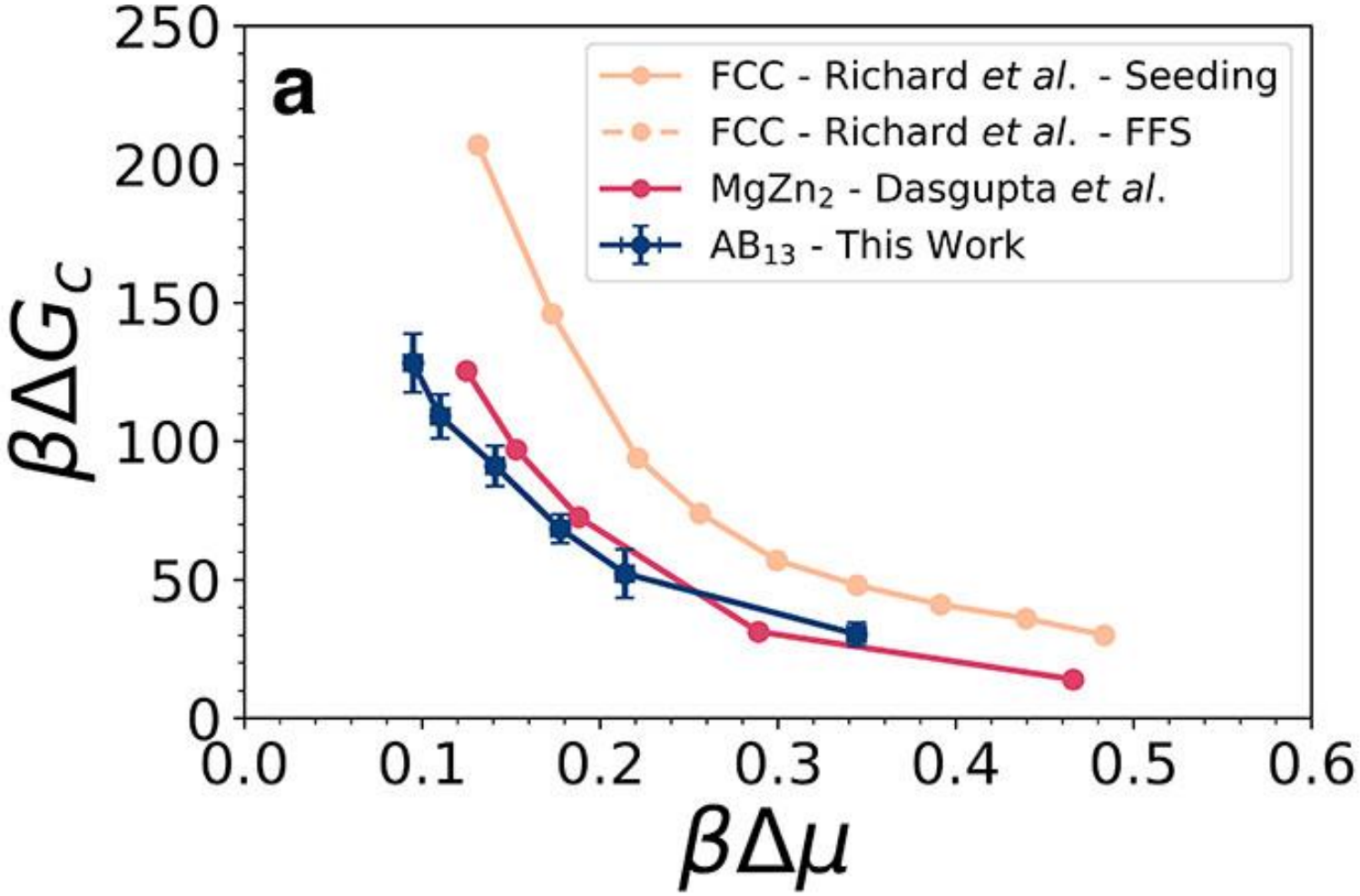
Pressure too high



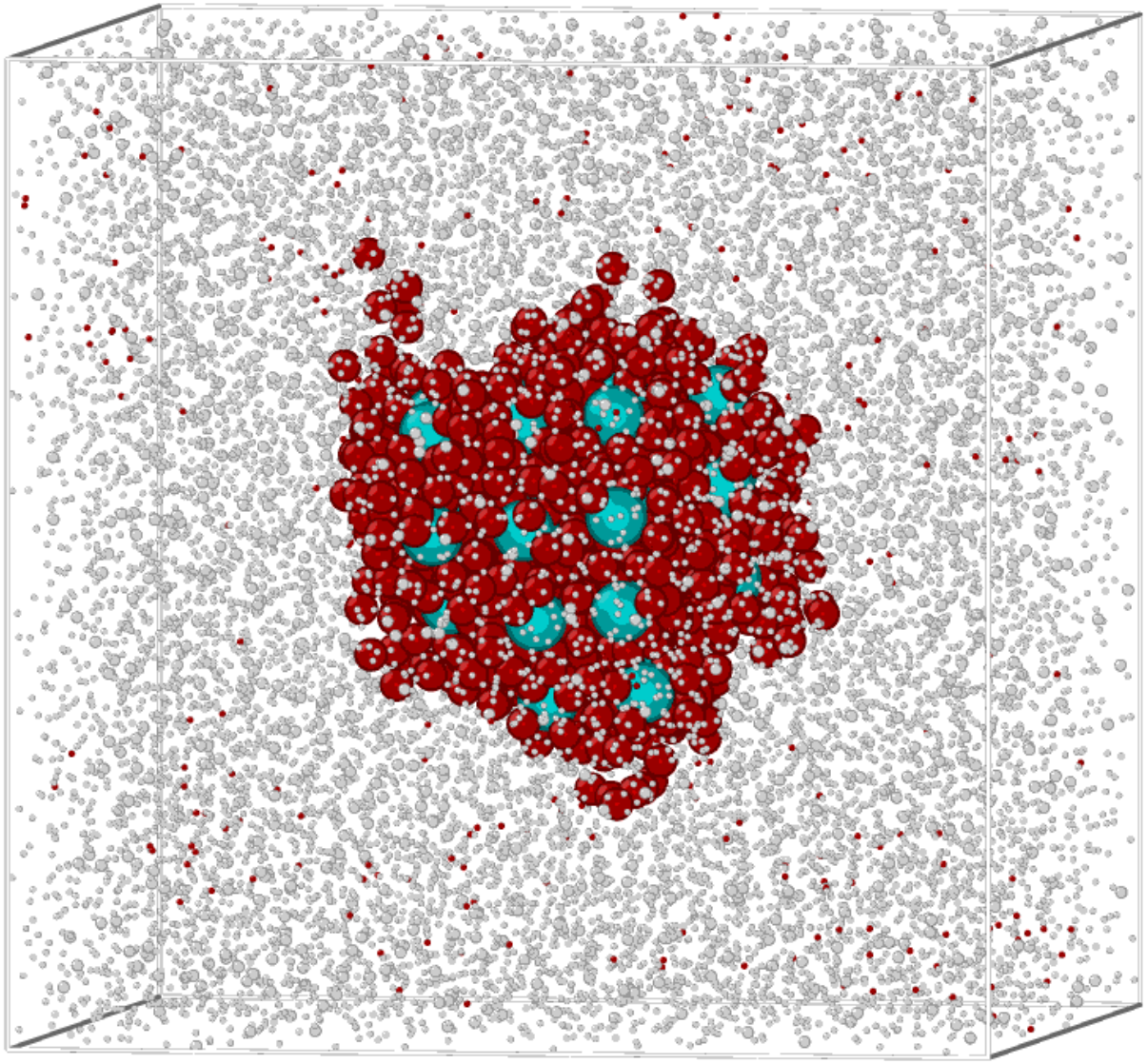
Pressure too low

Seeding approach: Espinosa, Vega, Valeriani, Sanz, JCP 144, 034501 (2016)

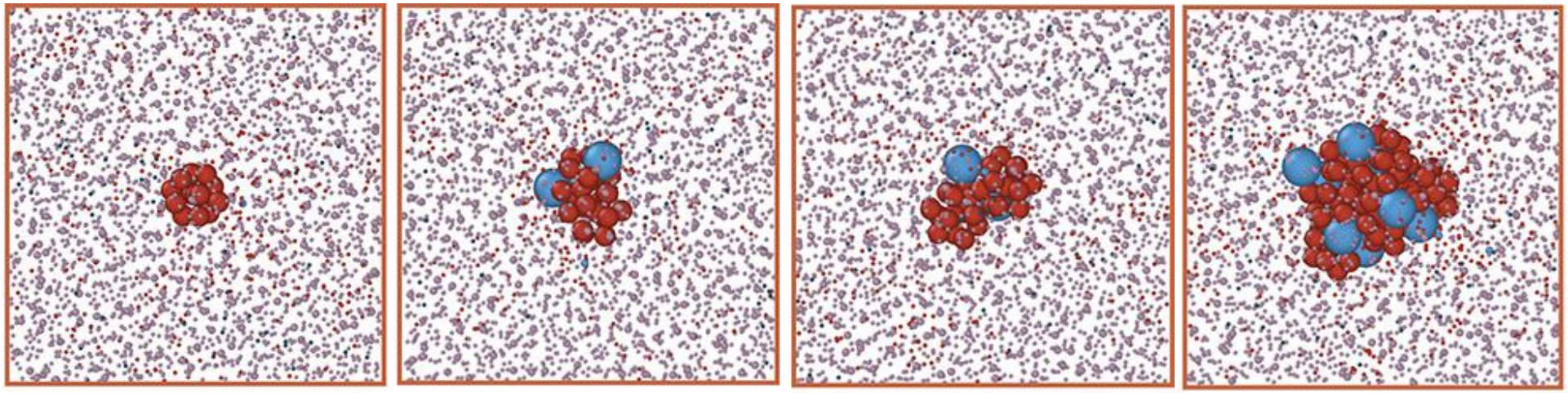
Nucleation barrier height of AB_{13}

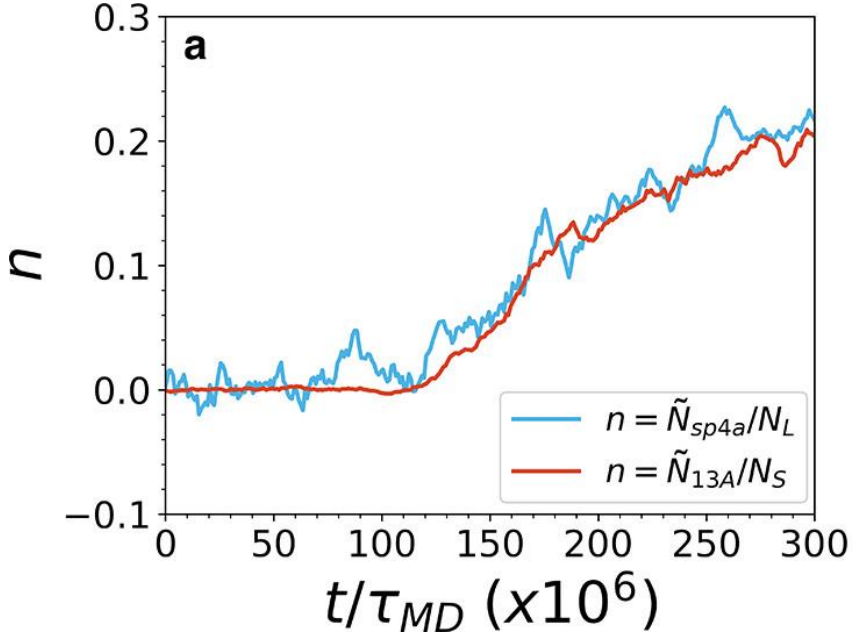


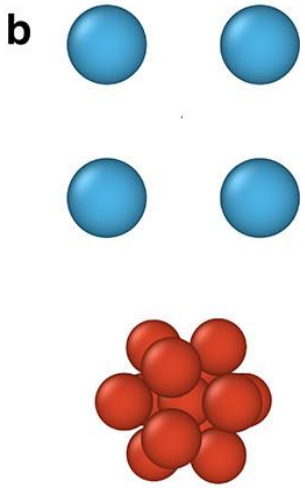
Spontaneous nucleation of AB_{13}



Birth of an AB_{13} Nucleus







Co-assembly of large spheres and icosahedral clusters of small spheres

Coli and Dijkstra, ACSNano 15, 4335 (2021)

Outline

- Soft Matter

- Classification and Finding Order parameters
  - Nucleation of Hard Spheres
  - Birth of a AB_{13} crystal of hard spheres

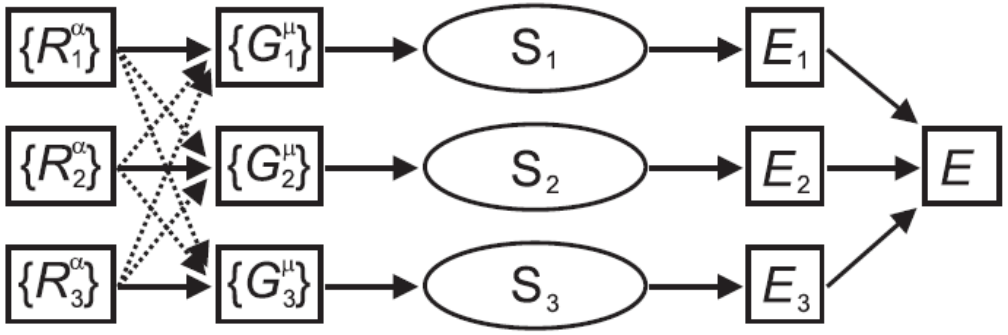
- Coarse-graining
  - Machine Learning potentials for colloid-polymer mixtures

