UNIVERSITY OF LIVERPOOL

LEVERHULME TRUST

The Leverhulme Research Centre for Functional Materials Design

Aspects of Uncertainty Quantification in Materials Discovery

Miloslav Torda

Knowledge bite

diamond

Hartree Centre
Science & Technology Facilities Council

Imperial College London

ICNS

King Abdullah University of Science and Technology

University of Southampton

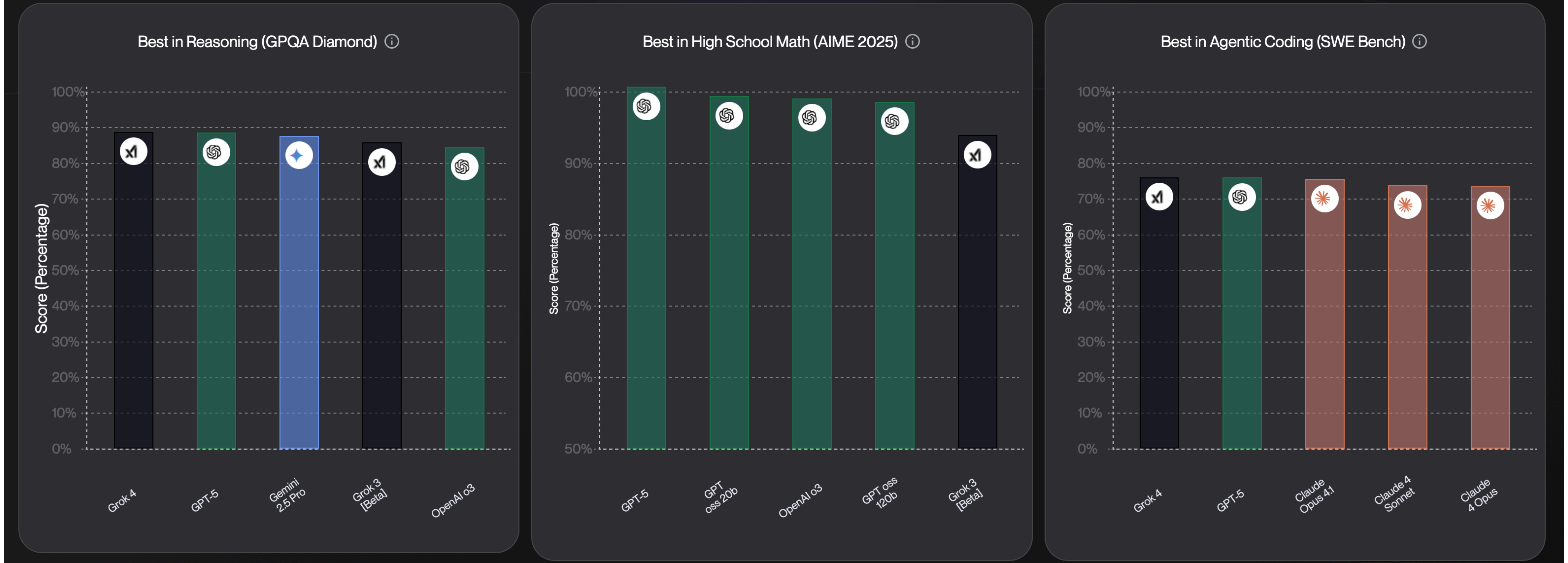
Are You Confident?

UPDATED 11 Sept 2025

LLM Leaderboard

This LLM leaderboard displays the latest public benchmark performance for SOTA model versions released after April 2024. The data comes from model providers as well as independently run evaluations by Vellum or the open-source community. We feature results from non-saturated benchmarks, excluding outdated benchmarks (e.g. MMLU). If you want to evaluate these models on your use-cases, try Vellum Evals.

Top models per tasks



Are You Confident?

write me a script visualizing an icosahedron in matlab

Thought for 39s >

Here's a compact, modern MATLAB script that builds an icosahedron from first principles (golden-ratio coordinates), then renders it nicely with lighting, vertex markers, and a wire overlay. It uses `patch` with Faces/Vertices as documented by MathWorks. MathWorks

matlab   Copy

```matlab
% ICOSAHEDRON_VIS.m
% Visualize a regular icosahedron with shading, edges, and vertex markers.
% Cleve-style: short, clear, and vectorized.

% Geometry (golden rectangles)
phi = (1+sqrt(5))/2;
V = [
     0,  1,  phi;
     0, -1,  phi;
     0,  1, -phi;
```

Are You Confident?

write me a script visualizing an icosahedron in matlab

Thought for 39s >

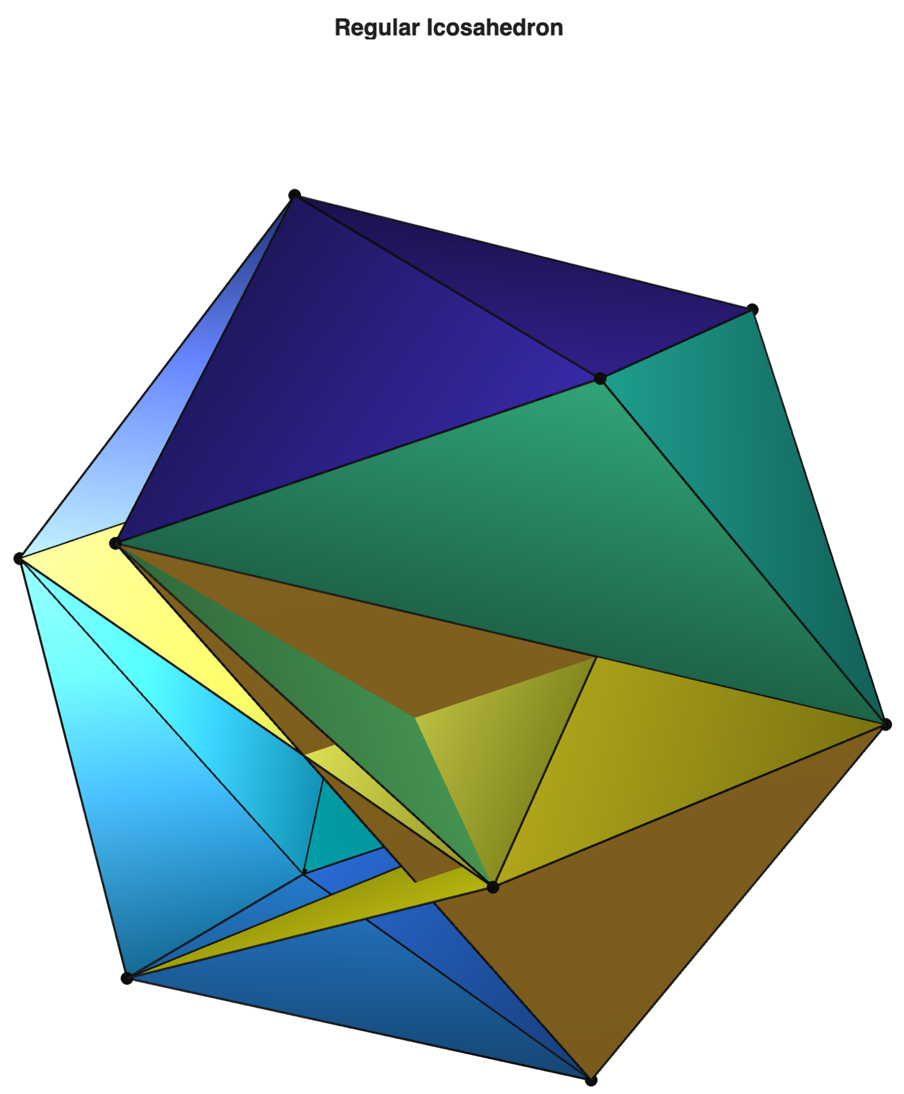
Here's a compact, modern MATLAB script that builds an icosahedron from first principles (golden-ratio coordinates), then renders it nicely with lighting, vertex markers, and a wire overlay. It uses `patch` with Faces/Vertices as documented by MathWorks. MathWorks

matlab

Copy

```
% ICOSAHEDRON_VIS.m
% Visualize a regular icosahedron with shading, edges, and vertex markers.
% Cleve-style: short, clear, and vectorized.

% Geometry (golden rectangles)
phi = (1+sqrt(5))/2;
V = [
     0,  1,  phi;
     0, -1,  phi;
     0,  1, -phi;
```



Are You Confident?

write me a script visualizing an icosahedron in matlab

Thought for 39s >

Here's a compact, modern MATLAB script that builds an icosahedron from first principles (golden-ratio coordinates), then renders it nicely with lighting, vertex markers, and a wire overlay. It uses `patch` with Faces/Vertices as documented by MathWorks. MathWorks

matlab Copy

```
% ICOSAHEDRON_VIS.m
% Visualize a regular icosahedron with shading, edges, and vertex markers.
% Cleve-style: short, clear, and vectorized.

% Geometry (golden rectangles)
phi = (1+sqrt(5))/2;
V = [
     0,  1,  phi;
     0, -1,  phi;
     0,  1, -phi;
```

Sources

nope you got something wrong

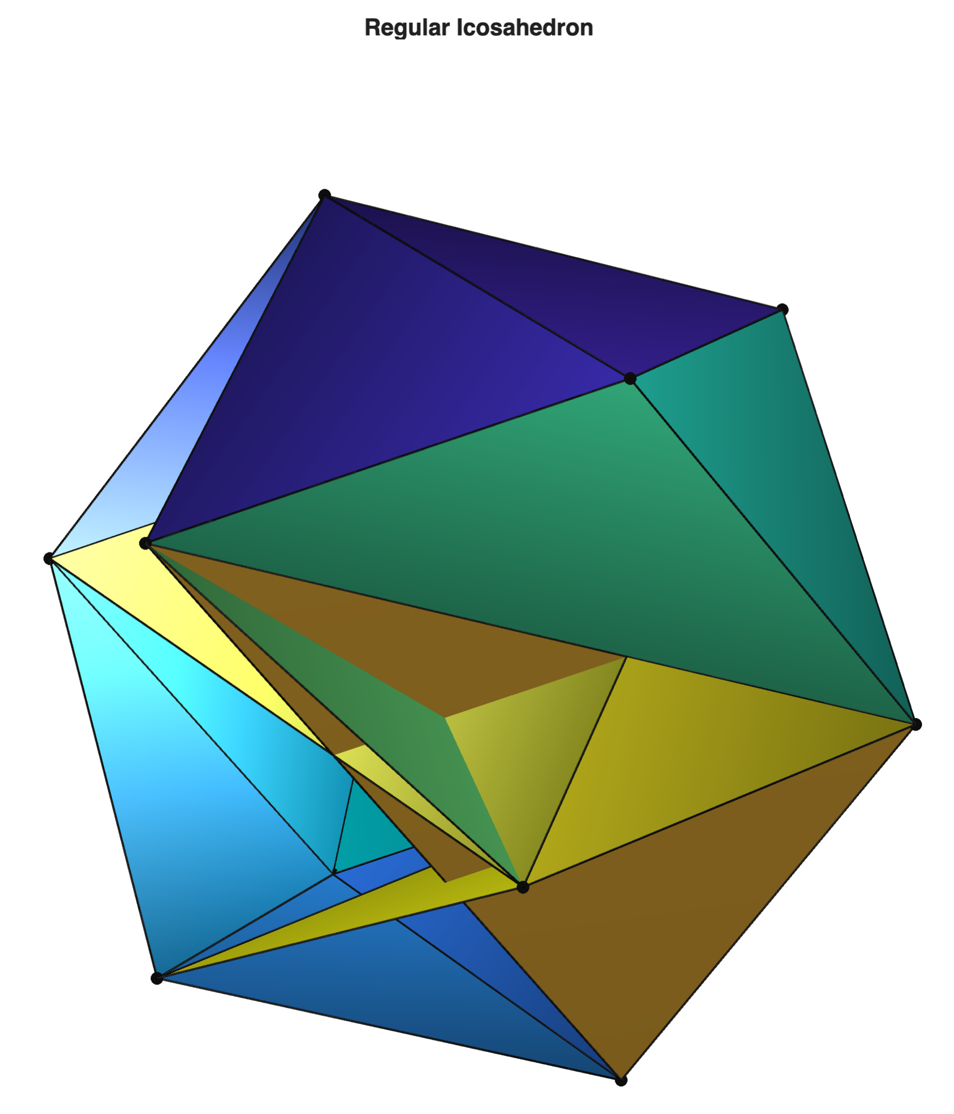
Thought for 33s >

You're right—my face list was off. Here's a rock-solid version that **derives the faces automatically** from the vertex set using `convhulln`, so the connectivity is guaranteed correct for a regular icosahedron.

matlab Copy

```
% ICOSAHEDRON_VIS.m
% Visualize a regular icosahedron (faces from convex hull).
% Cleve-style: compact, clear, vectorized.

% Geometry: golden ratio vertices
phi = (1+sqrt(5))/2;
V = [...
     0   1   phi;
     0  -1   phi;
     0   1  -phi;
```



Are You Confident?