- Inverse Design
  - Quasicrystals, liquid crystals, crystals

PRL 98, 146401 (2007) PHYSICAL REVIEW LETTERS week ending 6 APRIL 2007

Generalized Neural-Network Representation of High-Dimensional Potential-Energy Surfaces

Jörg Behler and Michele Parrinello
Department of Chemistry and Applied Biosciences, ETH Zurich, USI-Campus, Via Giuseppe Buffi 13, CH-6900 Lugano, Switzerland
(Received 27 September 2006; published 2 April 2007)

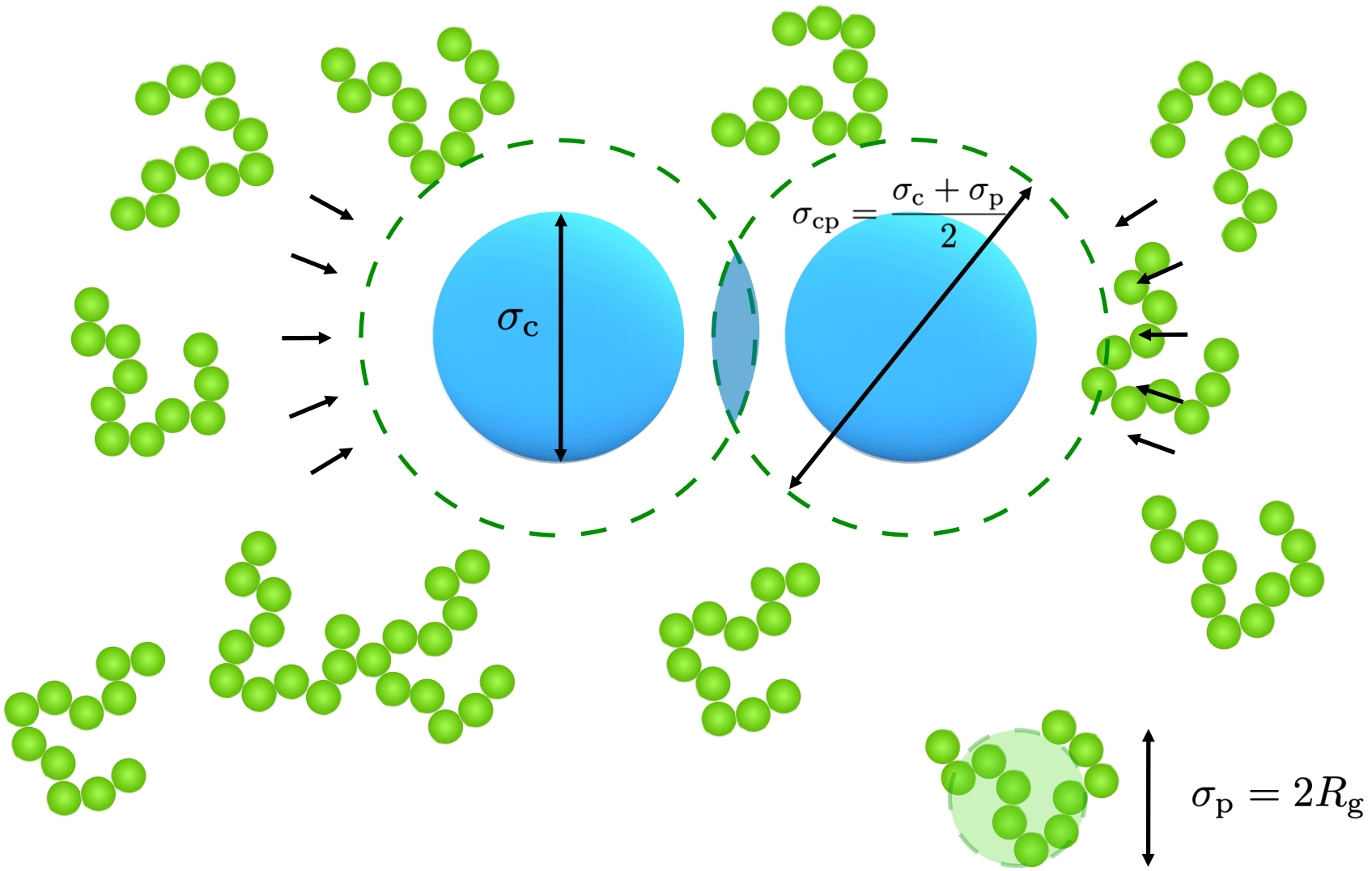


Radial symmetry function $$G_i^1 = \sum_{j \neq i}^{\text{all}} e^{-\eta(R_{ij}-R_s)^2} f_c(R_{ij}).$$

Angular symmetry function $$G_i^2 = 2^{1-\zeta} \sum_{j,k \neq i}^{\text{all}} (1 + \lambda \cos\theta_{ijk})^\zeta \times e^{-\eta(R_{ij}^2+R_{ik}^2+R_{jk}^2)} f_c(R_{ij}) f_c(R_{ik}) f_c(R_{jk}),$$

Cut-off function $$f_c(R_{ij}) = \begin{cases} 0.5 \times \left[\cos\left(\frac{\pi R_{ij}}{R_c}\right) + 1\right] & \text{for } R_{ij} \leq R_c, \\ 0 & \text{for } R_{ij} > R_c. \end{cases}$$

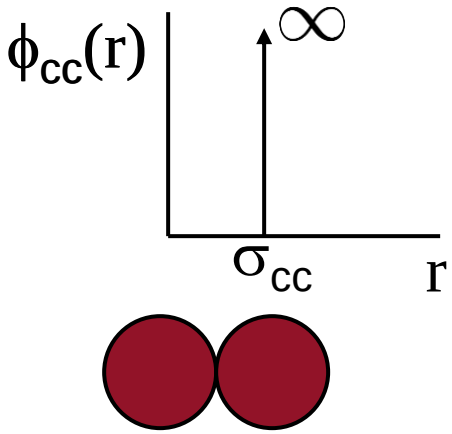
Colloid-polymer mixtures: depletion



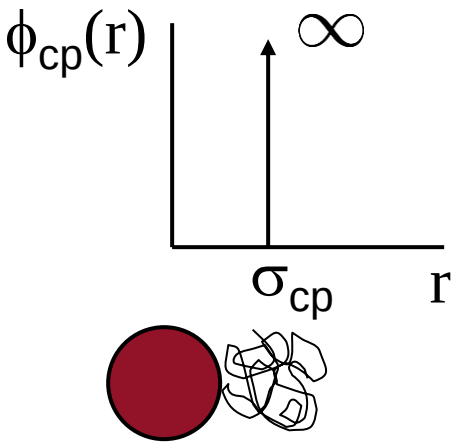
Size ratio
$q = \sigma_p/\sigma_c$

Asakura-Oosawa-(Vrij) model for colloid-polymer mixtures

Colloid-colloid



Colloid-polymer



Polymer-polymer

$\phi_{pp}(r) = 0$

Effective Hamiltonian

The total Hamiltonian of a binary mixture consists of a sum of interaction terms:

$$\mathcal{H} = \mathcal{H}_{\rm cc} + \mathcal{H}_{\rm cp} + \mathcal{H}_{\rm pp}$$

$$\mathcal{H}_{\rm cc} = \sum_{i<j}^{N_{\rm c}} \phi_{\rm cc}(R_{ij}), \qquad \mathcal{H}_{\rm cp} = \sum_{i=1}^{N_{\rm c}} \sum_{j=1}^{N_{\rm p}} \phi_{\rm cp}(|\boldsymbol{R}_i - \boldsymbol{r}_j|), \qquad \mathcal{H}_{\rm pp} \equiv 0.$$

It is convenient to consider the system in the (N_c,V,z_p,T) ensemble. The thermodynamic potential is:

$$\exp[-\beta F] = \sum_{N_p=0}^{\infty} \frac{z_p^{N_p}}{N_c! \Lambda_c^{3N_c} N_P!} \int_V d\boldsymbol{R}^{N_{\rm c}} \int_V d\boldsymbol{r}^{N_{\rm p}} \exp[-\beta(\mathcal{H}_{\rm cc} + \mathcal{H}_{\rm cp})]$$

$$= \frac{1}{N_c! \Lambda_c^{3N_c}} \int_V d\boldsymbol{R}^{N_{\rm c}} \exp[-\beta(\underbrace{\mathcal{H}_{\rm cc} + \Omega}_{\mathcal{H}_{\rm eff}})],$$

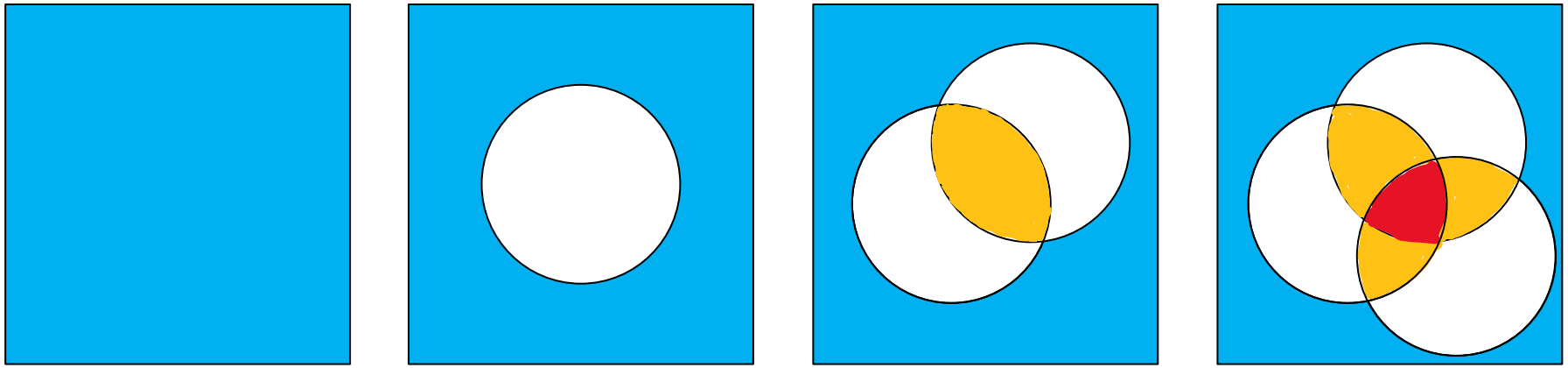
$$\mathcal{H}_{\text{eff}} = \mathcal{H}_{\text{cc}} + \Omega$$

Ω is the grand potential of the sea of polymer in the external field of a fixed configuration of N_c colloids at coordinates $\{\mathbf{R}\}$.

$$\Omega = -z_p V_{\text{f}}(\{\boldsymbol{R}_i\})$$

Free volume: available volume for the ideal polymer

$$V_{\text{f}} = V_{\text{f}}^{(0)} + \sum_{i=1}^{N_{\text{c}}} V_{\text{f}}^{(1)}(\boldsymbol{R}_i) + \sum_{i<j}^{N_{\text{c}}} V_{\text{f}}^{(2)}(\boldsymbol{R}_i, \boldsymbol{R}_j) + V_{\text{f}}^{(3+)}$$

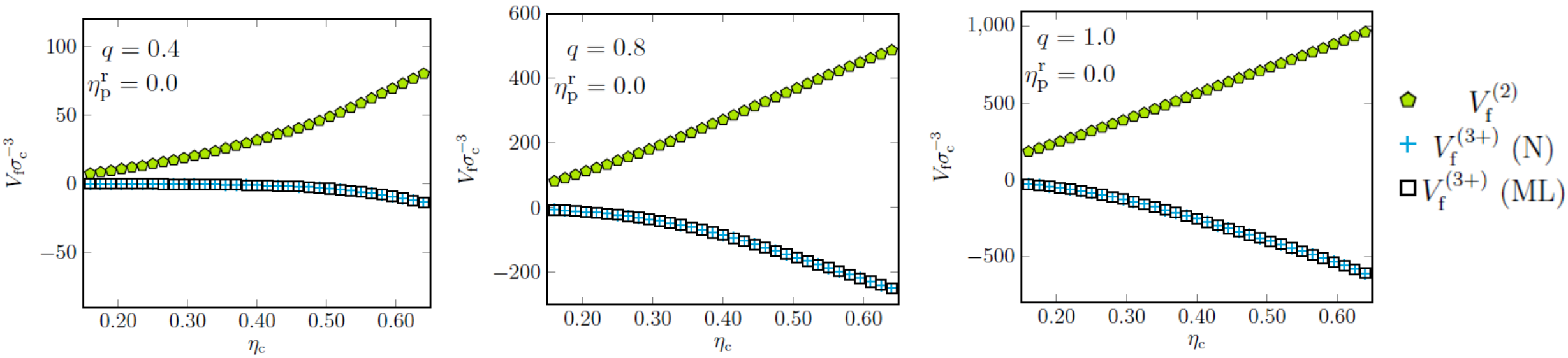


$$V_{\text{f}}^{(0)} \qquad \sum_{i=1}^{N_{\text{c}}} V_{\text{f}}^{(1)}(\boldsymbol{R}_i) \qquad \sum_{i<j}^{N_{\text{c}}} V_{\text{f}}^{(2)}(\boldsymbol{R}_i, \boldsymbol{R}_j) \qquad V_{\text{f}}^{(3+)}$$