write me a script visualizing an icosahedron in matlab

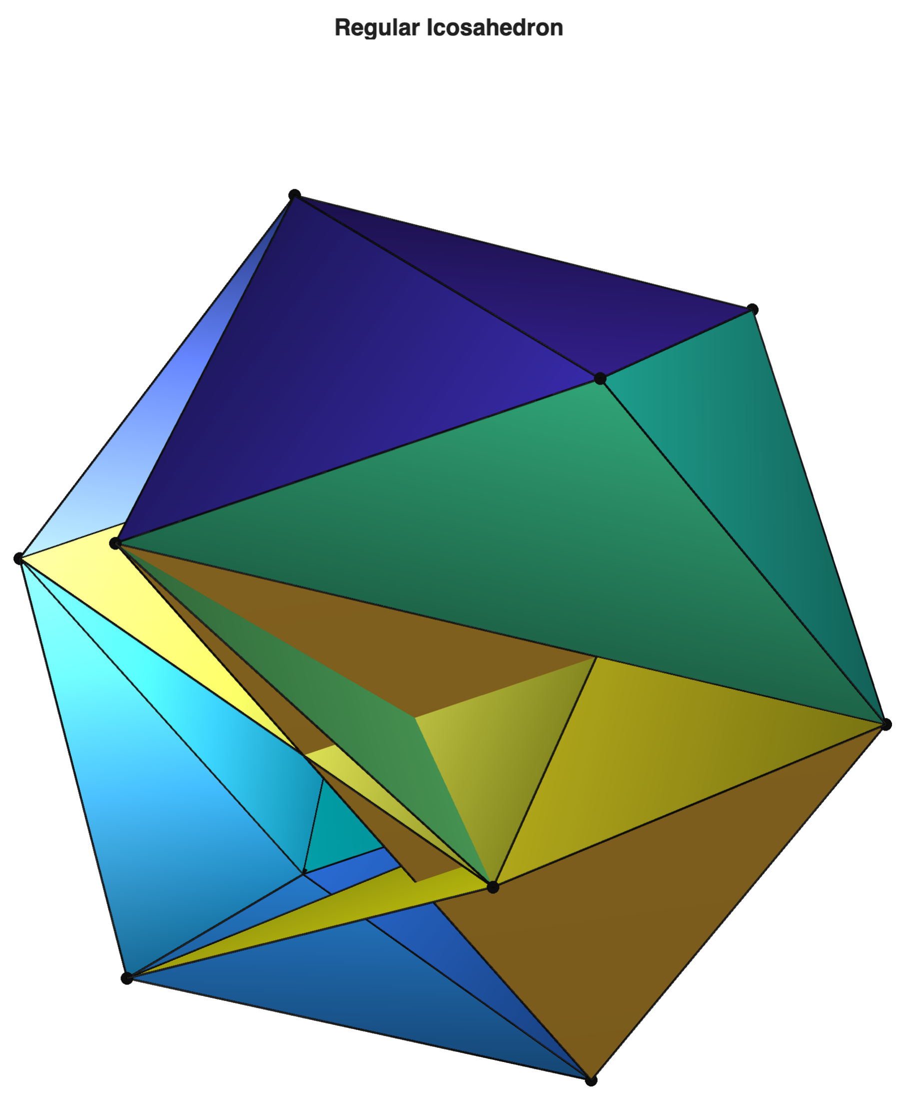
Thought for 39s >

Here's a compact, modern MATLAB script that builds an icosahedron from first principles (golden-ratio coordinates), then renders it nicely with lighting, vertex markers, and a wire overlay. It uses `patch` with Faces/Vertices as documented by MathWorks. MathWorks

matlab Copy

```matlab
% ICOSAHEDRON_VIS.m
% Visualize a regular icosahedron with shading, edges, and vertex markers.
% Cleve-style: short, clear, and vectorized.

% Geometry (golden rectangles)
phi = (1+sqrt(5))/2;
V = [
      0,  1,  phi;
      0, -1,  phi;
      0,  1, -phi;
```



Sources

nope you got something wrong

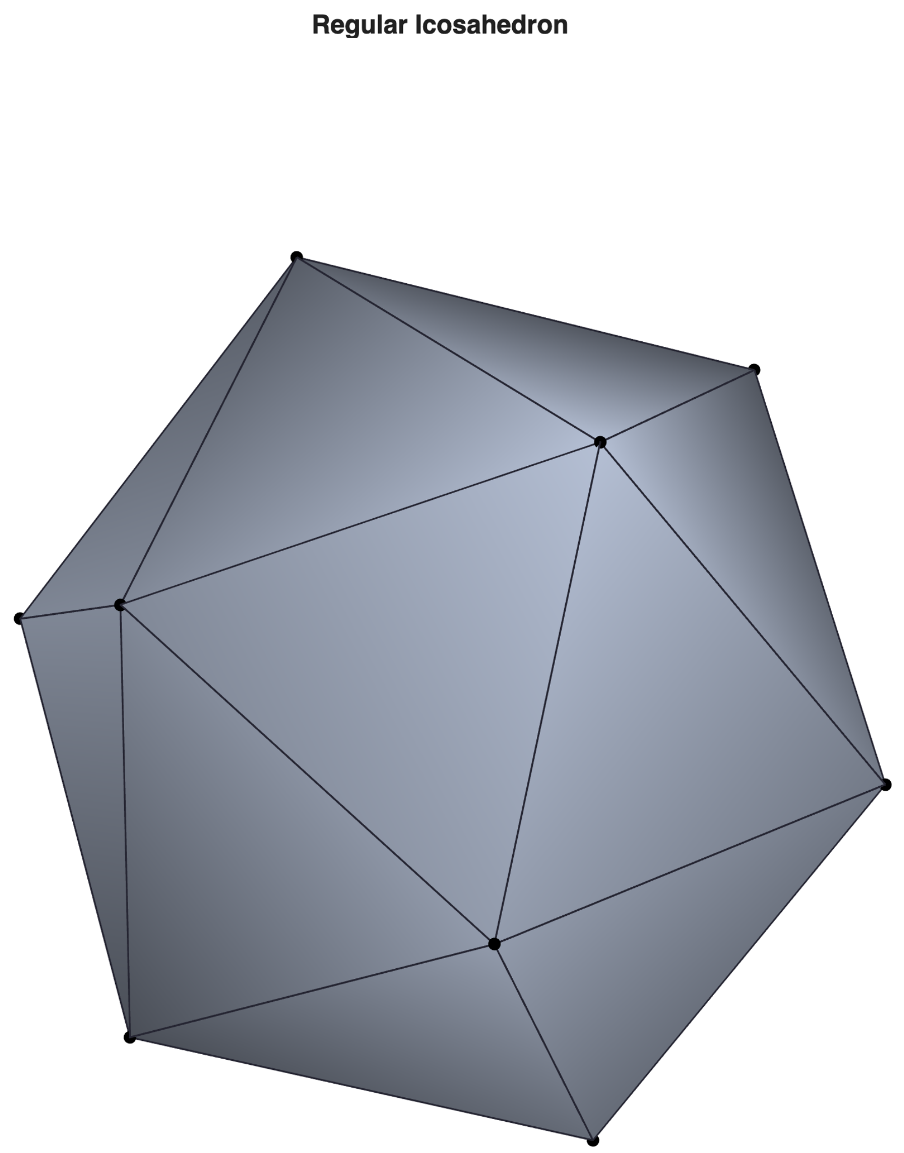
Thought for 33s >

You're right—my face list was off. Here's a rock-solid version that **derives the faces automatically** from the vertex set using `convhulln`, so the connectivity is guaranteed correct for a regular icosahedron.

matlab Copy

```matlab
% ICOSAHEDRON_VIS.m
% Visualize a regular icosahedron (faces from convex hull).
% Cleve-style: compact, clear, vectorized.

% Geometry: golden ratio vertices
phi = (1+sqrt(5))/2;
V = [...
      0   1   phi;
      0  -1   phi;
      0   1  -phi;
```

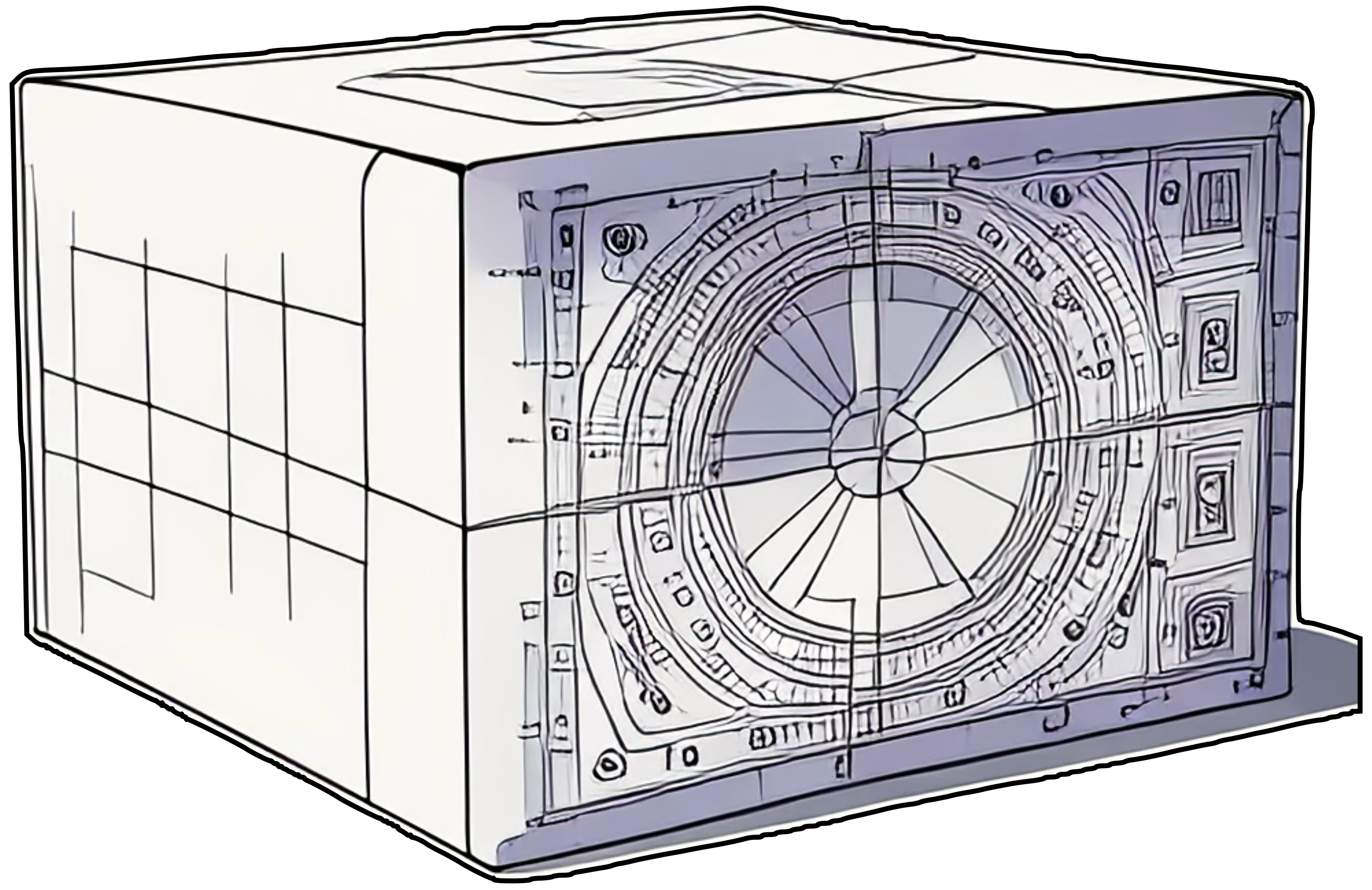


Epistemic or Systematic Uncertainties

- Model errors, model discrepancies, or model inadequacies

Epistemic or Systematic Uncertainties

- Model errors, model discrepancies, or model inadequacies
- The famous quote of Georg E.P. Box

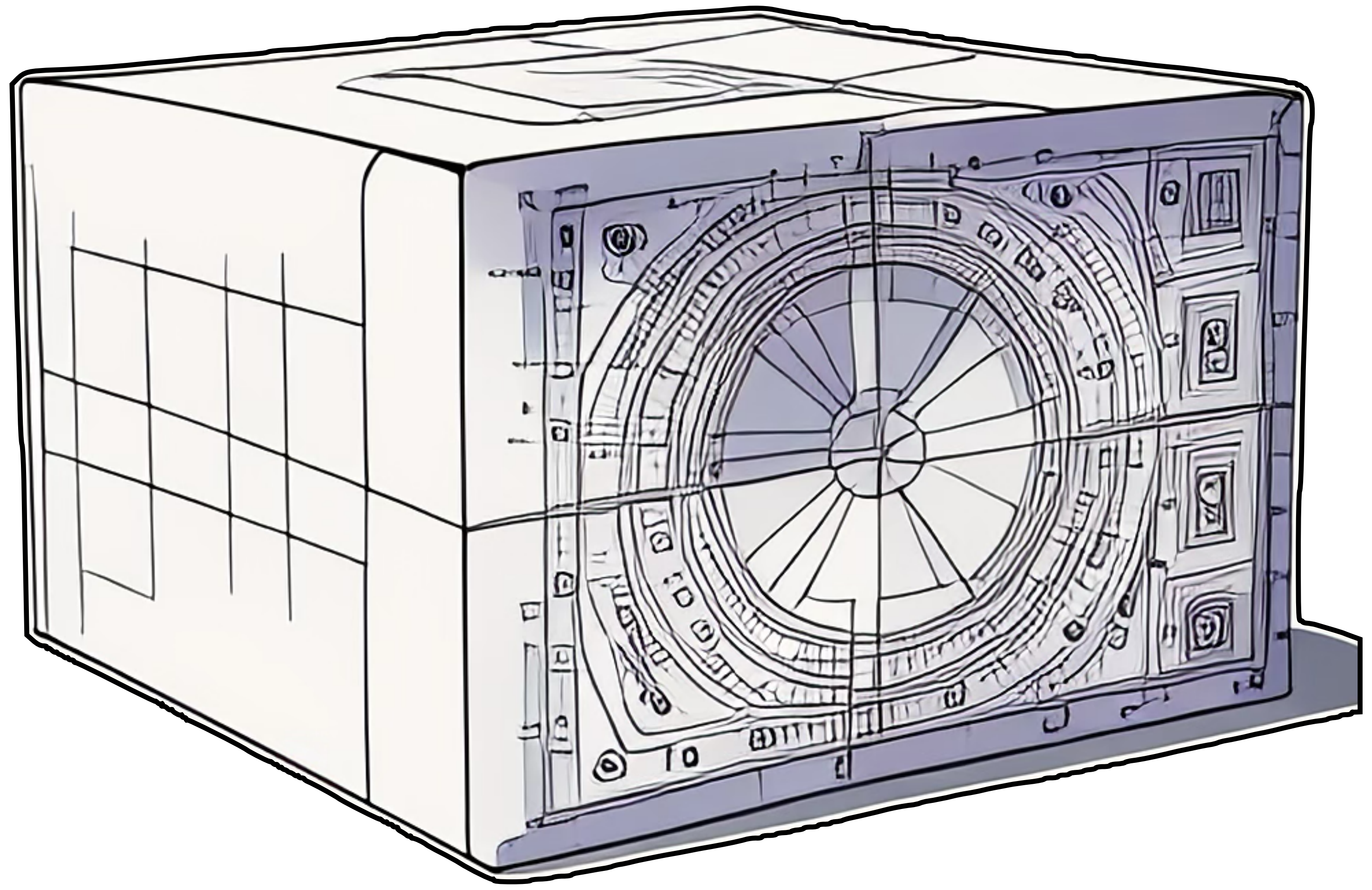


“Essentially, all models are wrong, but some are useful,“

G.E.P. Box and N.R. Draper, Empirical Model-Building and Response Surfaces, John Wiley and Sons, New York, 1987.

Epistemic or Systematic Uncertainties

- Model errors, model discrepancies, or model inadequacies
- The famous quote of Georg E.P. Box
    - **Invented and developed the idea of robustness**

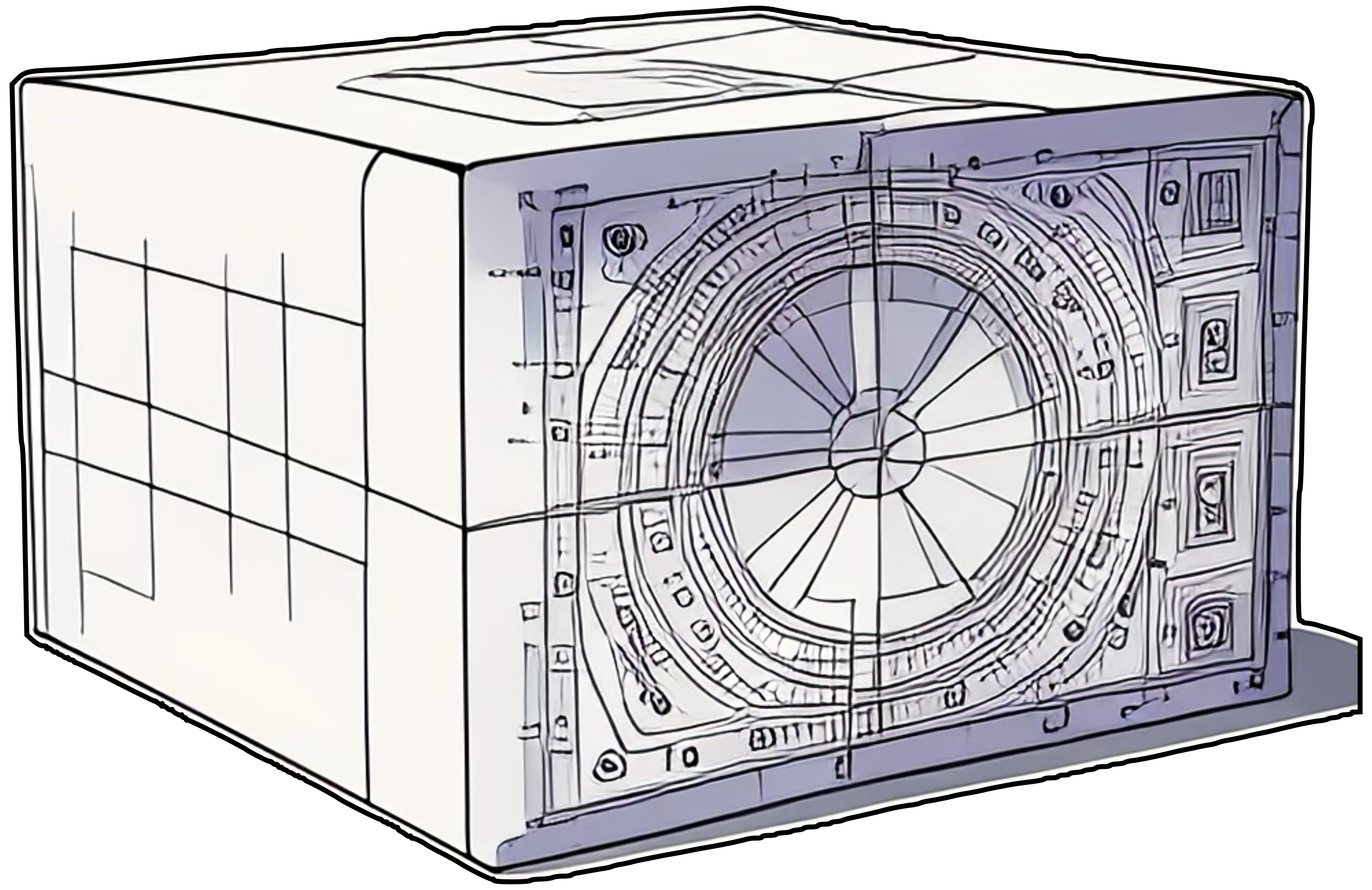


"Essentially, all models are wrong, but some are useful,"

G.E.P. Box and N.R. Draper, Empirical Model-Building and Response Surfaces, John Wiley and Sons, New York, 1987.

Epistemic or Systematic Uncertainties

- Model errors, model discrepancies, or model inadequacies
- The famous quote of Georg E.P. Box
  - **Invented and developed the idea of robustness**
  - **Had a key role in the integration of the classical and Bayesian schools of thought in Statistical Inference.**

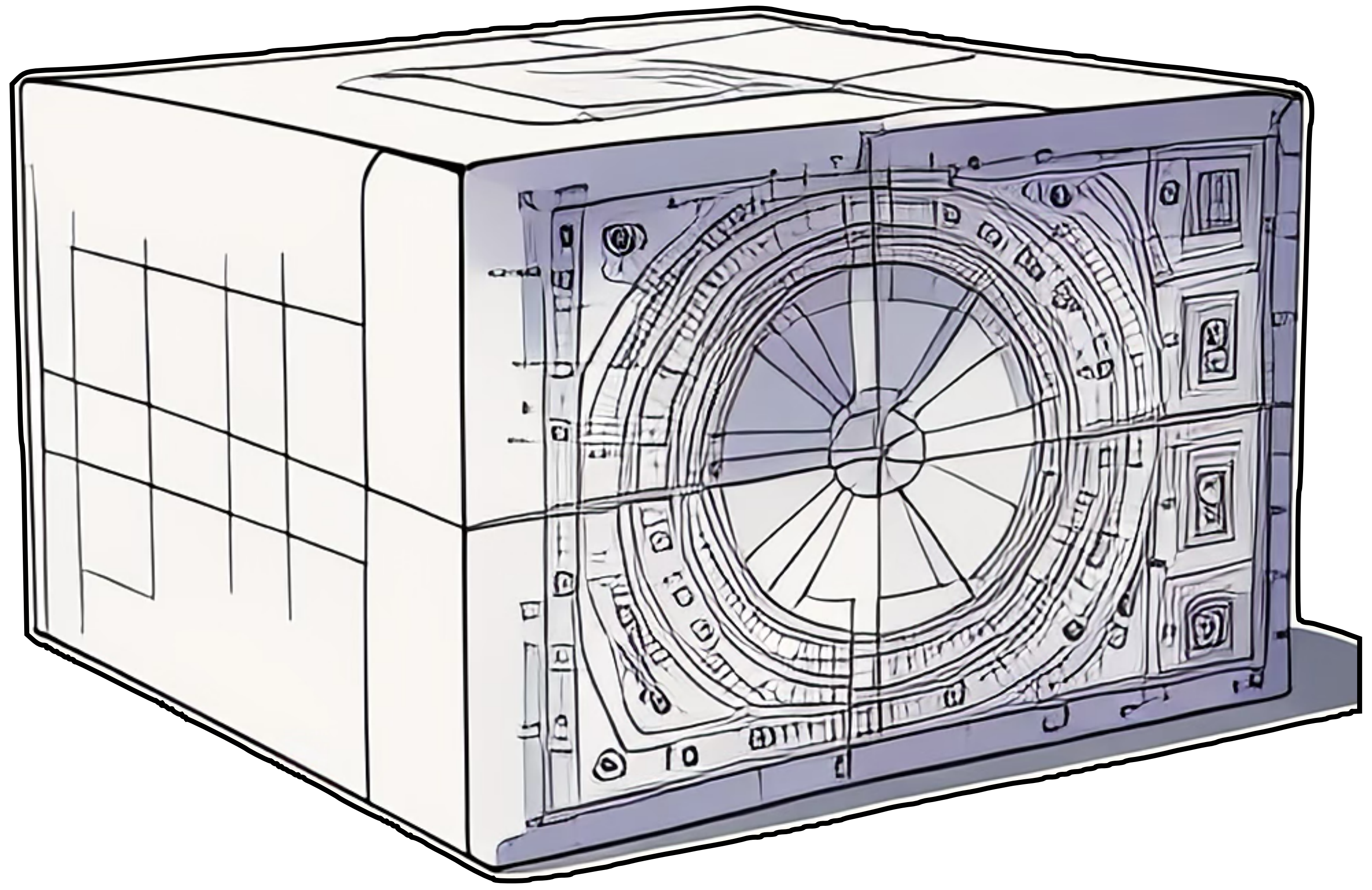


“Essentially, all models are wrong, but some are useful,“

G.E.P. Box and N.R. Draper, Empirical Model-Building and Response Surfaces, John Wiley and Sons, New York, 1987.

Epistemic or Systematic Uncertainties

- Model errors, model discrepancies, or model inadequacies
- The famous quote of Georg E.P. Box
    - **Invented and developed the idea of robustness**
    - **Had a key role in the integration of the classical and Bayesian schools of thought in Statistical Inference.**
- Sources of Systematic Uncertainties:

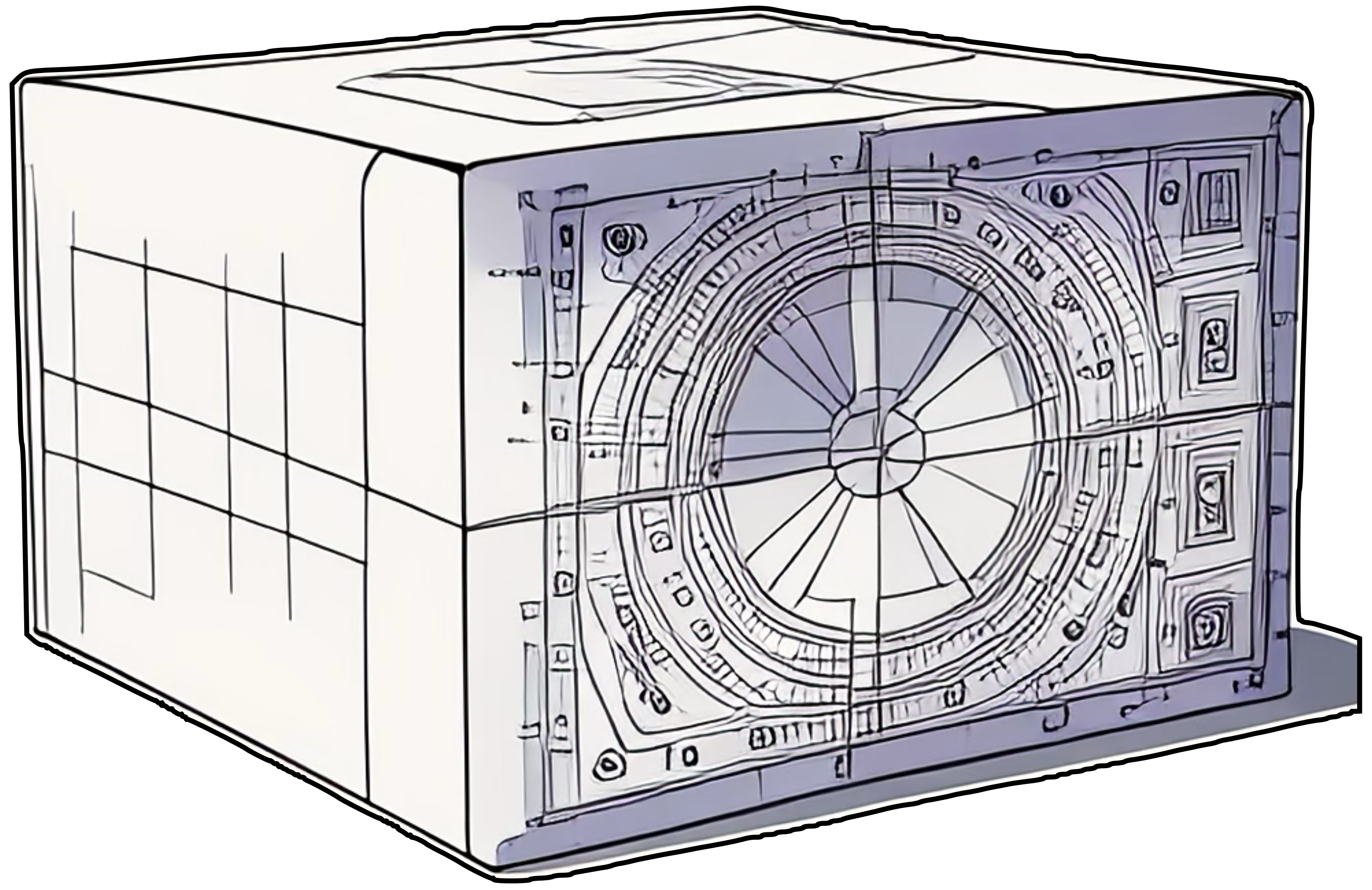


"Essentially, all models are wrong, but some are useful,"

G.E.P. Box and N.R. Draper, Empirical Model-Building and Response Surfaces, John Wiley and Sons, New York, 1987.