Fitting with Behler-Parrinello Symmetry functions

We calculate V_f^{3+} using a numerical evaluation for a system at polymer fugacity $z_p=0$ and colloid packing fraction for η_c = [0.15,0.65]

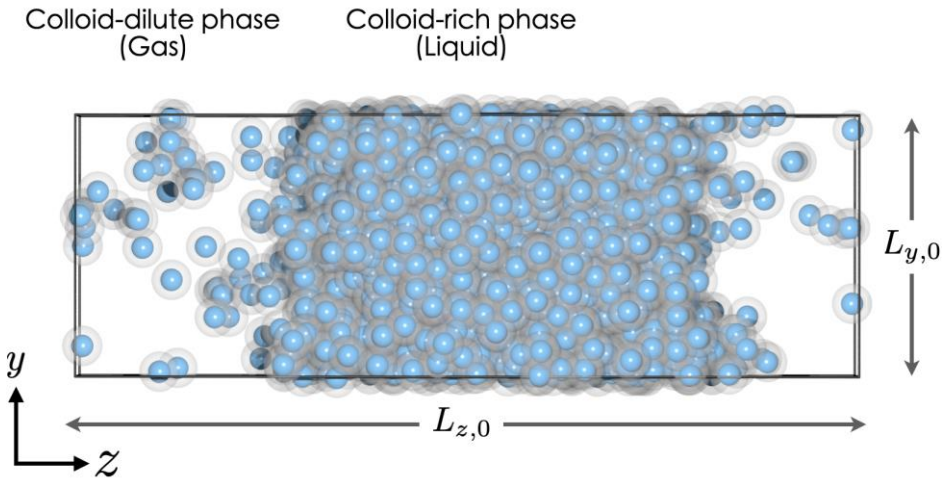
Many different λ, η, ξ $$G_I^3 = 2^{1-\xi} \sum_{j,k\neq i} (1+\lambda\cos\theta_{ijk})^{\xi}\, e^{\eta\left(r_{ij}^2+r_{ik}^2+r_{jk}^2\right)} f_c(r_{ij}) f_c(r_{ik}) f_c(r_{jk})$$

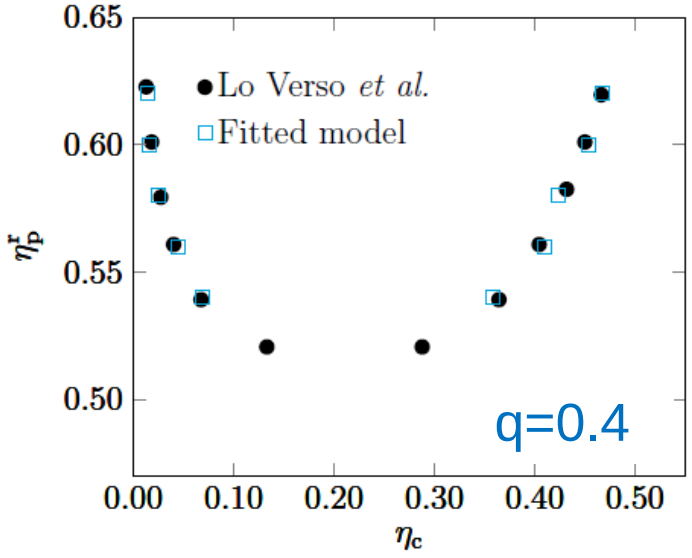


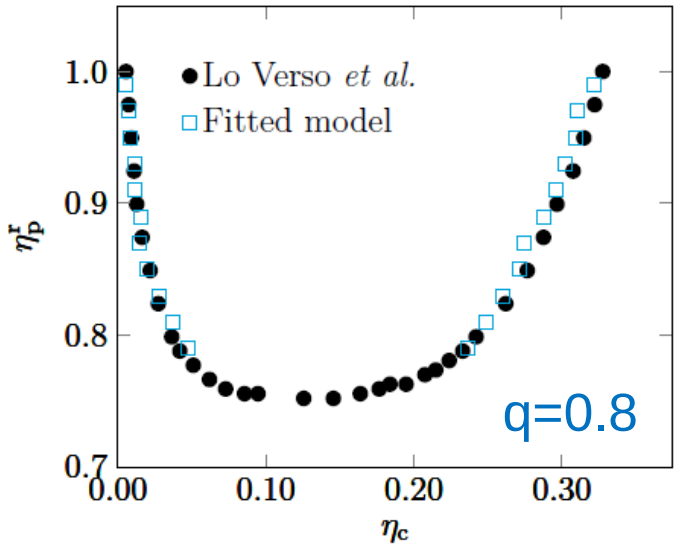
ML potentials speed up the calculations with 2 till 4 orders of magnitude and can be assumed to be state-independent

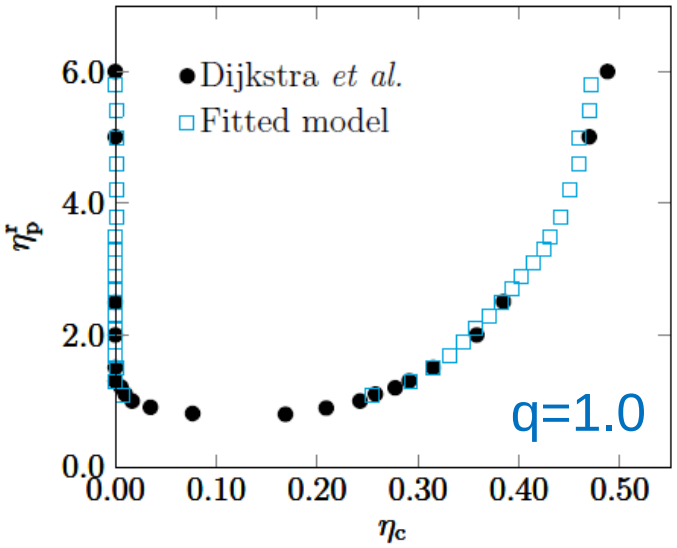
Phase behavior: Colloidal Gas-Liquid binodals

Campos Villalobos, Boattini, Filion, Dijkstra, J. Chem. Phys. (2021)









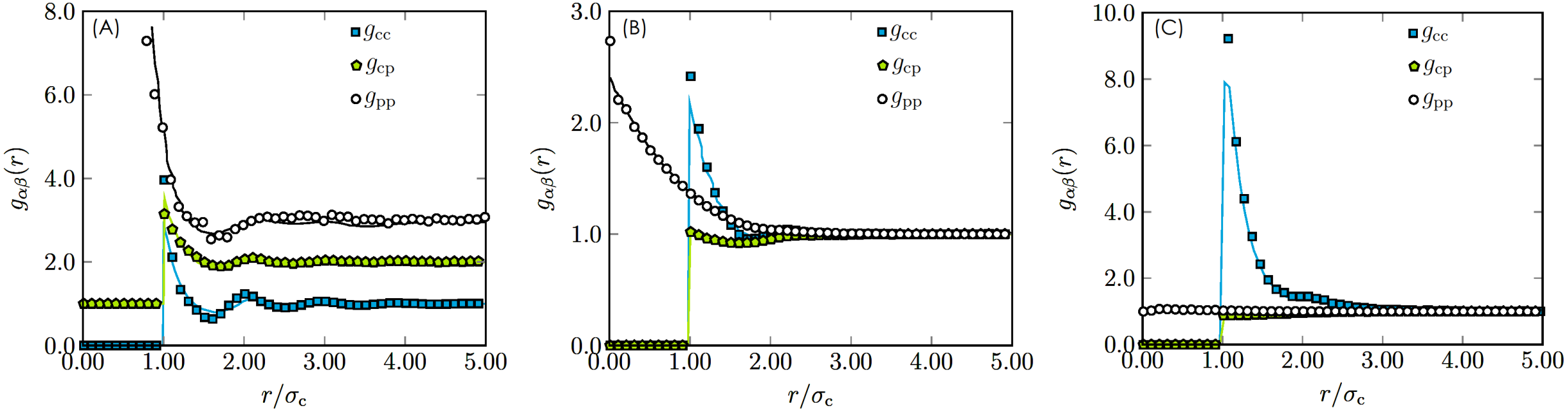
Black dots: Binary mixture[1] or an effective one-component system [2]
Blue squares: Our ML model

[1]Phys. Rev. E 73, 061407 (2006)
[2]Phys. Rev. E, 73, 041404 (2006)

Structure: Pair correlation functions

q=1.0

Lines: Numerical model
Symbols: Our ML model



Campos Villalobos, Boattini, Filion, Dijkstra, J. Chem. Phys. (2021)

[1]Phys. Rev. E, 73, 041404 (2006)

Outline

- Soft Matter

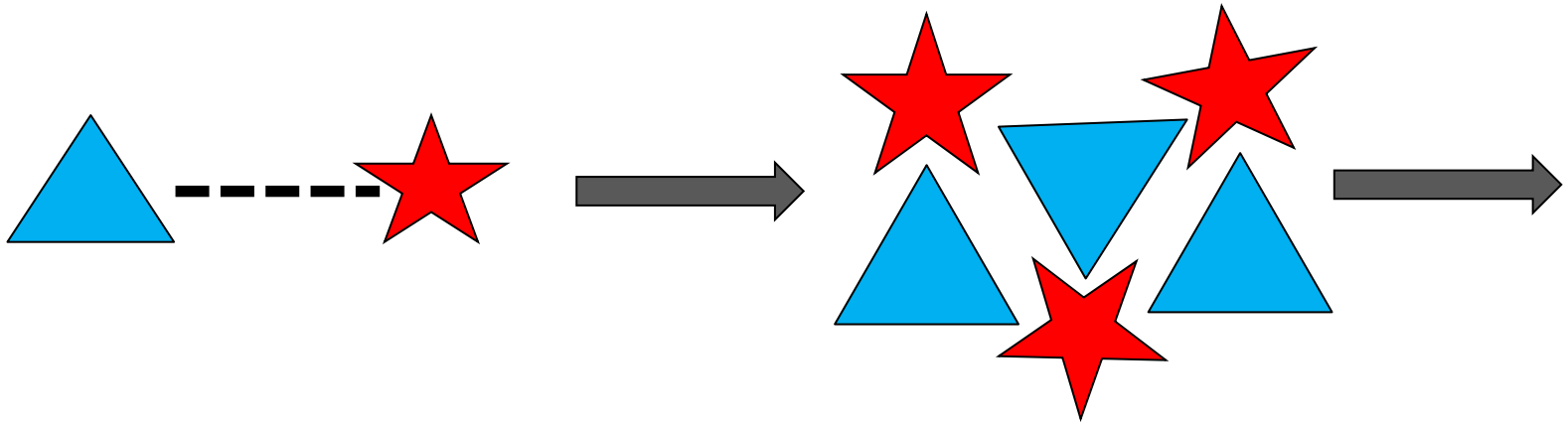
- Classification and Finding Order parameters
  - Nucleation of Hard Spheres
  - Birth of a AB_{13} crystal of hard spheres

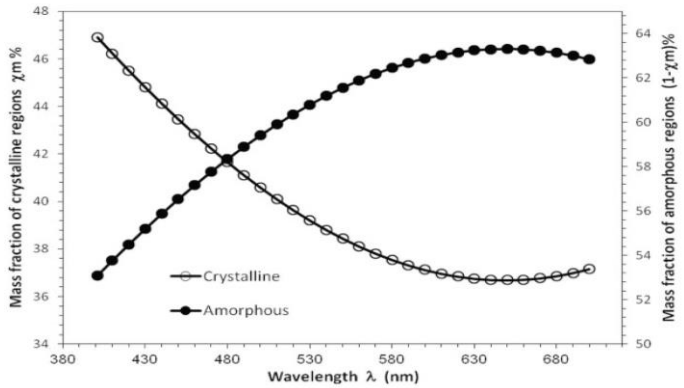
- Coarse-graining
  - Machine Learning potentials for colloid-polymer mixtures