Epistemic or Systematic Uncertainties

- Model errors, model discrepancies, or model inadequacies
- The famous quote of Georg E.P. Box
  - **Invented and developed the idea of robustness**
  - **Had a key role in the integration of the classical and Bayesian schools of thought in Statistical Inference.**
- Sources of Systematic Uncertainties:
  - **Simplifying model assumptions**

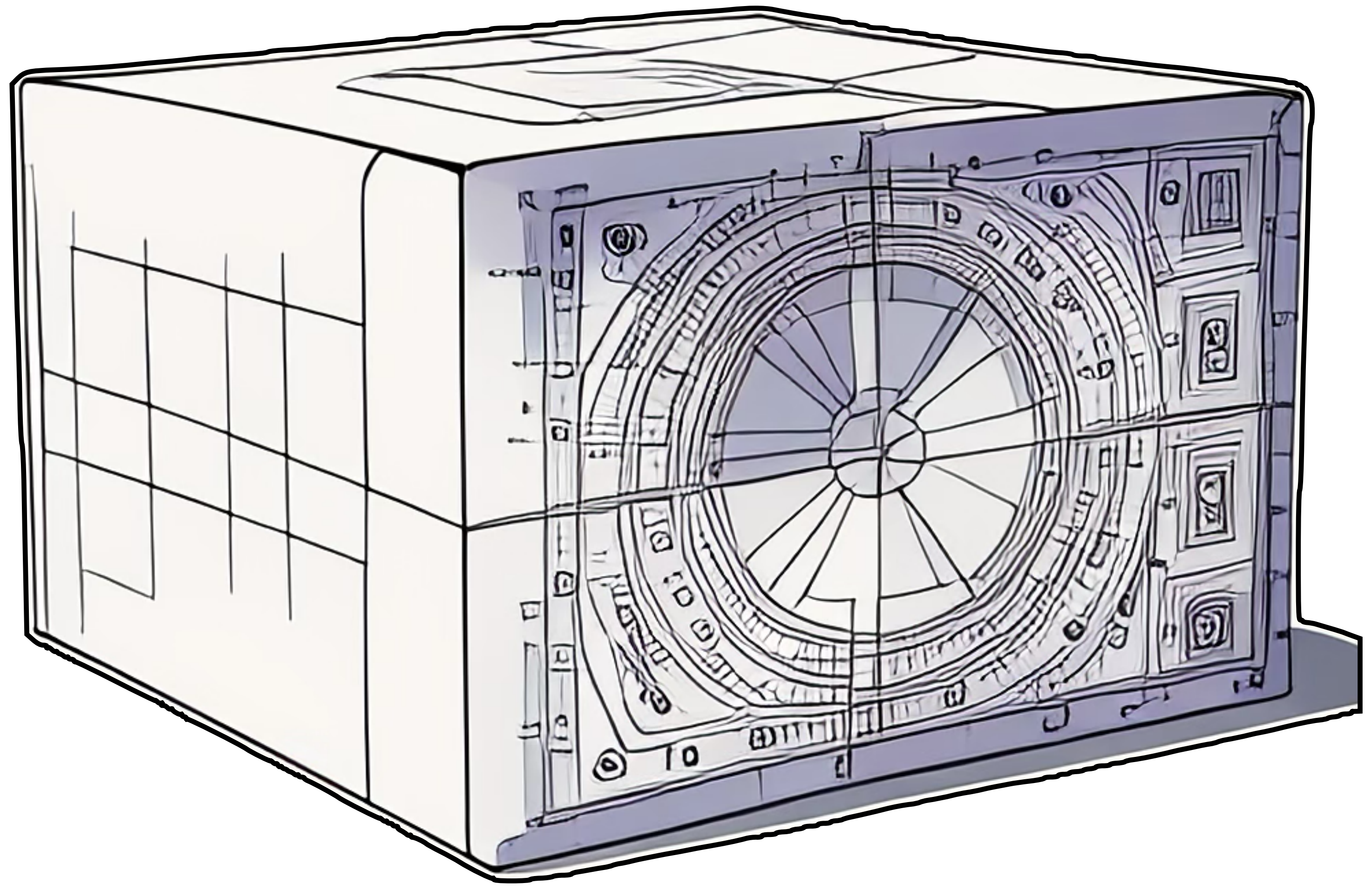


"Essentially, all models are wrong, but some are useful,"

G.E.P. Box and N.R. Draper, Empirical Model-Building and Response Surfaces, John Wiley and Sons, New York, 1987.

Epistemic or Systematic Uncertainties

- Model errors, model discrepancies, or model inadequacies
- The famous quote of Georg E.P. Box
  - **Invented and developed the idea of robustness**
  - **Had a key role in the integration of the classical and Bayesian schools of thought in Statistical Inference.**
- Sources of Systematic Uncertainties:
  - **Simplifying model assumptions**
  - **Missing physics**

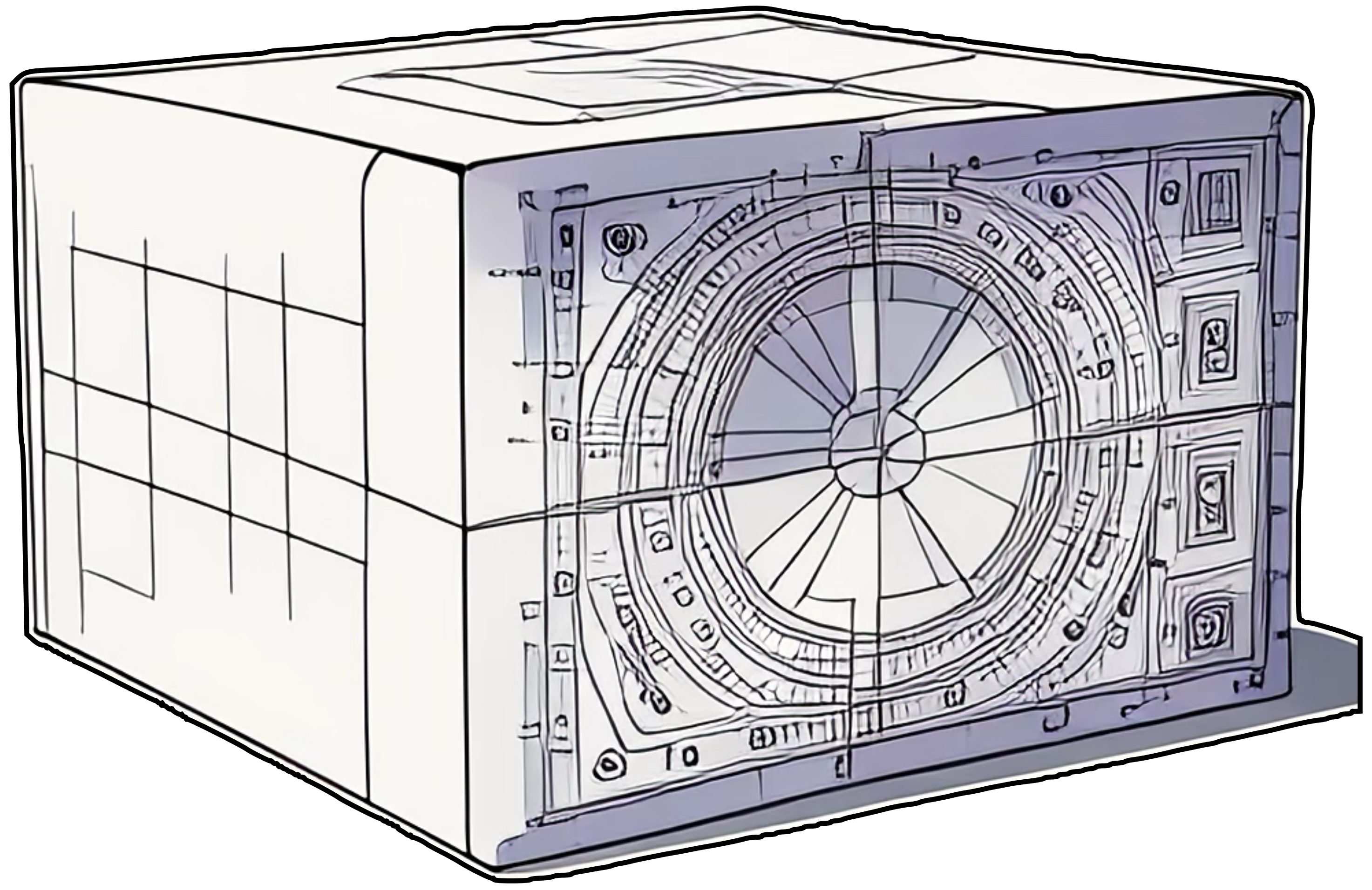


“Essentially, all models are wrong, but some are useful,“

G.E.P. Box and N.R. Draper, Empirical Model-Building and Response Surfaces, John Wiley and Sons, New York, 1987.

Epistemic or Systematic Uncertainties

- Model errors, model discrepancies, or model inadequacies
- The famous quote of Georg E.P. Box
    - **Invented and developed the idea of robustness**
    - **Had a key role in the integration of the classical and Bayesian schools of thought in Statistical Inference.**
- Sources of Systematic Uncertainties:
    - **Simplifying model assumptions**
    - **Missing physics**
    - **Basic lack of knowledge**

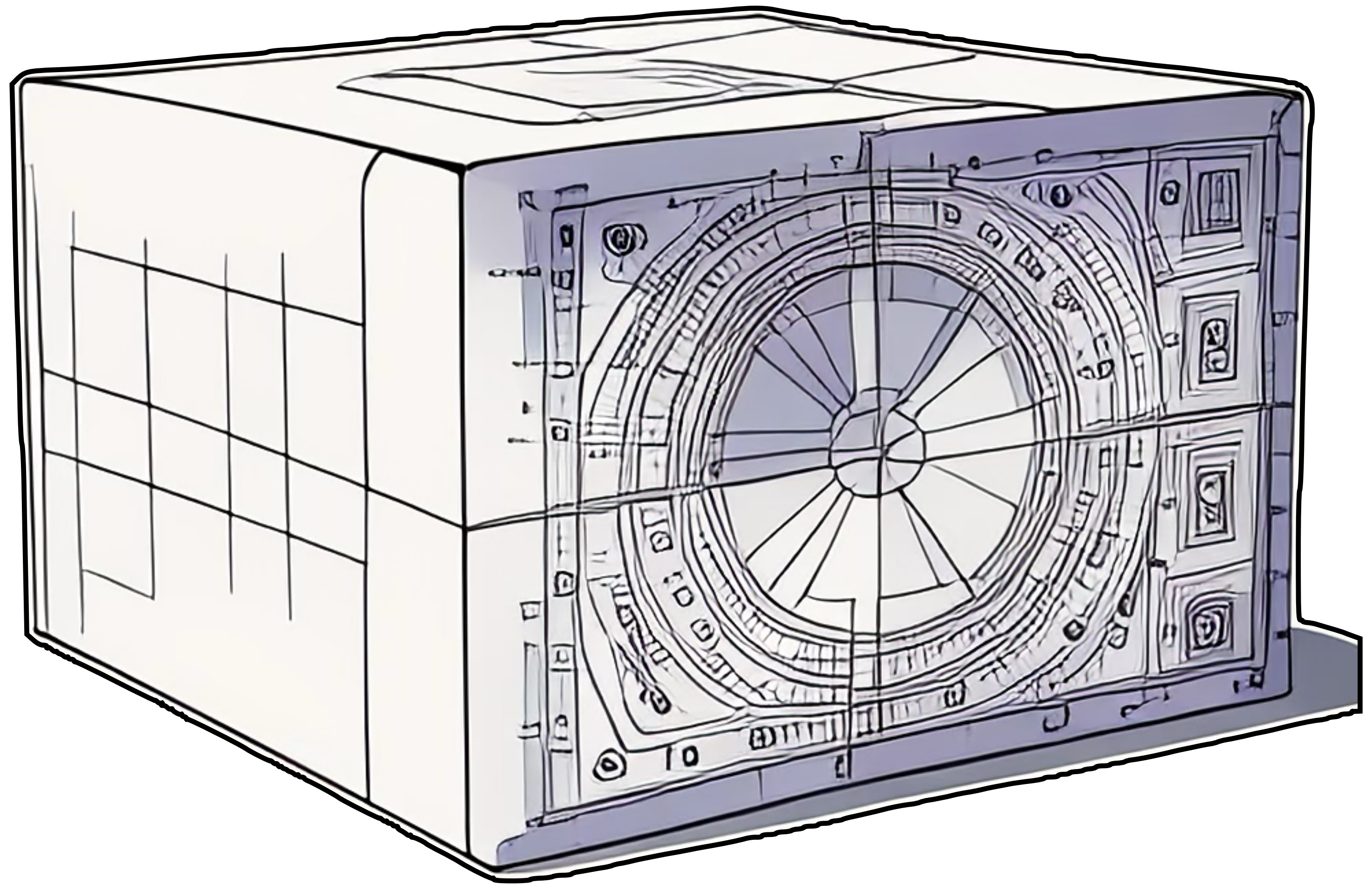


“Essentially, all models are wrong, but some are useful,“

G.E.P. Box and N.R. Draper, Empirical Model-Building and Response Surfaces, John Wiley and Sons, New York, 1987.