- Inverse Design
  - Quasicrystals, liquid crystals, crystals

Forward Design

Standard approach in Statistical mechanics





Shape and interactions

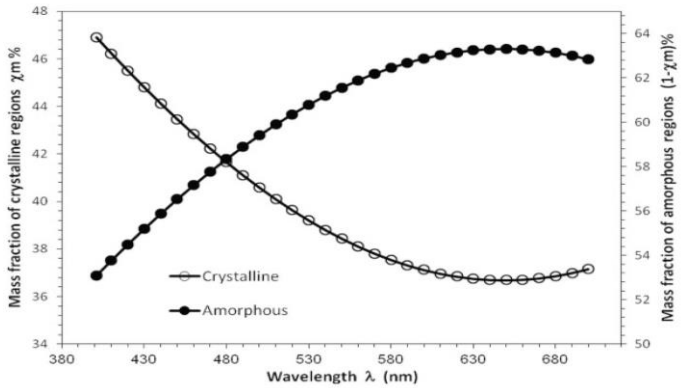
Structure

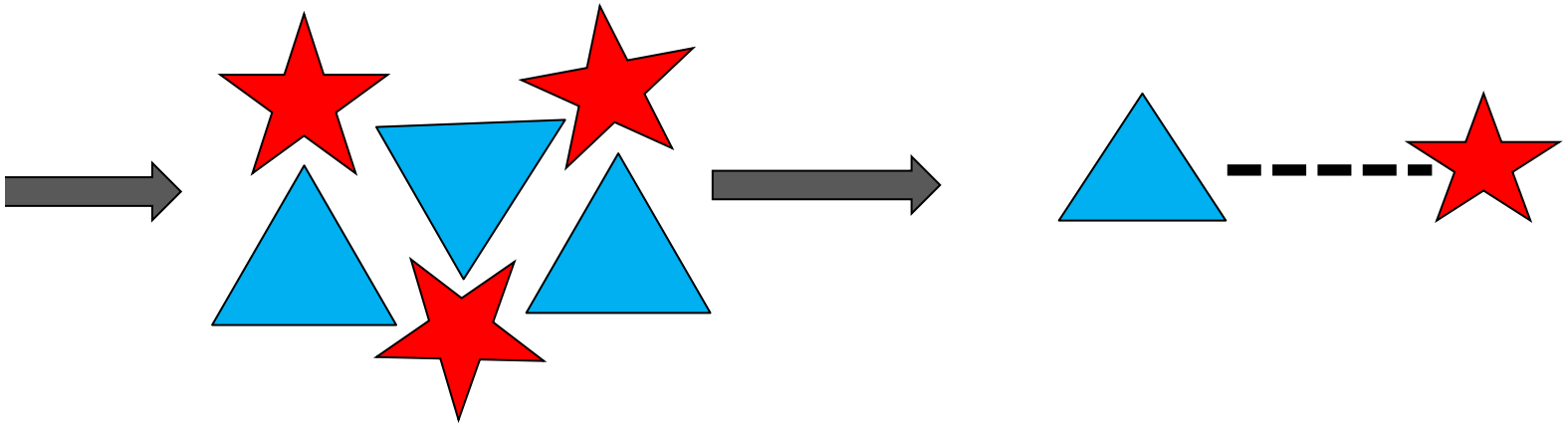
Properties

are limitless quantities!!!!

Forward Design

Standard approach in Statistical mechanics





Properties

Structure

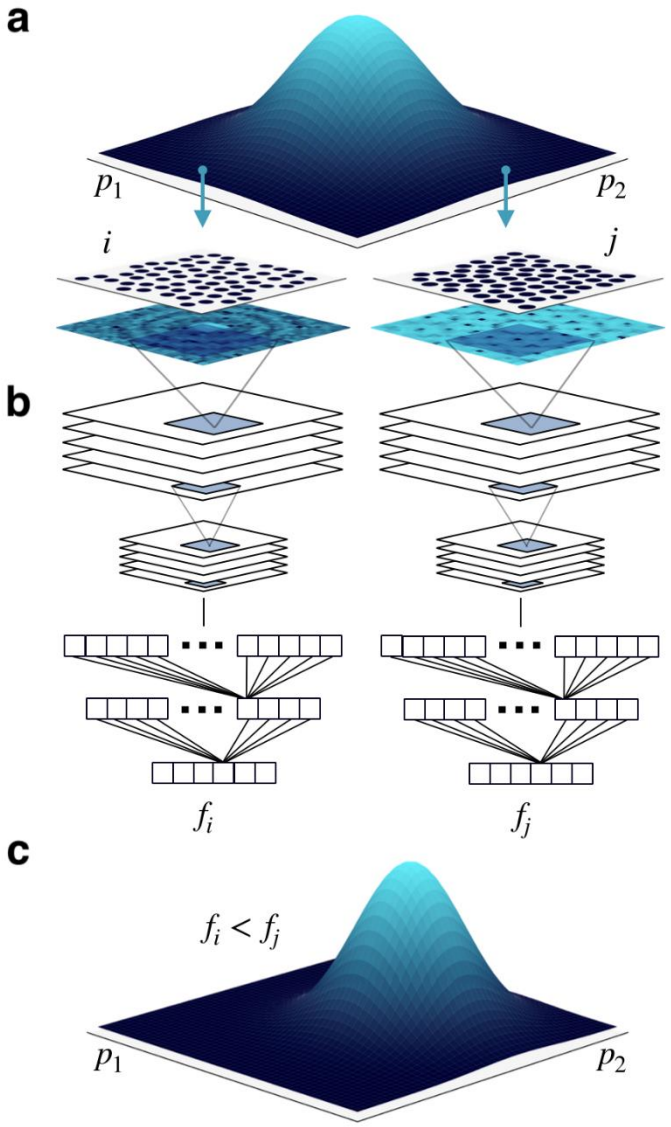
Shape and interactions

Can we inverse design a target structure?

Inverse Design Method

Three ingredients:

- The Model with free design parameters that will be optimized
- Scheme to update the design parameters
- Order parameter to distinguish the target phase from competing phases – fitness function

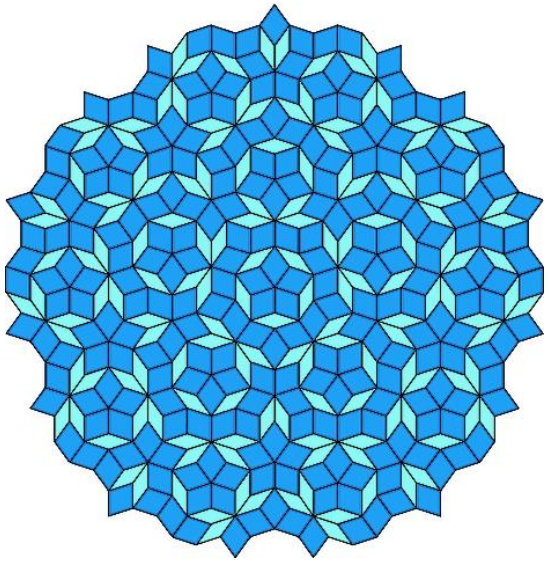


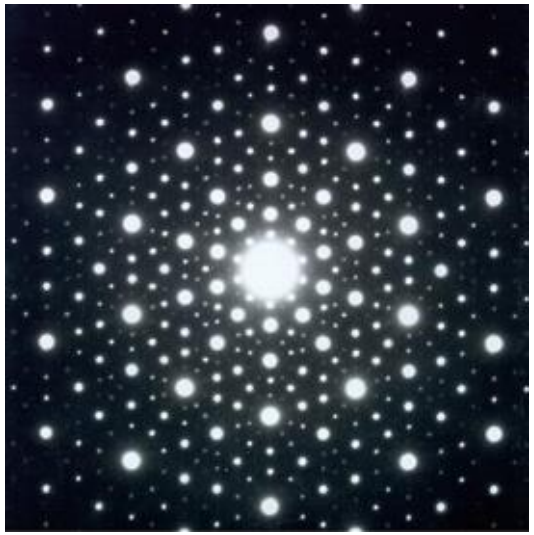
Quasicrystals

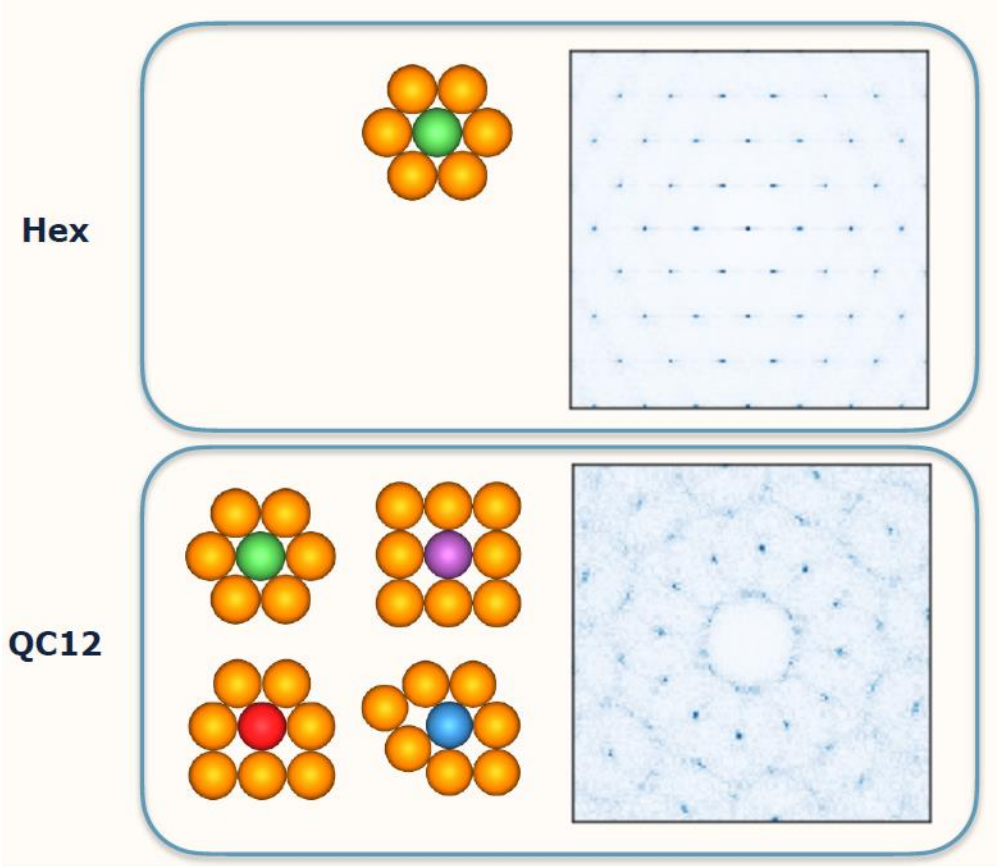
- Nature's impossible phase

- lacks translational symmetry: their is no repeating unit

- exhibit rotational symmetry as characterized by sharp peaks in their diffraction patterns





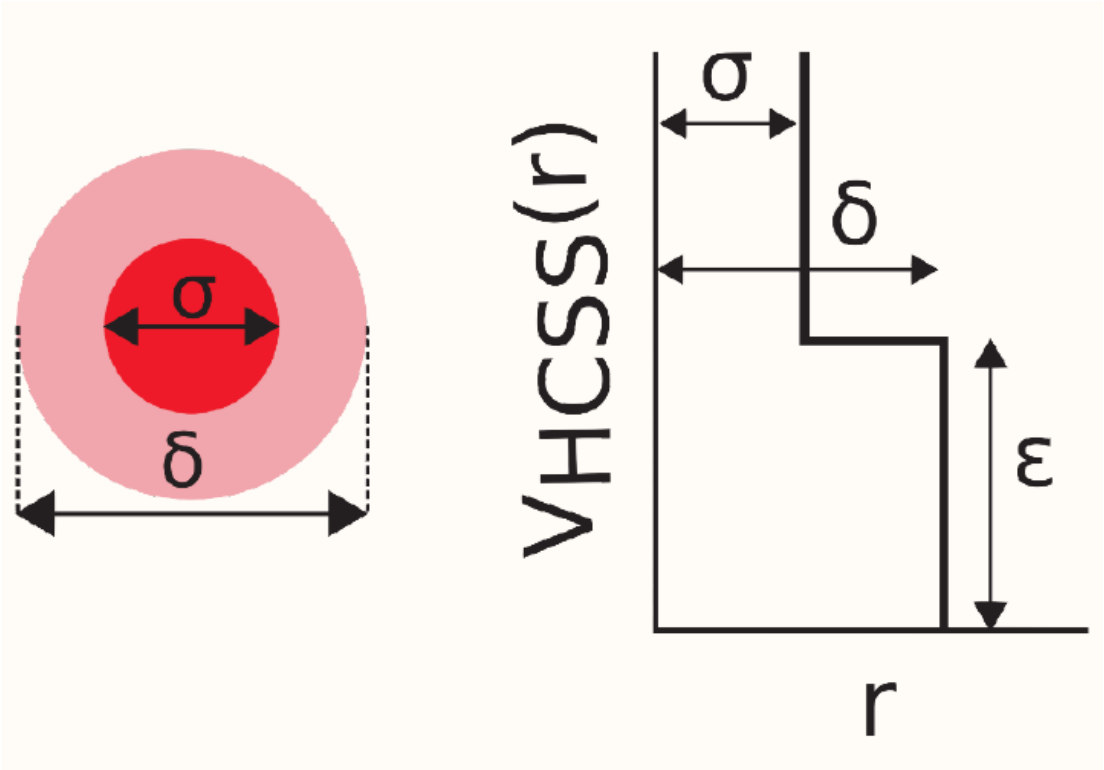


Model of Ligand-Stabilized Nanoparticles

Hard-Core Square-Shoulder potential:

Three parameters

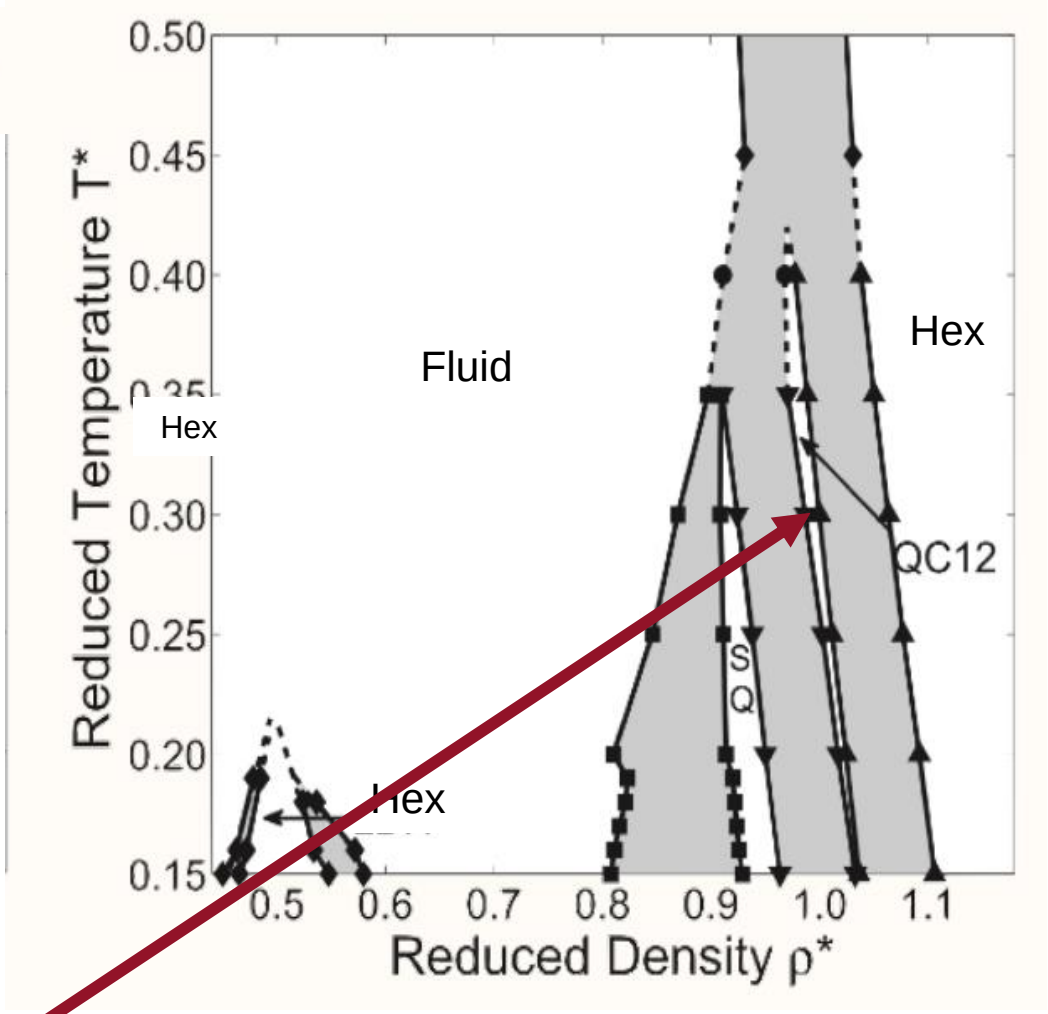
- The shoulder width δ/σ
- The temperature $k_B T/\epsilon$
- The pressure $\beta P \sigma^2$ (or density $\rho\sigma^2$)



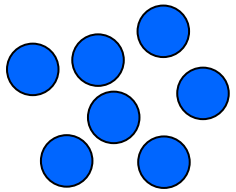
Two competing length scales: Quasicrystalline order

Phase Diagrams

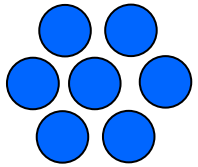
$\delta = 1.40\sigma$



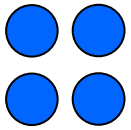
Fluid



Hexagonal (Hex)



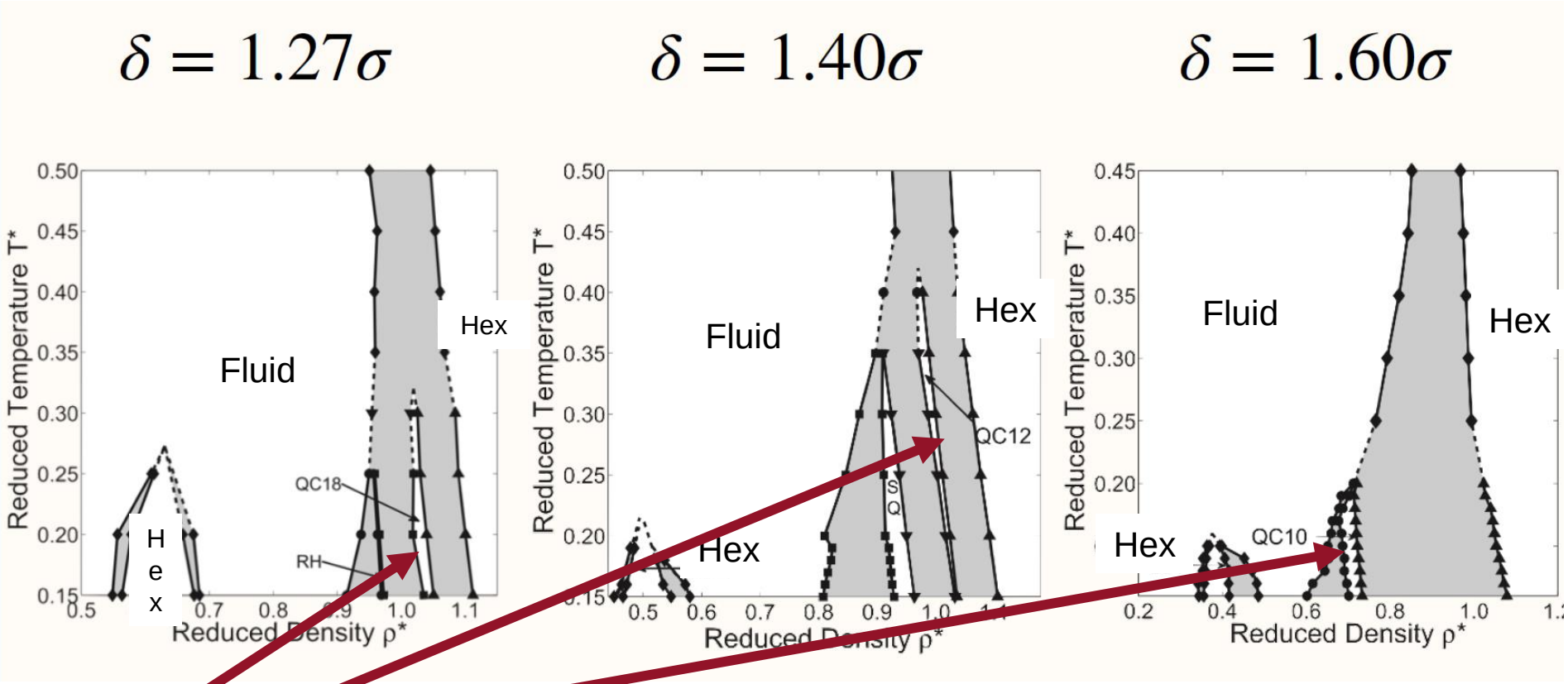
Square (Sq)



12-fold Quasicrystals in a tiny region in the phase diagram

H. Pattabhiraman, A.P. Gantapara, and M. Dijkstra, *The Journal of Chemical Physics* **143**, 164905 (2015); ibid **146**, 114901 (2017).

Phase Diagrams



18-, 12-, 10-fold Quasicrystals in a tiny region in the phase diagram

H. Pattabhiraman, A.P. Gantapara, and M. Dijkstra, *The Journal of Chemical Physics* **143**, 164905 (2015); ibid **146**, 114901 (2017).

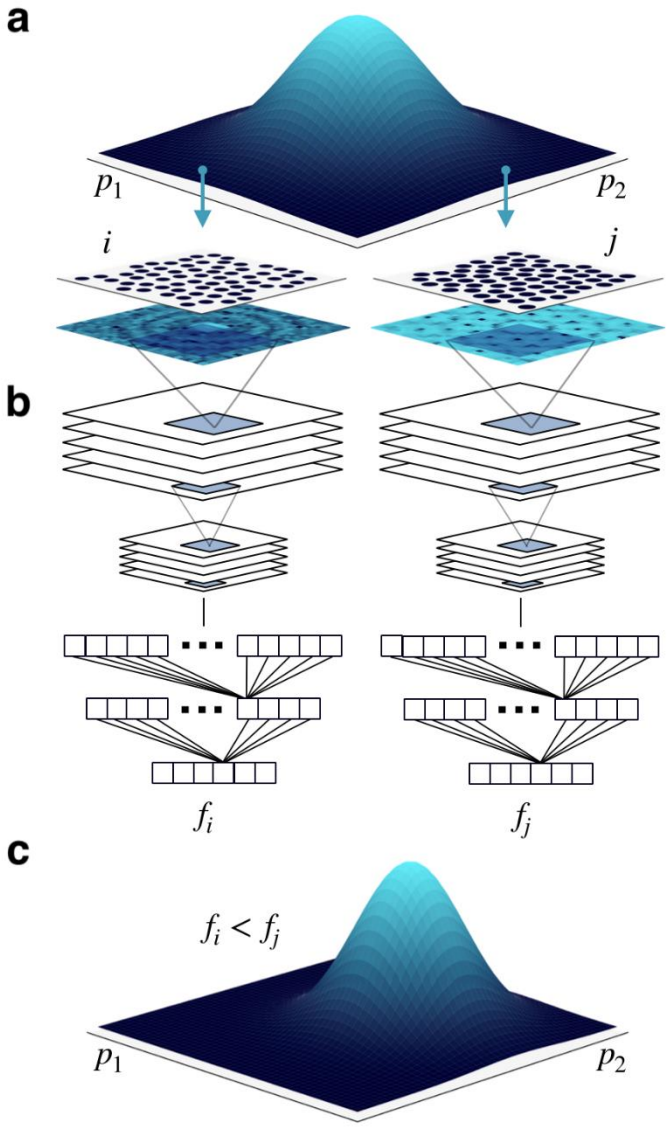
Inverse Design Method

Three ingredients:

- The Model with free design parameters that will be optimized
- Scheme to update the design parameters

  Covariance Matrix Adaptation Evolutionary Strategy

- Order parameter to distinguish the target phase from competing phases – fitness function



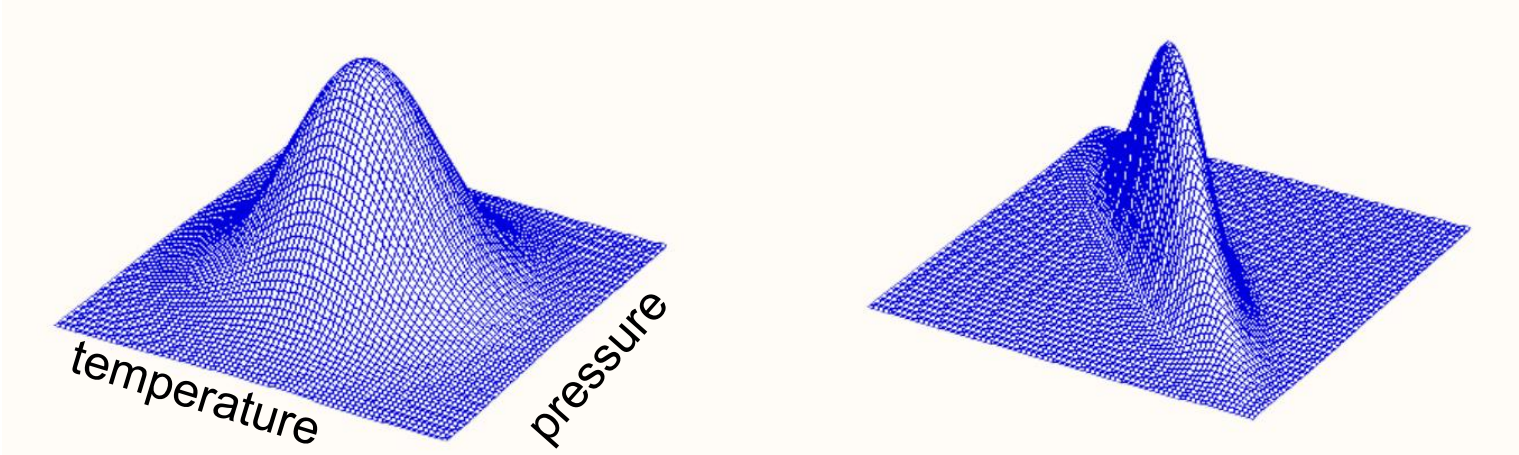
The Optimizer

- The CMA-ES (Covariance Matrix Adaption – Evolutionary Strategy) algorithm works by evolving the **mean vector** and the **covariance matrix** of a **multivariate gaussian distribution**;

- The **dimension** *D* of the distribution is set by the number of parameters we want to tune;

- At each generation we draw *n* **samples** from the multivariate gaussian;

- **By evaluating their performances** in terms of the fitness function, we **evolve the parameters** of the distribution and we repeat the process!