Epistemic or Systematic Uncertainties

- Model errors, model discrepancies, or model inadequacies
- The famous quote of Georg E.P. Box
    - **Invented and developed the idea of robustness**
    - **Had a key role in the integration of the classical and Bayesian schools of thought in Statistical Inference.**
- Sources of Systematic Uncertainties:
    - **Simplifying model assumptions**
    - **Missing physics**
    - **Basic lack of knowledge**
- Uncertainty Propagation

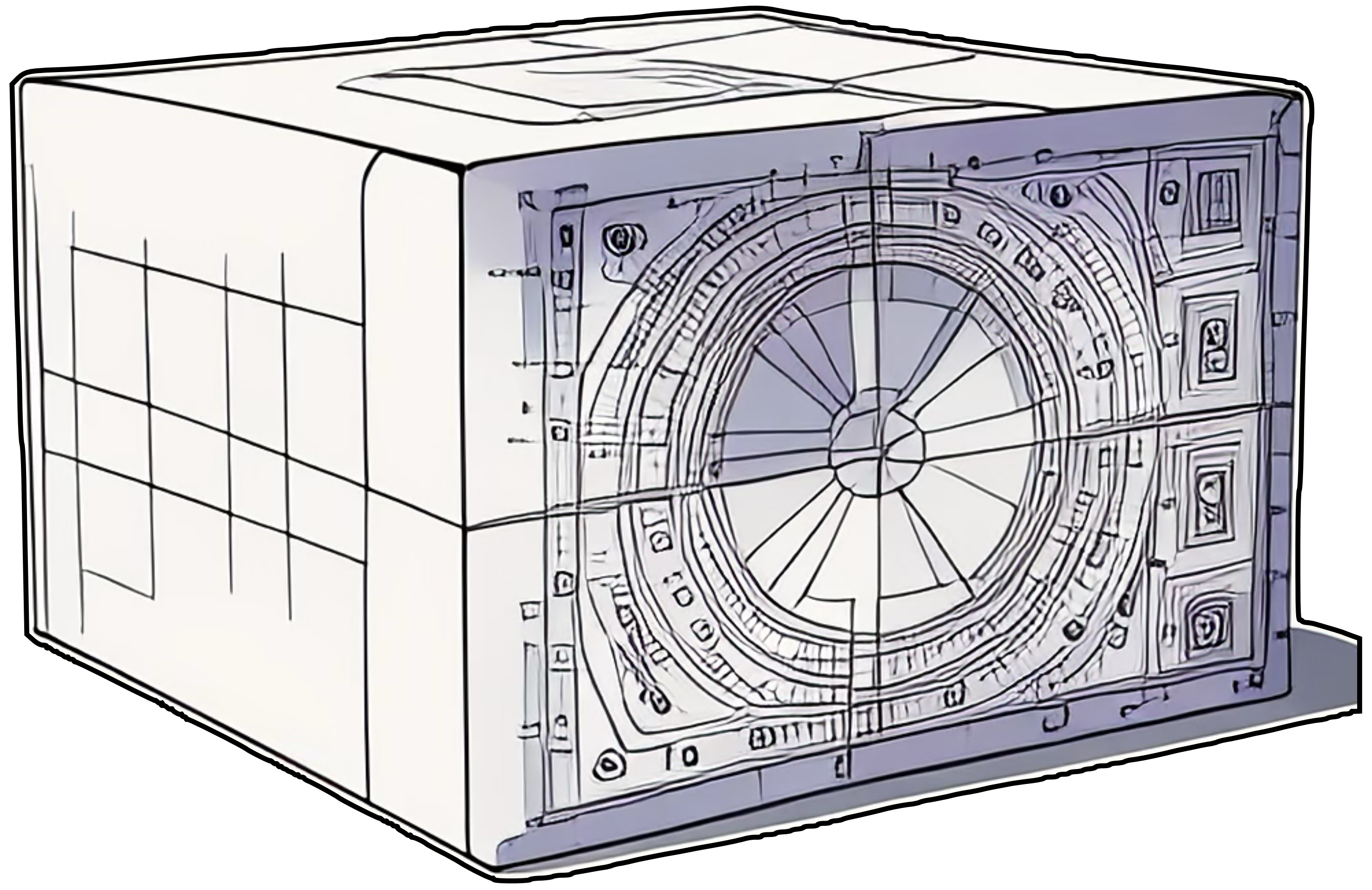


“Essentially, all models are wrong, but some are useful,“

G.E.P. Box and N.R. Draper, Empirical Model-Building and Response Surfaces, John Wiley and Sons, New York, 1987.

Epistemic or Systematic Uncertainties

- Model errors, model discrepancies, or model inadequacies
- The famous quote of Georg E.P. Box
  - **Invented and developed the idea of robustness**
  - **Had a key role in the integration of the classical and Bayesian schools of thought in Statistical Inference.**
- Sources of Systematic Uncertainties:
  - **Simplifying model assumptions**
  - **Missing physics**
  - **Basic lack of knowledge**
- Uncertainty Propagation
  - **Propagate systematic uncertainties through models to construct interval estimates**

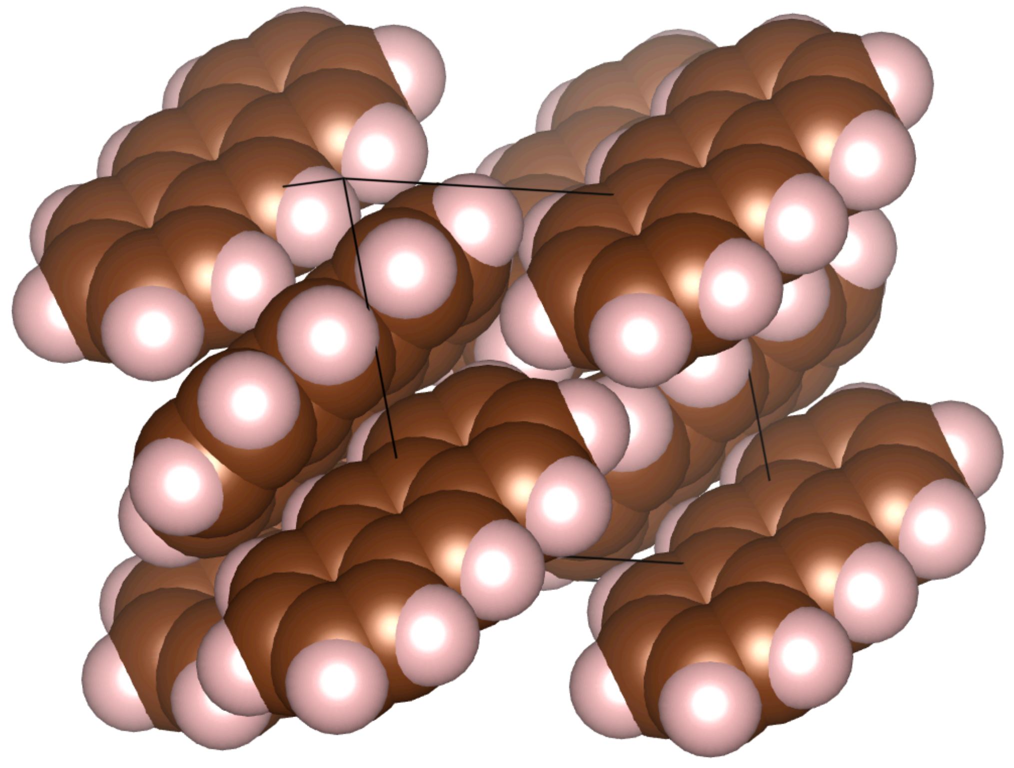


“Essentially, all models are wrong, but some are useful,“

G.E.P. Box and N.R. Draper, Empirical Model-Building and Response Surfaces, John Wiley and Sons, New York, 1987.

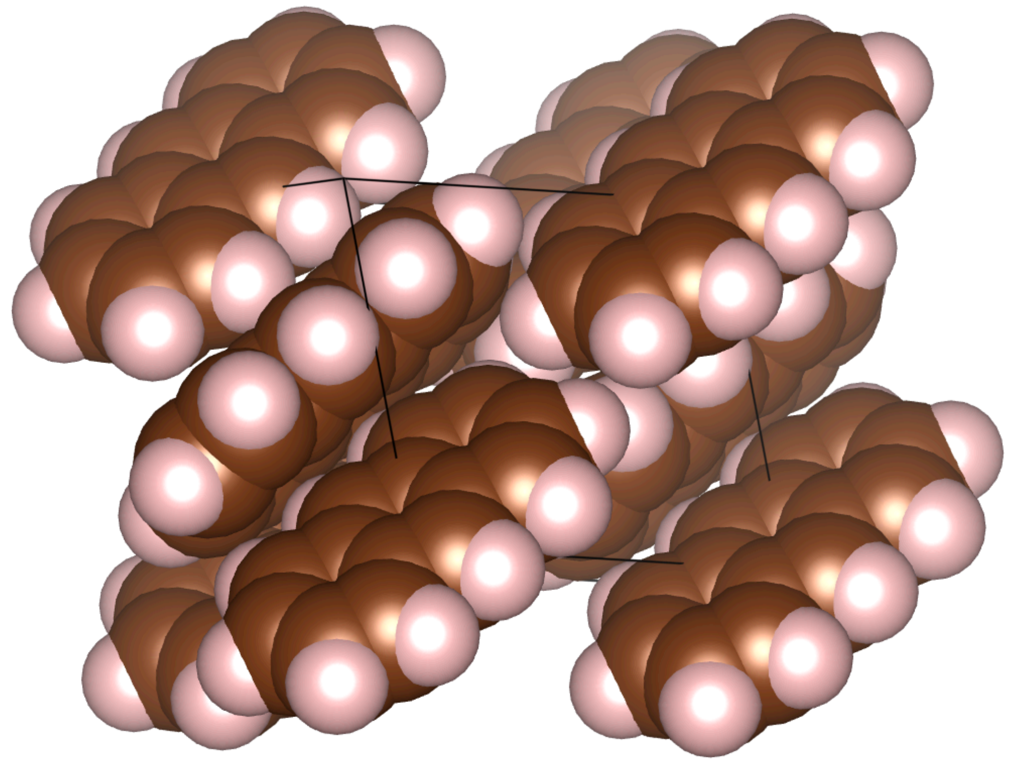
Example: Computing the Packing Coefficient

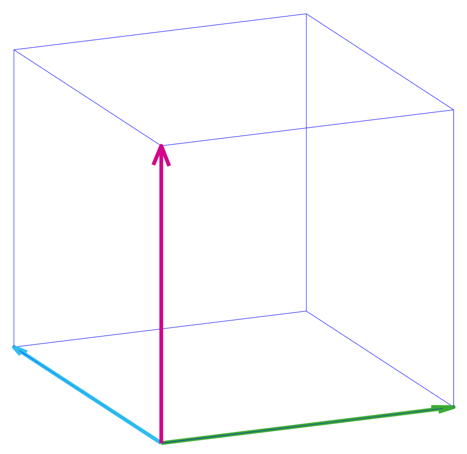
- Packing coefficient for periodic configurations: $\rho = \dfrac{\text{vol}(O)}{\text{vol}(U)}$



Example: Computing the Packing Coefficient

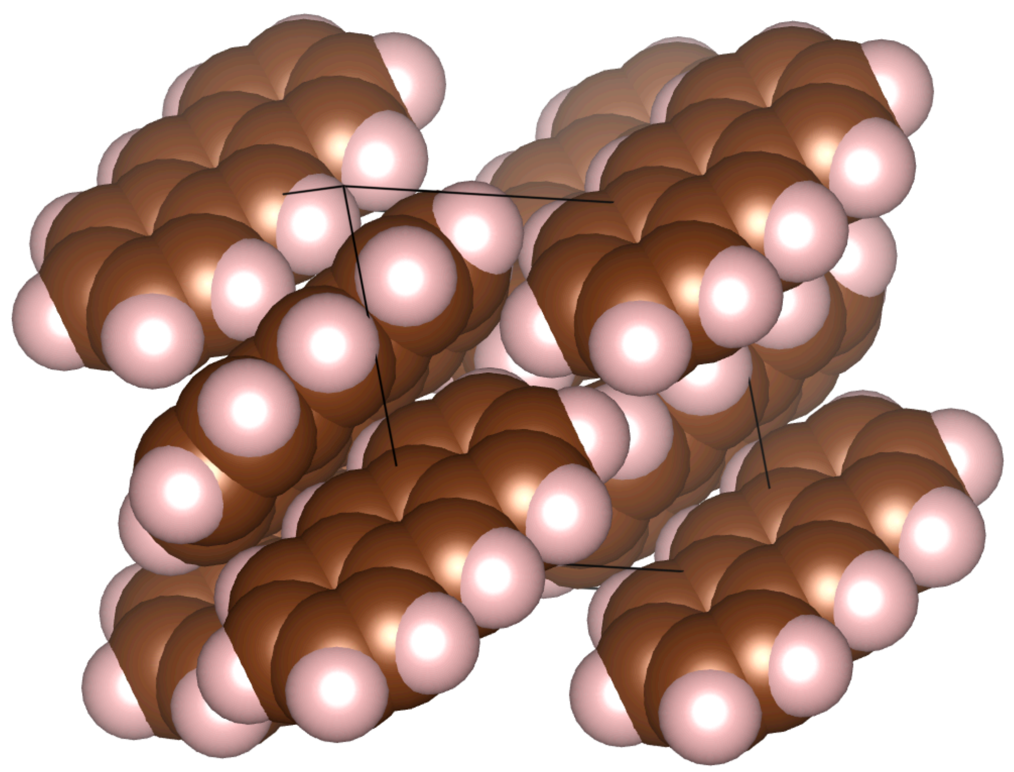
- Packing coefficient for periodic configurations: $\rho = \dfrac{\text{vol}(O)}{\text{vol}(U)}$



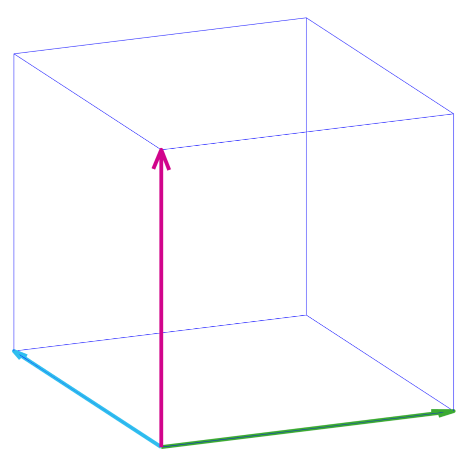


- $\text{vol}(U) = \det(U)$

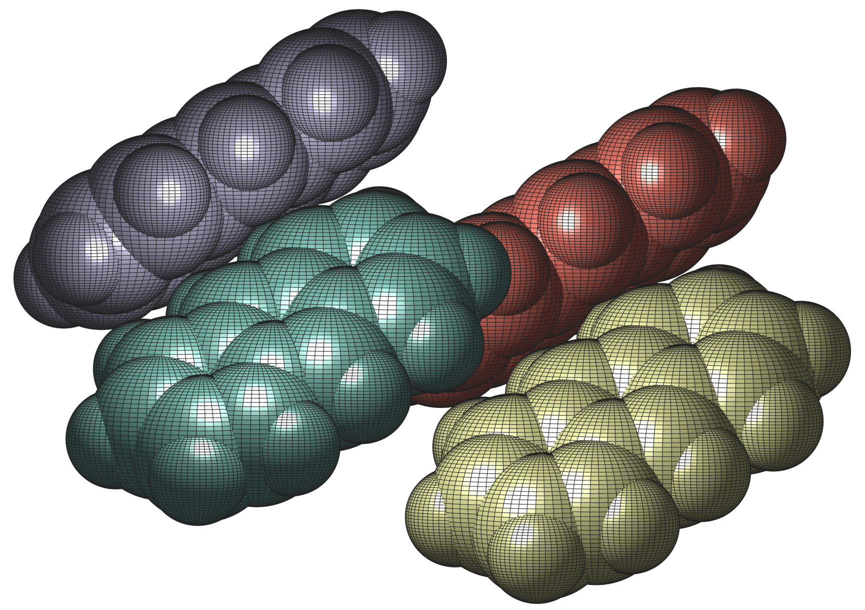
Example: Computing the Packing Coefficient



- Packing coefficient for periodic configurations: $\rho = \dfrac{\mathrm{vol}(O)}{\mathrm{vol}(U)}$



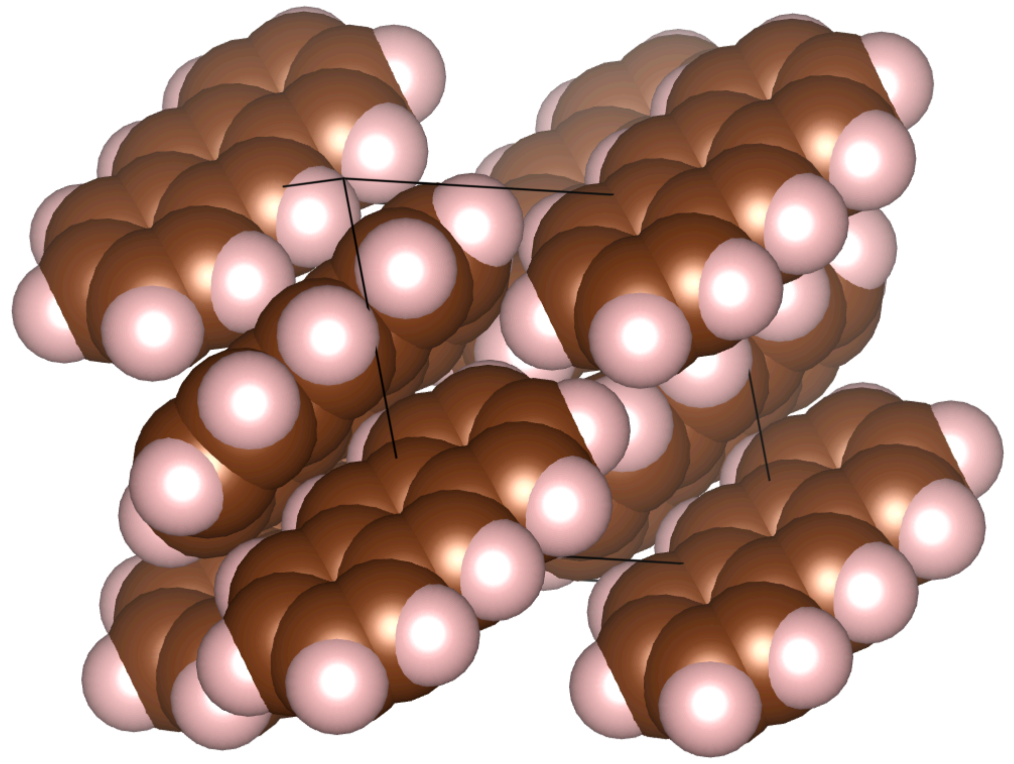
- $\mathrm{vol}(U) = \det(U)$



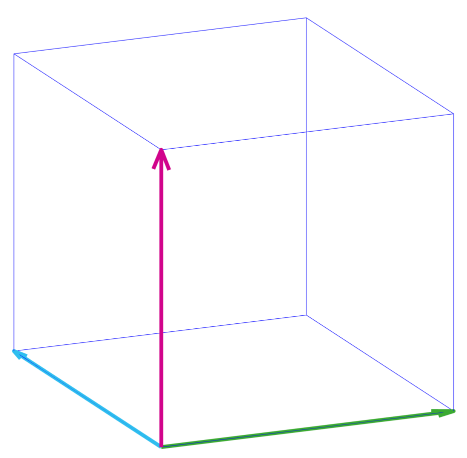
- $\mathrm{vol}(O) = \displaystyle\int_O dV$
  - **O - subset of occupied by the van der Waals spheres**
  - **dV - natural volume form on $\mathbf{R}^3$**

Example: Computing the Packing Coefficient

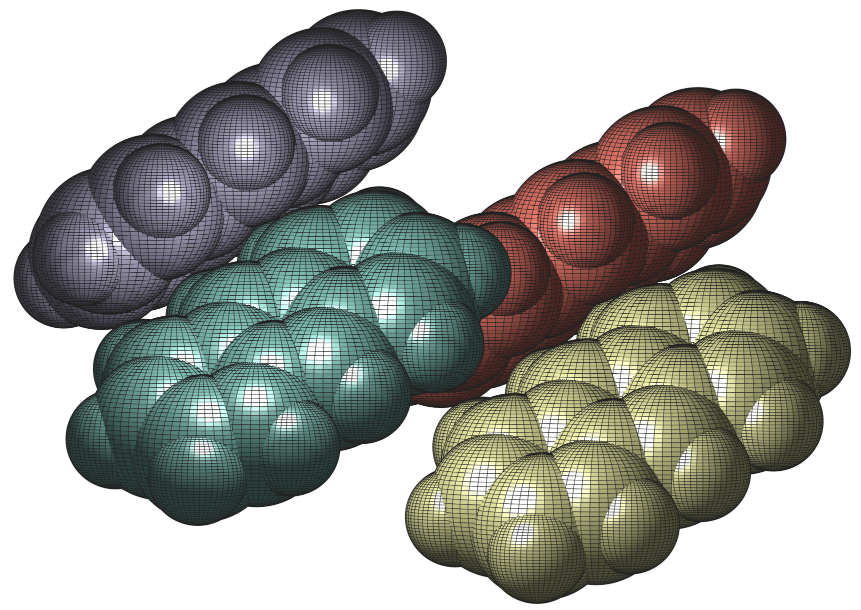
- Packing coefficient for periodic configurations: $\rho = \dfrac{\mathrm{vol}(O)}{\mathrm{vol}(U)}$

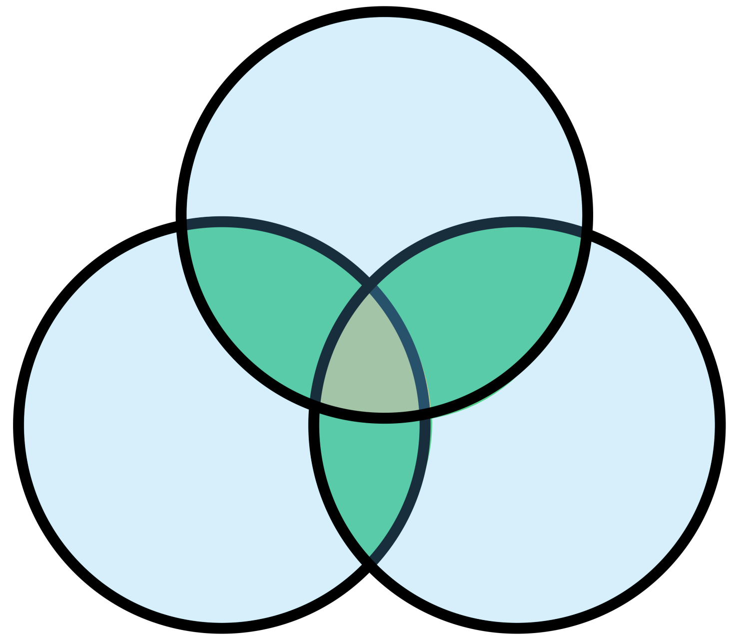


- $\mathrm{vol}(U) = \det(U)$



- $\mathrm{vol}(O) = \int_O dV$



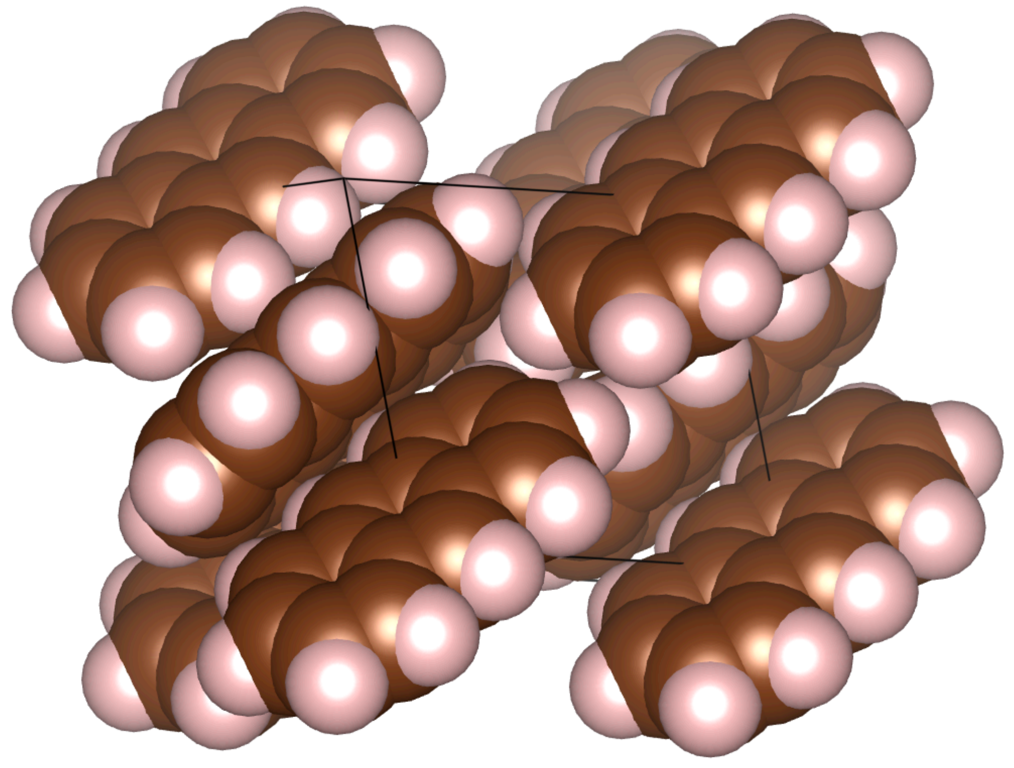


$S_1 + S_2 + S_3 - S_1 \cap S_2 - S_1 \cap S_3 - S_2 \cap S_3 + S_1 \cap S_2 \cap S_3$

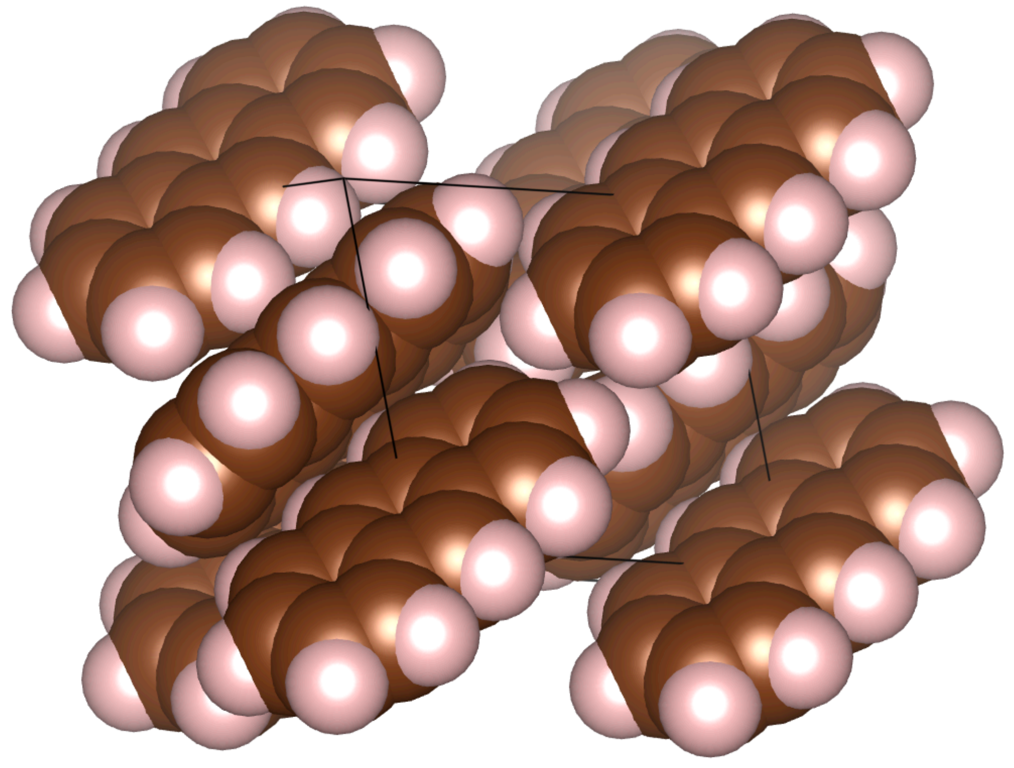
  - **O - subset of occupied by the van der Waals spheres**
  - **dV - natural volume form on $\mathbf{R}^3$**