Inverse Design Method

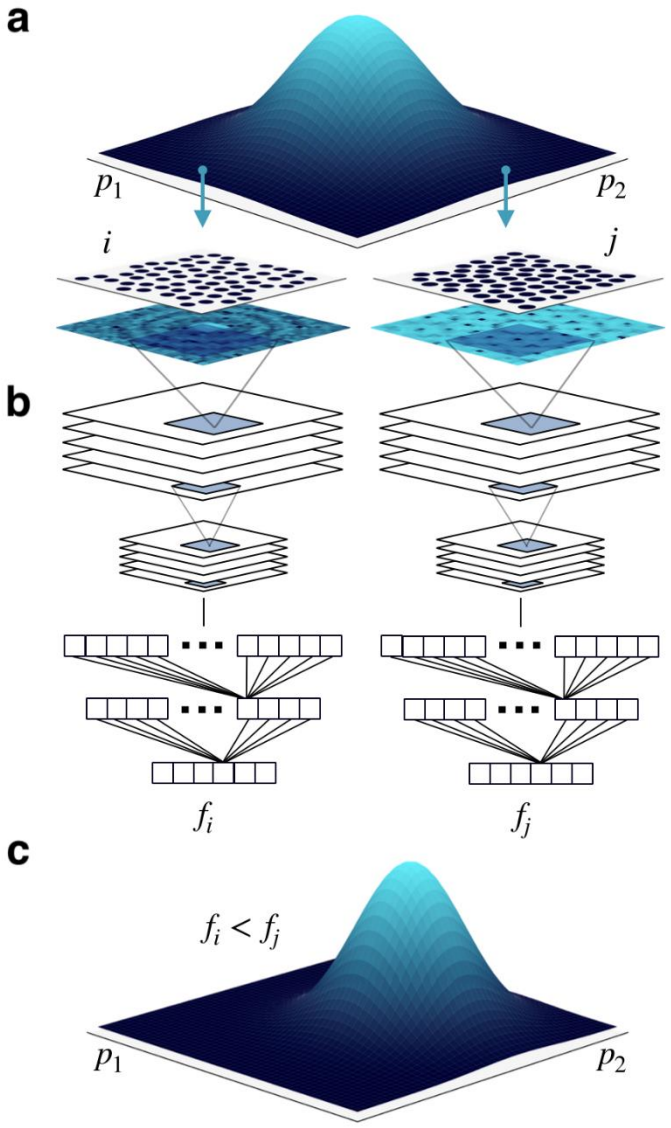
Three ingredients:

- The Model with free design parameters that will be optimized
- Scheme to update the design parameters

  Covariance Matrix Adaptation Evolutionary Strategy

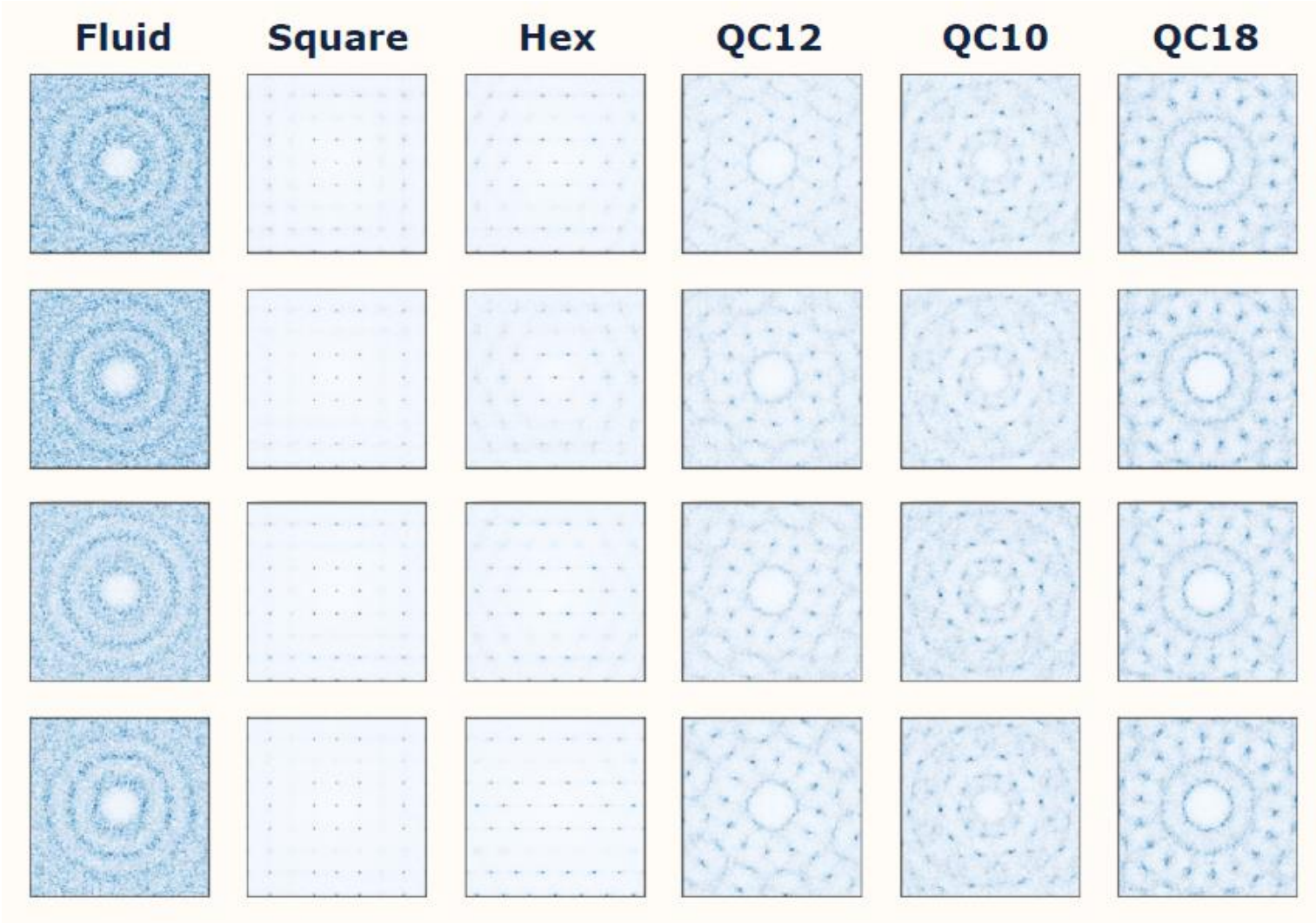
- Order parameter to distinguish the target phase from competing phases – fitness function

  Convolutional Neural Network that classifies the phases from their diffraction patterns



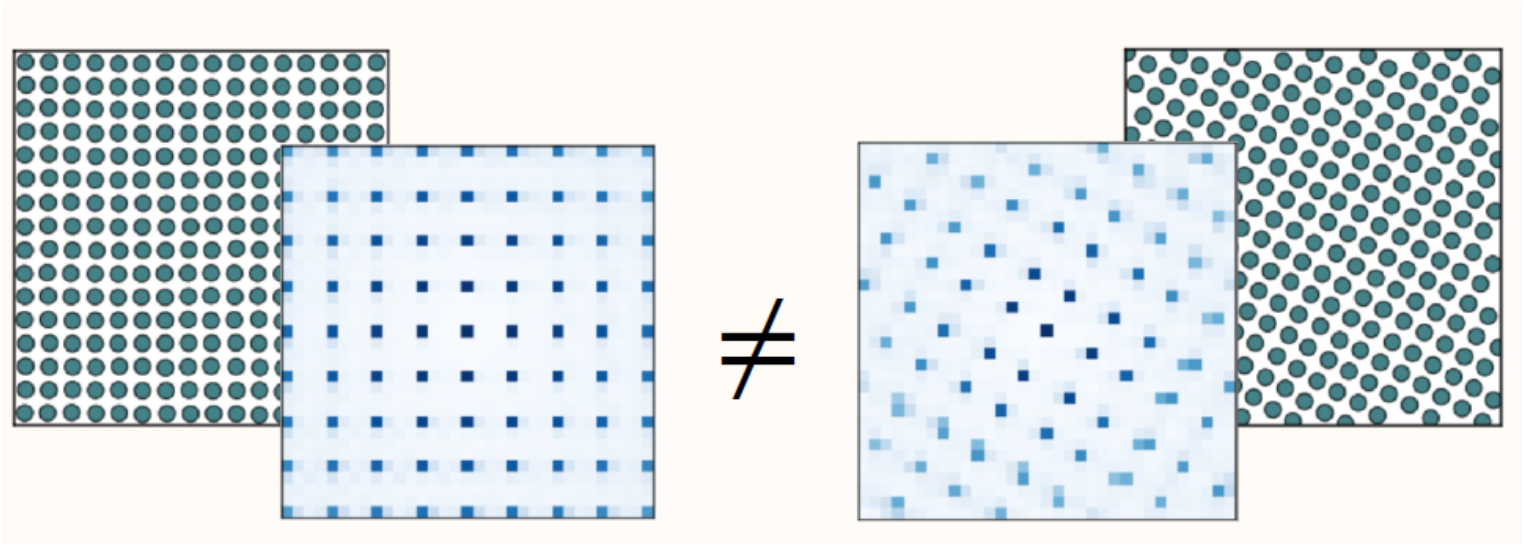
Building the training set

MC simulations of each phase at different state points and save 10000 independent configurations for each phase.



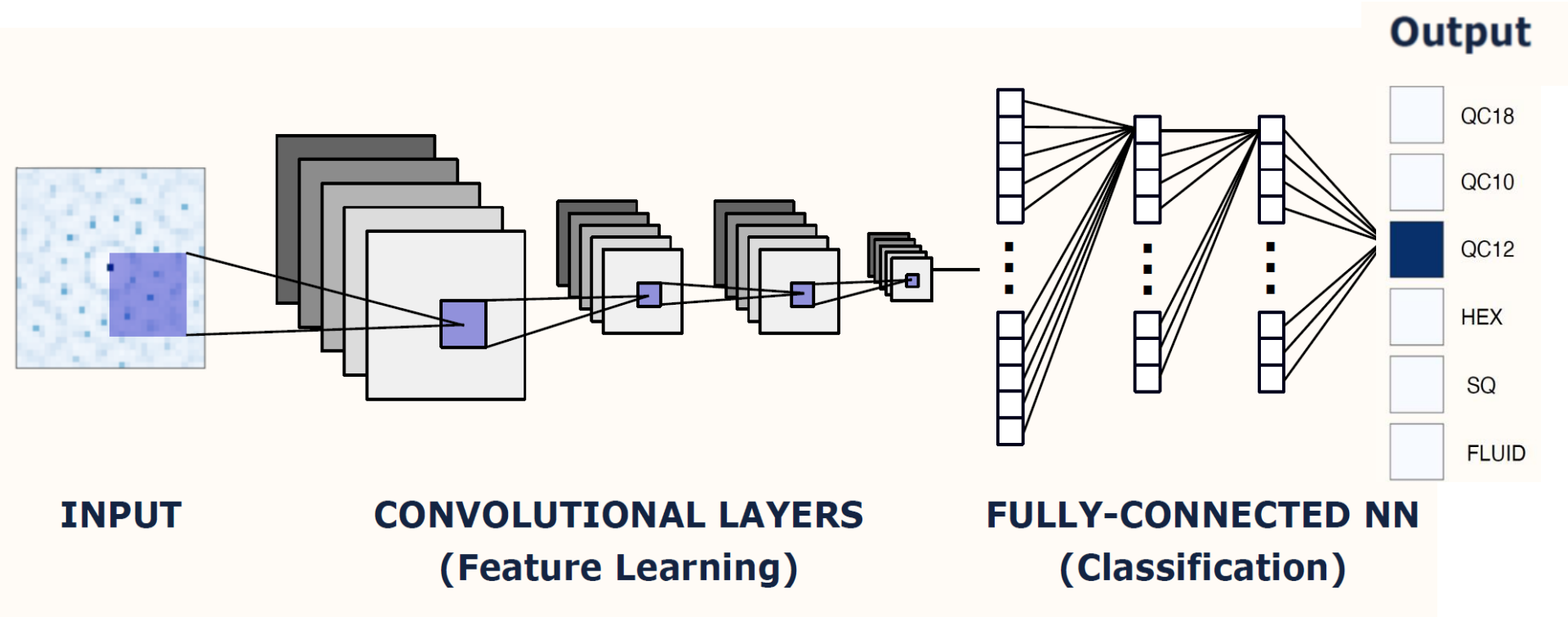
Data Augmentation

Diffraction patterns are **not rotationally invariant**



We **rotate each training example by a random angle** so that the network can learn to recognize each phase independent of its orientation