Alternative: Monte-Carlo Estimation

- $\text{vol}(O) = \int \mathbf{1}_O dU$
  - $\mathbf{1}_O$ - indicator function over U



Alternative: Monte-Carlo Estimation

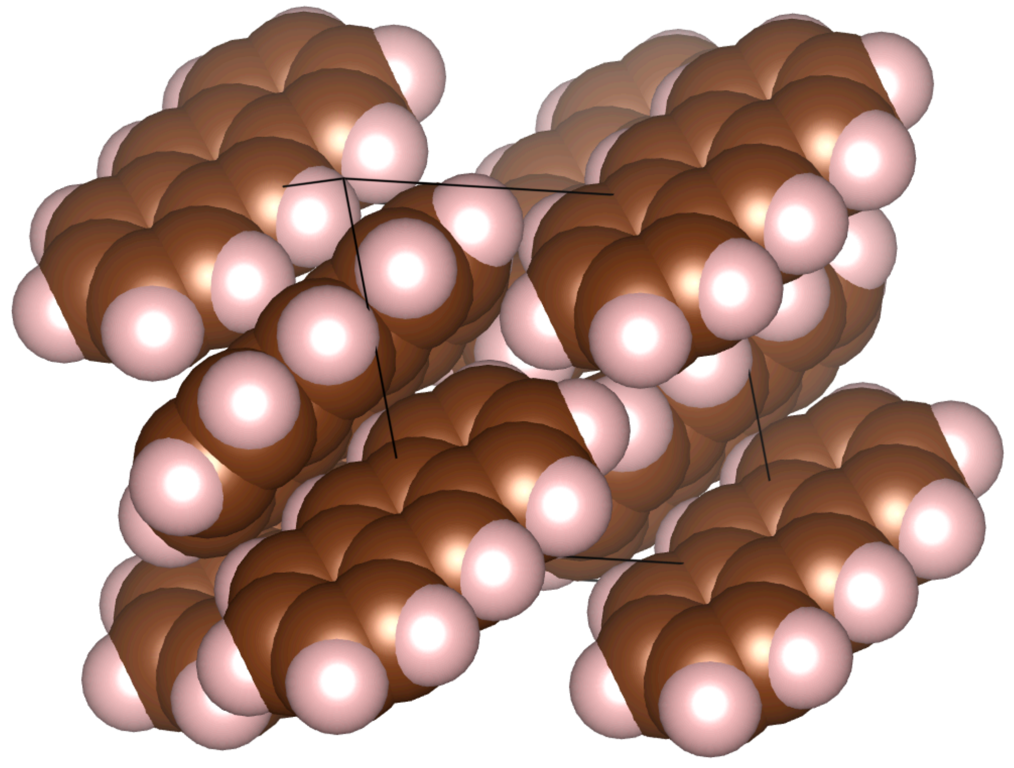


- $\text{vol}(O) = \int \mathbf{1}_O dU$
  - $\mathbf{1}_O$ **- indicator function over** U

- Change of coordinates to an integral over the unit cube C

$$\text{vol}(O) = \int \mathbf{1}_{U^{-1}O} \text{det}(U) dC$$

Alternative: Monte-Carlo Estimation



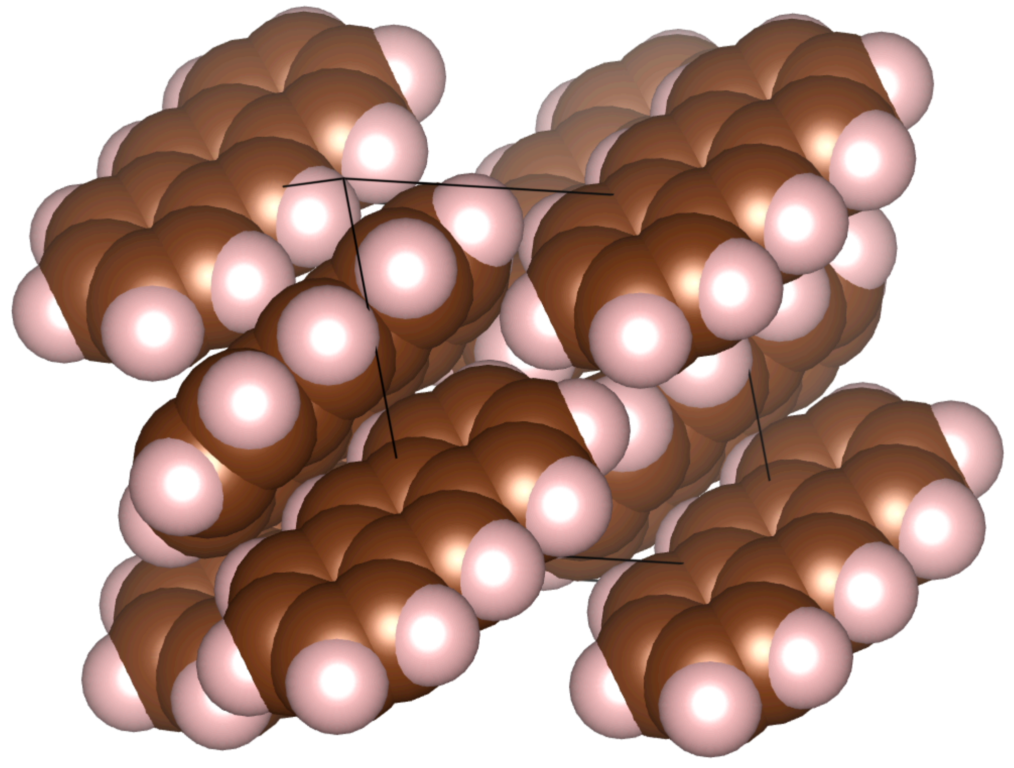
- $\text{vol}(O) = \int \mathbf{1}_O dU$
  - $\mathbf{1}_O$ **- indicator function over** U

- Change of coordinates to an integral over the unit cube C

$$\text{vol}(O) = \int \mathbf{1}_{U^{-1}O} \det(U) dC$$

- $\rho = \dfrac{\text{vol}(O)}{\text{vol}(U)}$ $\longrightarrow$ $\rho = \int \mathbf{1}_{U^{-1}O}(U) dC$

Alternative: Monte-Carlo Estimation



- $\text{vol}(O) = \int \mathbf{1}_O dU$
  - $\mathbf{1}_O$ **- indicator function over** U

- Change of coordinates to an integral over the unit cube C

$$\text{vol}(O) = \int \mathbf{1}_{U^{-1}O} \text{det}(U) dC$$

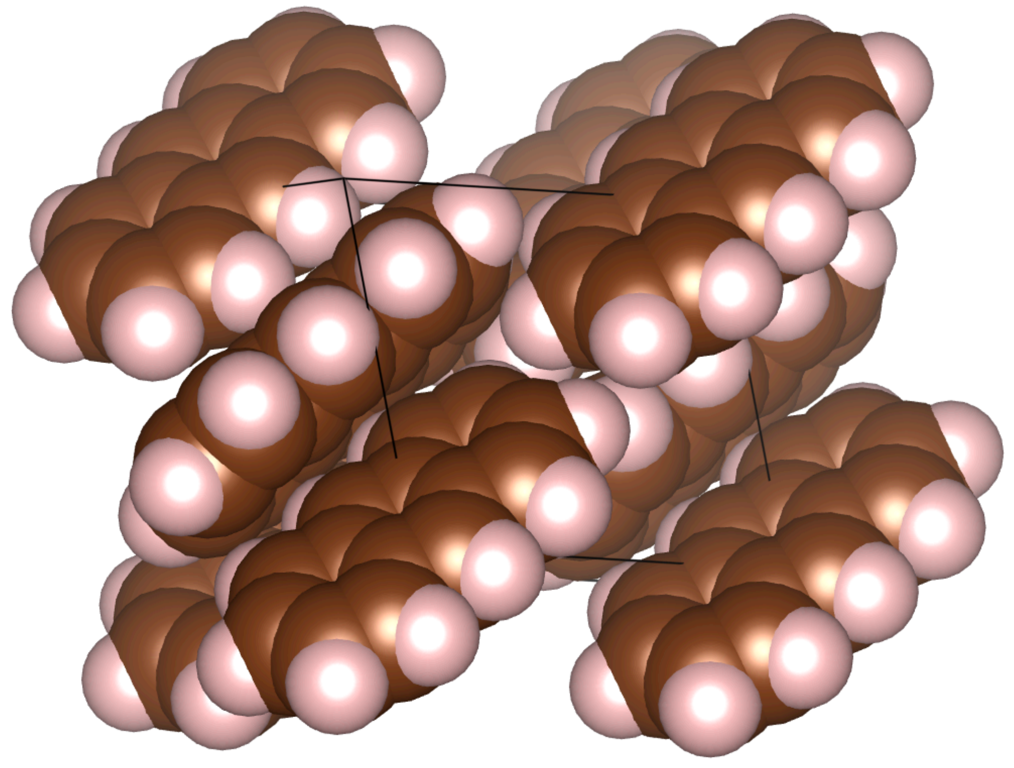
- $\rho = \dfrac{\text{vol}(O)}{\text{vol}(U)}$ ➡ $\rho = \int \mathbf{1}_{U^{-1}O}(U) dC$

- Uniform random variable on a unit cube $C \sim \mathbf{X}$

$$\rho = \mathbf{E}\left[1_{U^{-1}O}\right] = \mathrm{P}\left(X \in U^{-1}O\right)$$

Alternative: Monte-Carlo Estimation

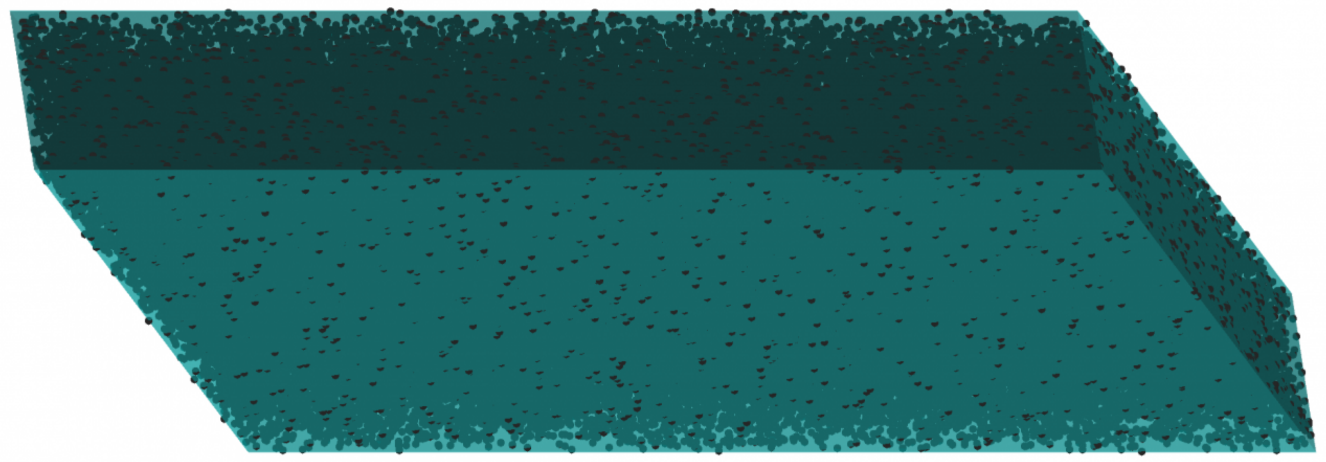
- vol$(O) = \int \mathbf{1}_O dU$
  - $\mathbf{1}_O$ - **indicator function over** U



- Change of coordinates to an integral over the unit cube C

$$\text{vol}(O) = \int \mathbf{1}_{U^{-1}O} \det(U) dC$$

- $\rho = \dfrac{\text{vol}(O)}{\text{vol}(U)}$ $\longrightarrow$ $\rho = \int \mathbf{1}_{U^{-1}O}(U) dC$