Our Convolutional Neural Network

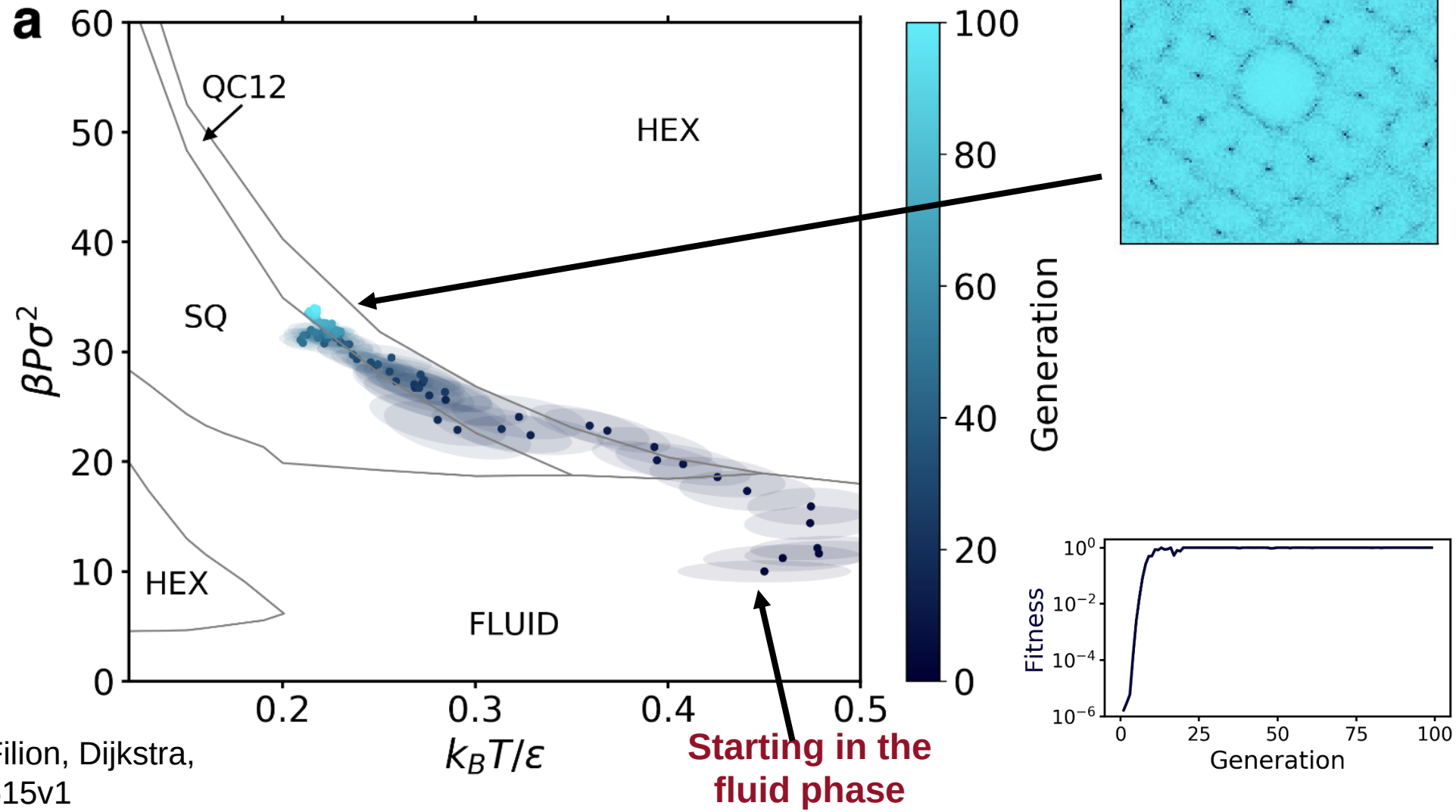


We use the probability of the target phase as the fitness function for the CMA-ES

The trained CNN can classify all phases with 100% accuracy!

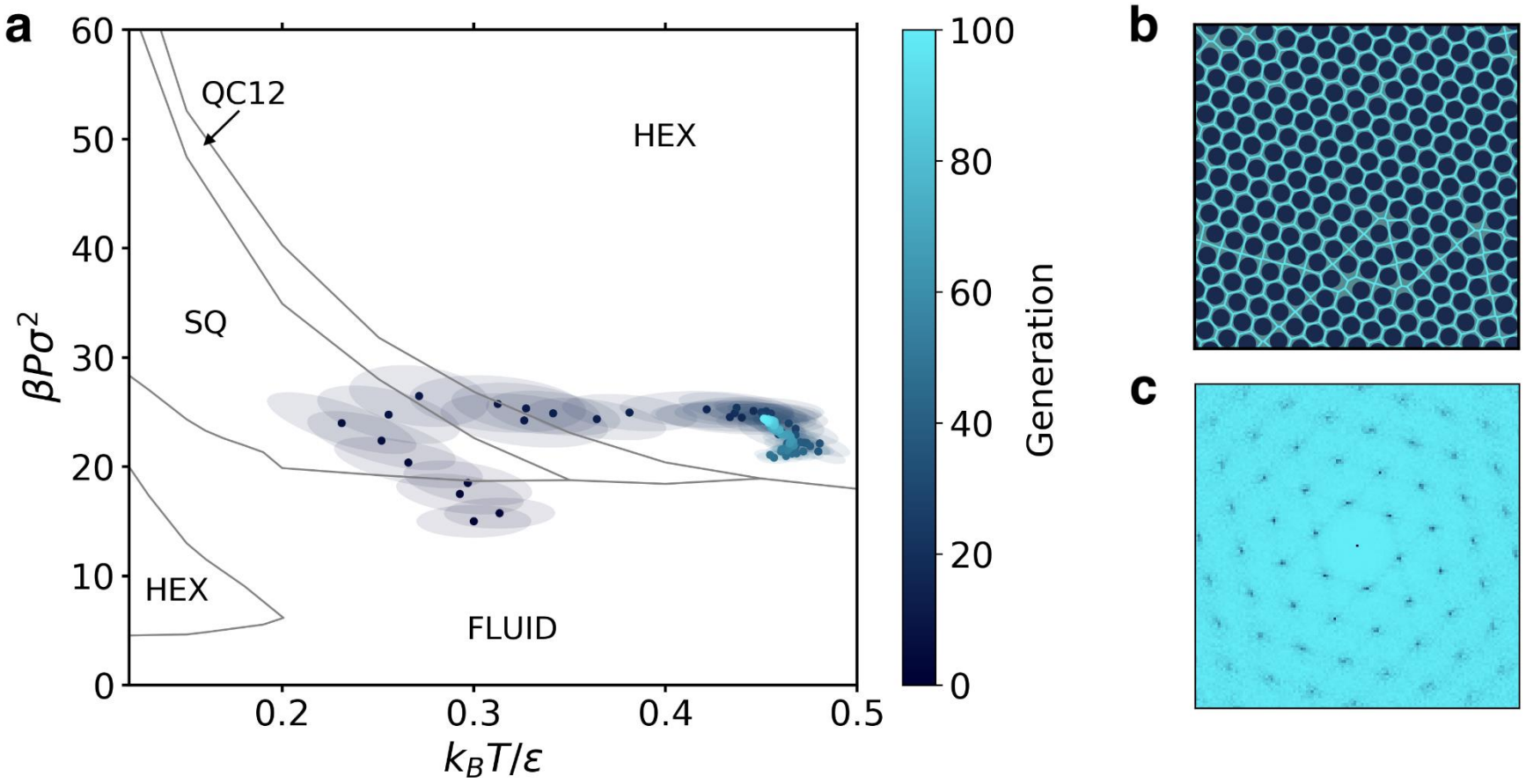
Inverse Design of the dodecagonal QC

Design Parameters: Pressure and Temperature
Shoulder width fixed (δ/σ = 1.4)



Coli, Boattini, Filion, Dijkstra, arXiv:2106.14615v1

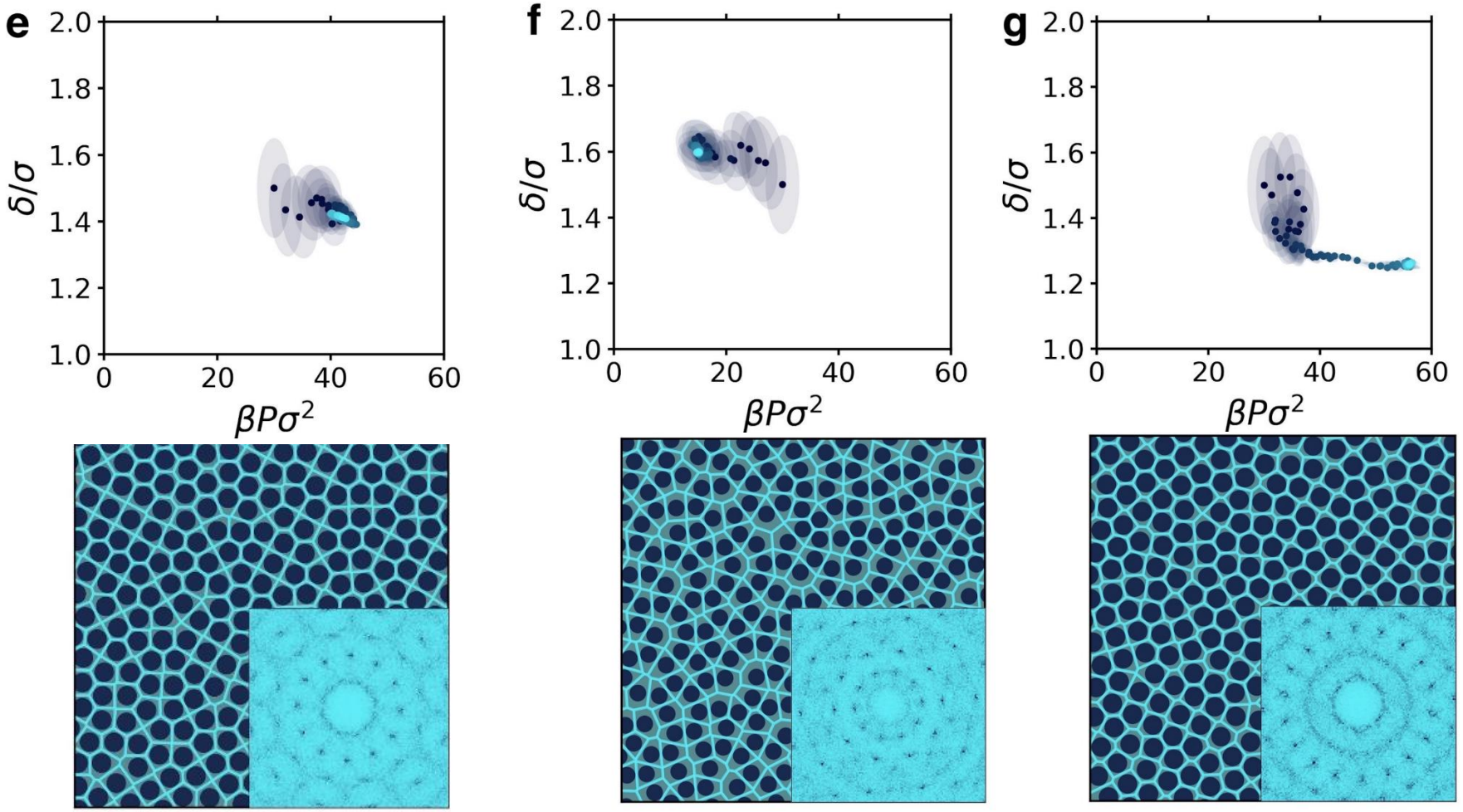
Inverse Design of the HEX crystal phase



Using the probability that it belongs to the HEX phase as determined by the CNN as fitness function in the CMA-ES

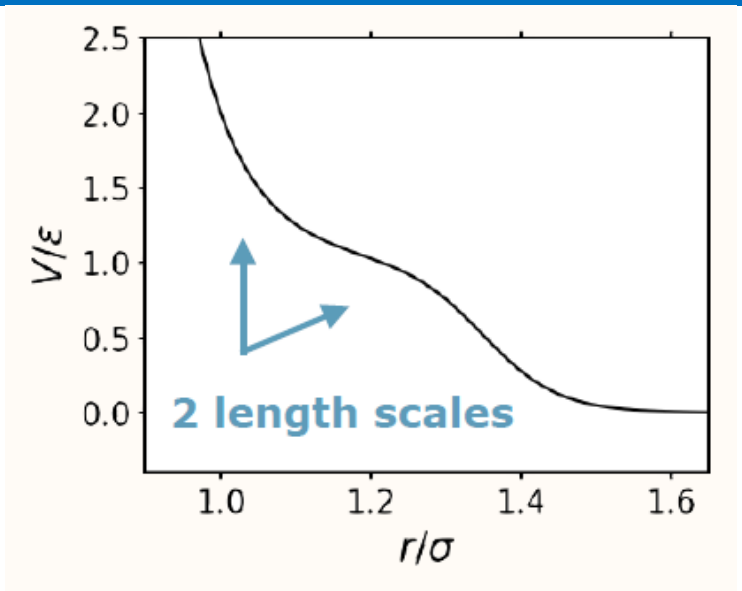
Inverse Design of 12-, 10-, and 18-fold QC

Design Parameters: Shoulder width δ/σ and pressure *P*



Fixed temperature *T*=0.17*

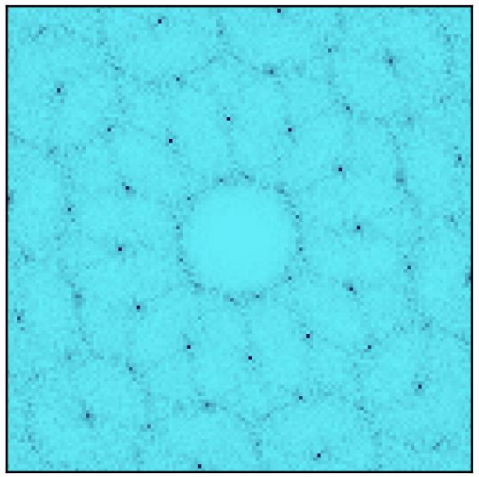
Transferability to other models

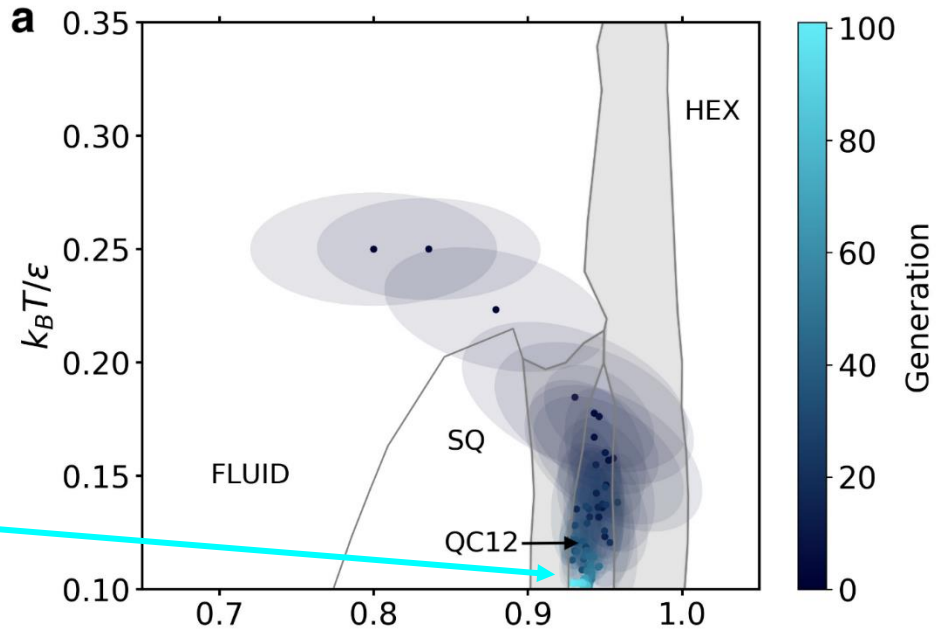


Softened-core-shoulder potential

$$V(r)/\epsilon = \left(\frac{\sigma}{r}\right)^{14} + \frac{1}{2}\left(1 - \tanh(k(r-\delta))\right)$$

Design Parameters: Temperature *T* and density ρ



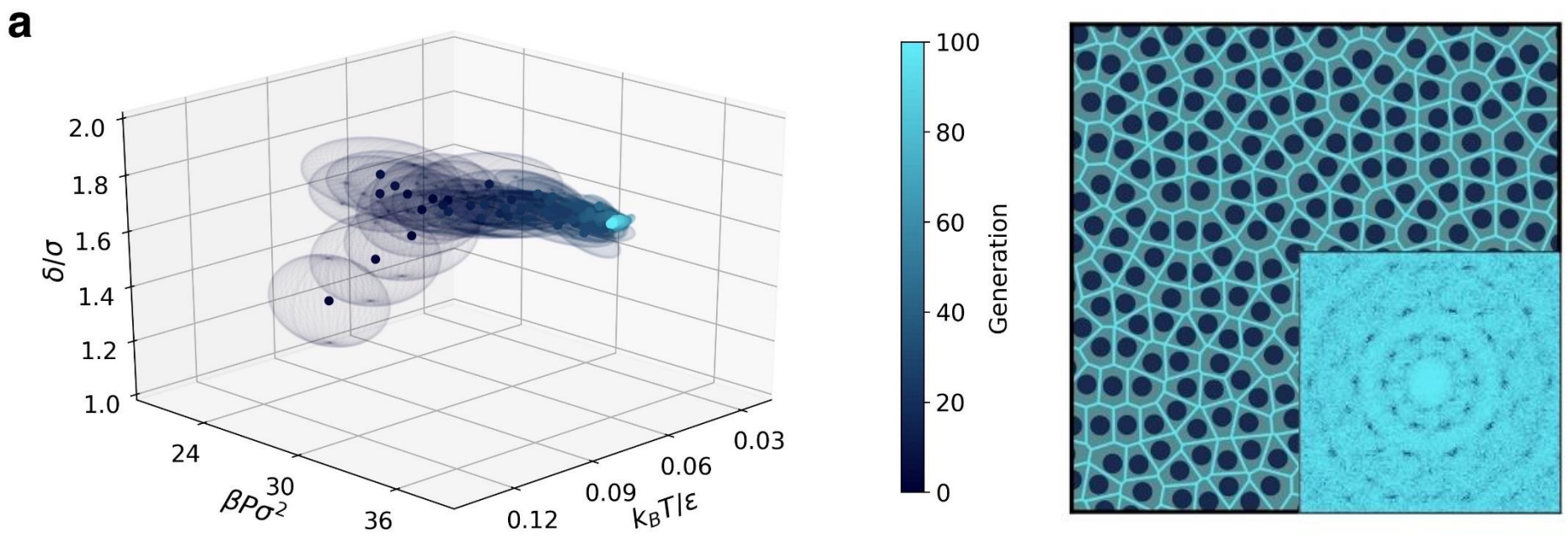


Fixed interaction parameters: κ and δ

P. Kryuchkov *et al.*, Soft Matter 14, 2152 (2018)

Discovery of a new 10-fold QC

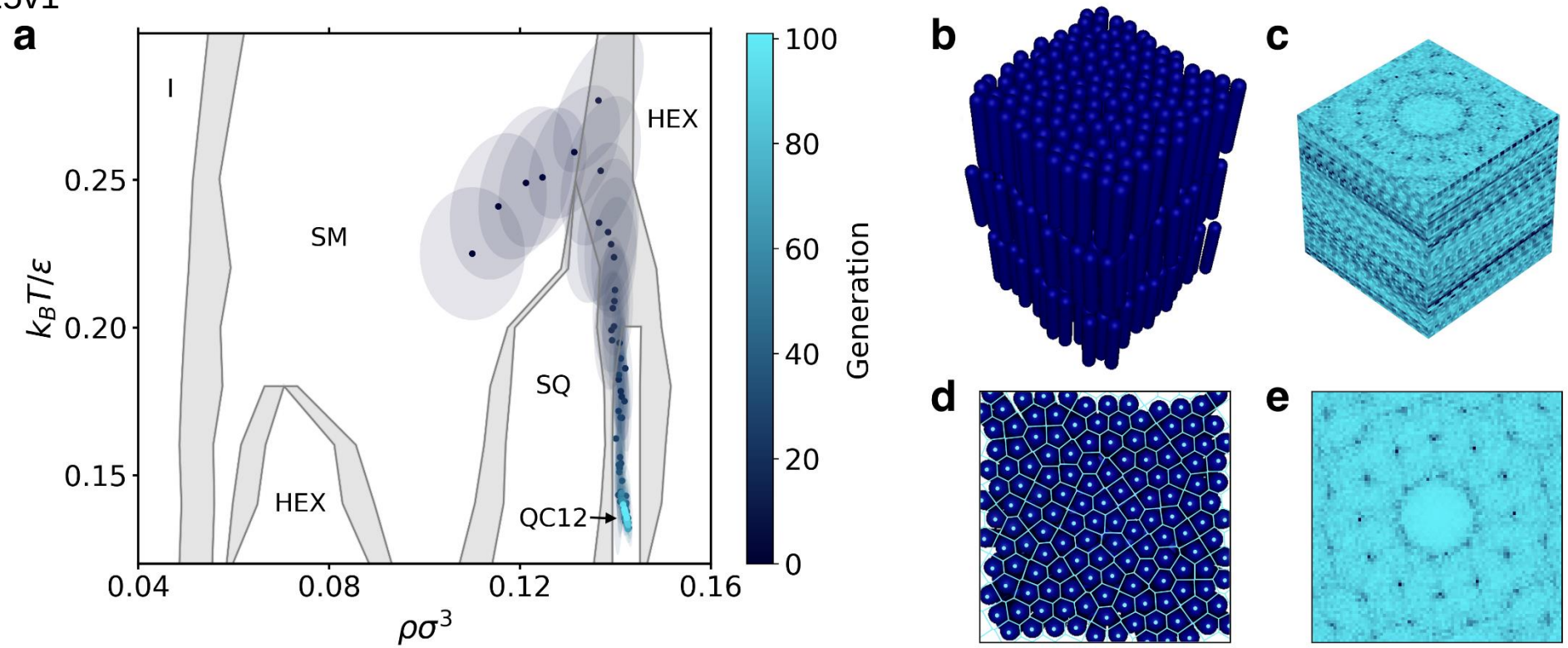
Design Parameters: Temperature *T*, and density ρ , Shoulder width δ



Fixed interaction parameters: κ

Inverse Design of QC in a 3D system

Coli, Boattini, Filion, Dijkstra,
arXiv:2106.14615v1

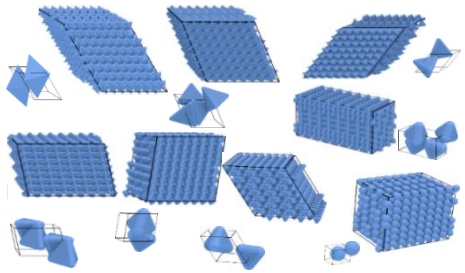


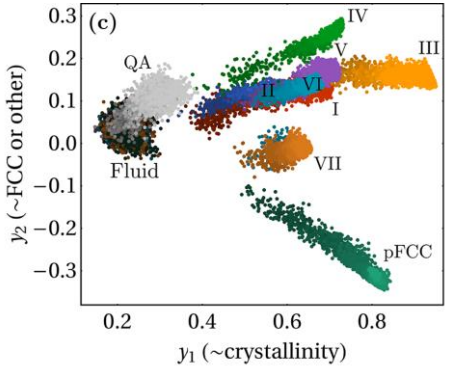
It also works in 3D with 3D diffraction patterns!!

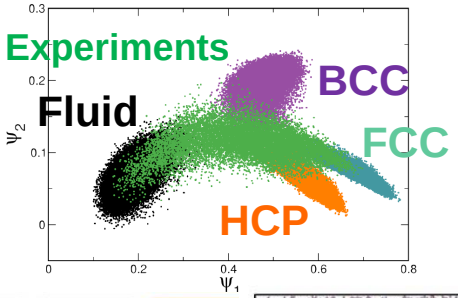
Phase diagram of soft-corona spherocylinders: G. Campos de Villalobos, M. Dijkstra, A. Patti, PRL 2021

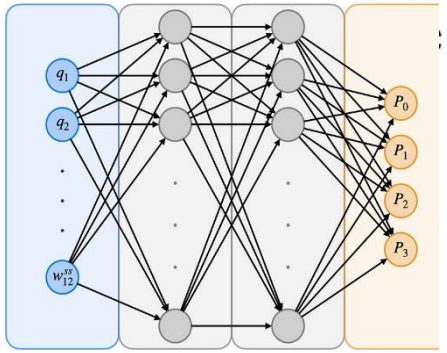
Conclusions

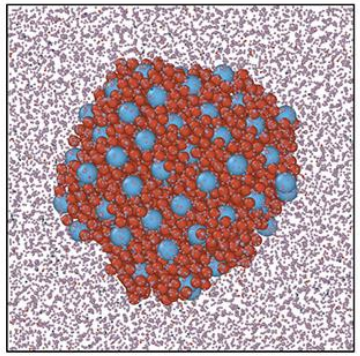
- Classification and Finding Order parameters



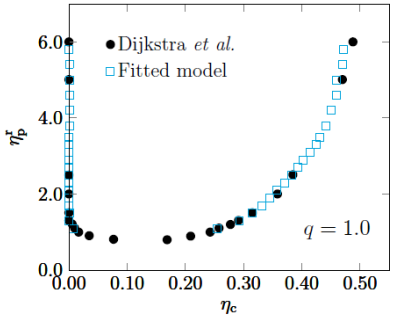


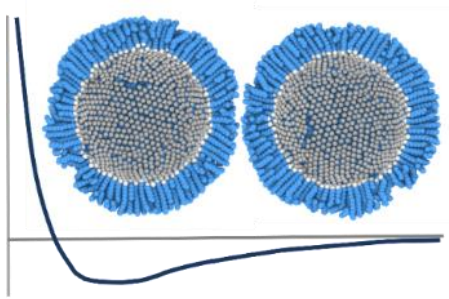




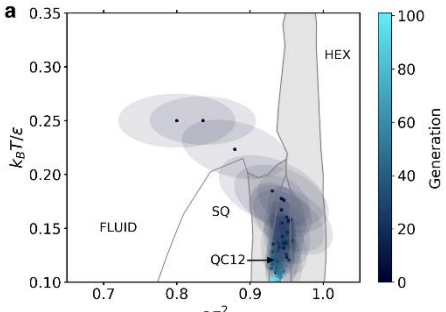


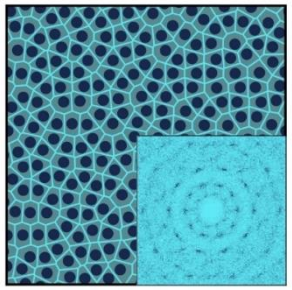
- Coarse-graining





- Inverse Design





Acknowledgements

Gabriele Coli

Robin van Damme

Emanuele Boattini

Gerardo Campos-Villalobos

Co-supervision

Laura Filion

René van Roij

Funding

erc Advanced Grant

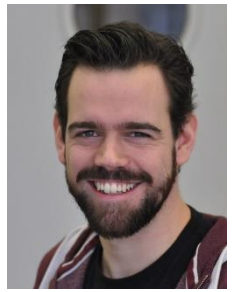
NWO

Experiments

Alfons van Blaaderen

Arnout Imhof

Patrick Baesjou



Ernest van der Wee

Anna Nikolaenkova