- Uniform random variable on a unit cube $C \sim \mathbf{X}$

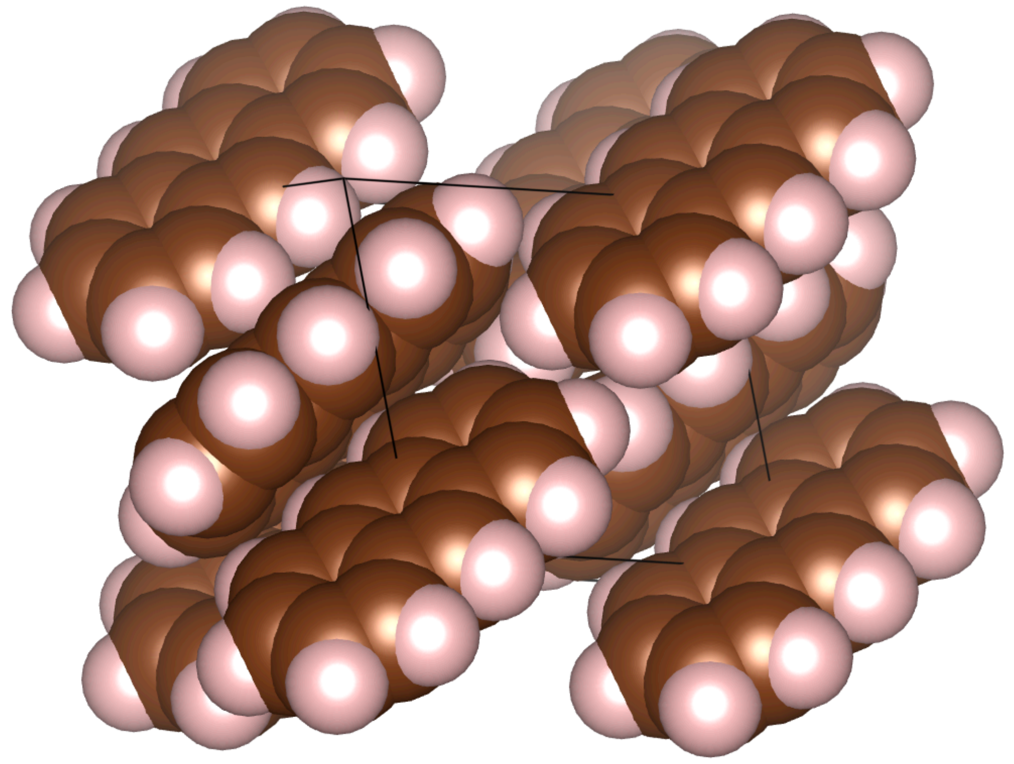
$$\rho = \mathbf{E}\left[1_{U^{-1}O}\right] = \mathrm{P}\left(X \in U^{-1}O\right)$$

- Draw realisations o a random vector

$$X_1, \ldots X_N \sim C$$

Alternative: Monte-Carlo Estimation

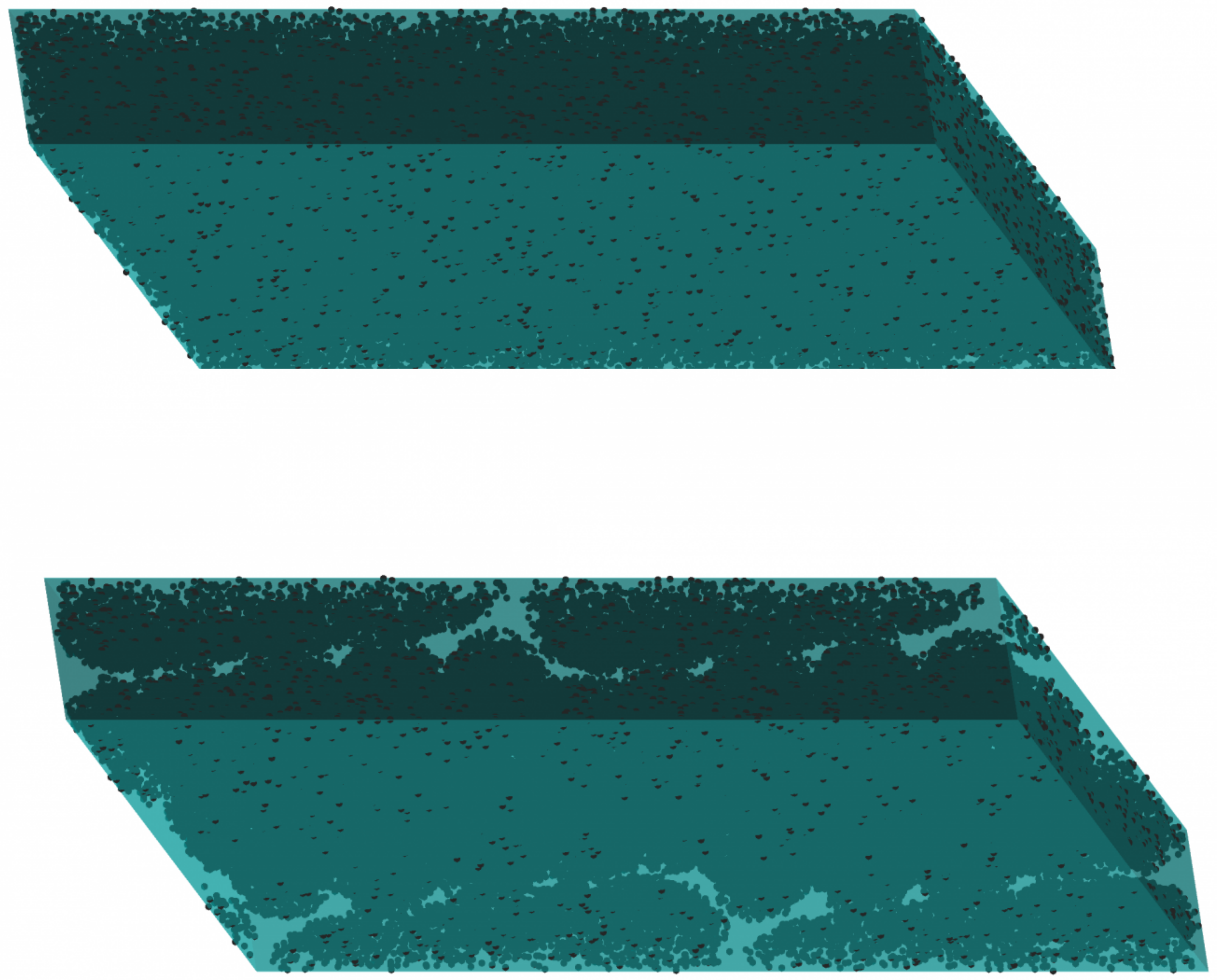
- $\mathrm{vol}(O) = \int \mathbf{1}_O dU$
  - $\mathbf{1}_O$ **- indicator function over** U



- Change of coordinates to an integral over the unit cube C

$$\mathrm{vol}(O) = \int \mathbf{1}_{U^{-1}O} \det(U) dC$$

- $\rho = \dfrac{\mathrm{vol}(O)}{\mathrm{vol}(U)}$ $\longrightarrow$ $\rho = \int \mathbf{1}_{U^{-1}O}(U) dC$



- Uniform random variable on a unit cube $C \sim \mathbf{X}$

$$\rho = \mathbf{E}\left[1_{U^{-1}O}\right] = \mathrm{P}\left(X \in U^{-1}O\right)$$

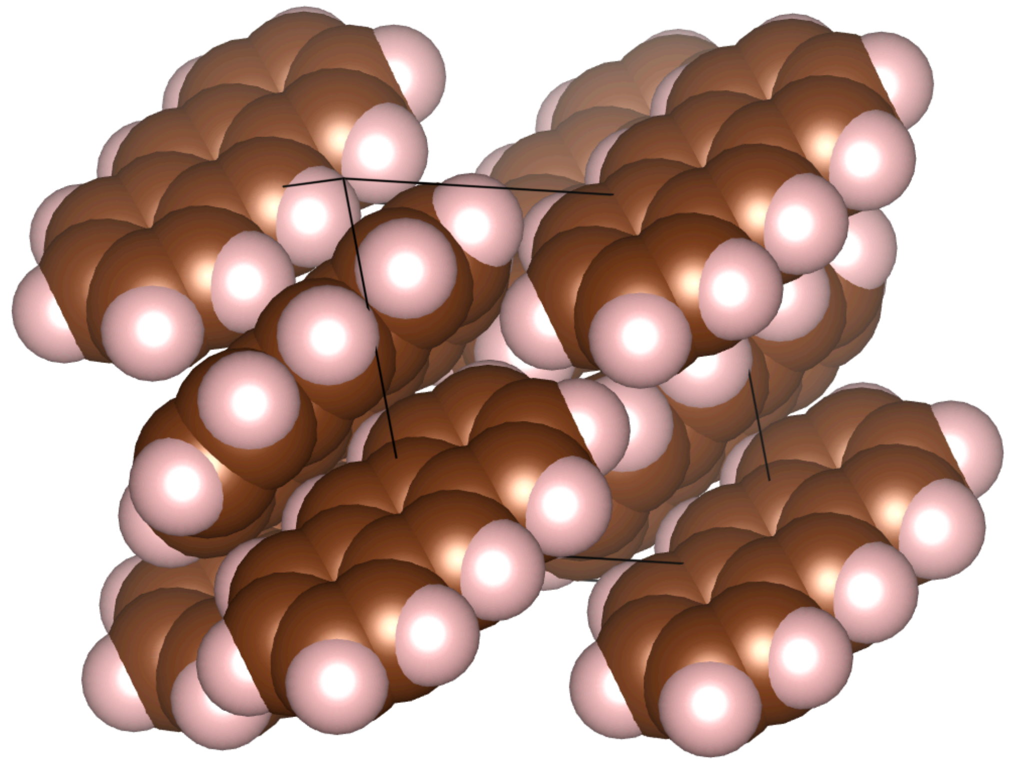
- Draw realisations o a random vector

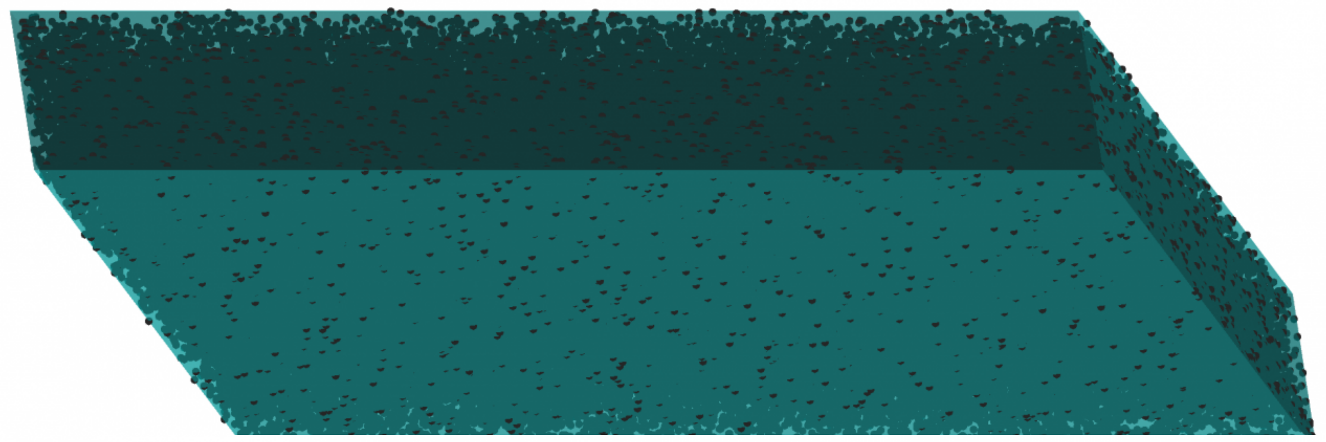
$$X_1, \ldots X_N \sim C$$

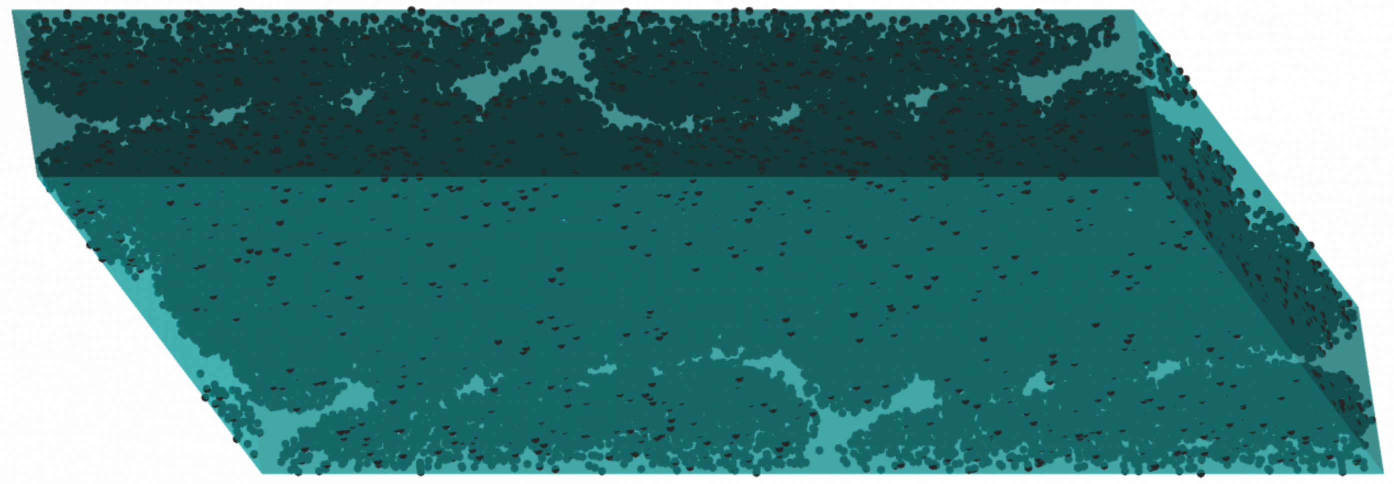
- Estimate the packing coefficient $\hat{\rho} = \dfrac{\#\{X_i \in U^{-1}O\}}{N}$

Alternative: Monte-Carlo Estimation

- N=100000 $\#\left\{X_i \in U^{-1}O\right\} = 76812$

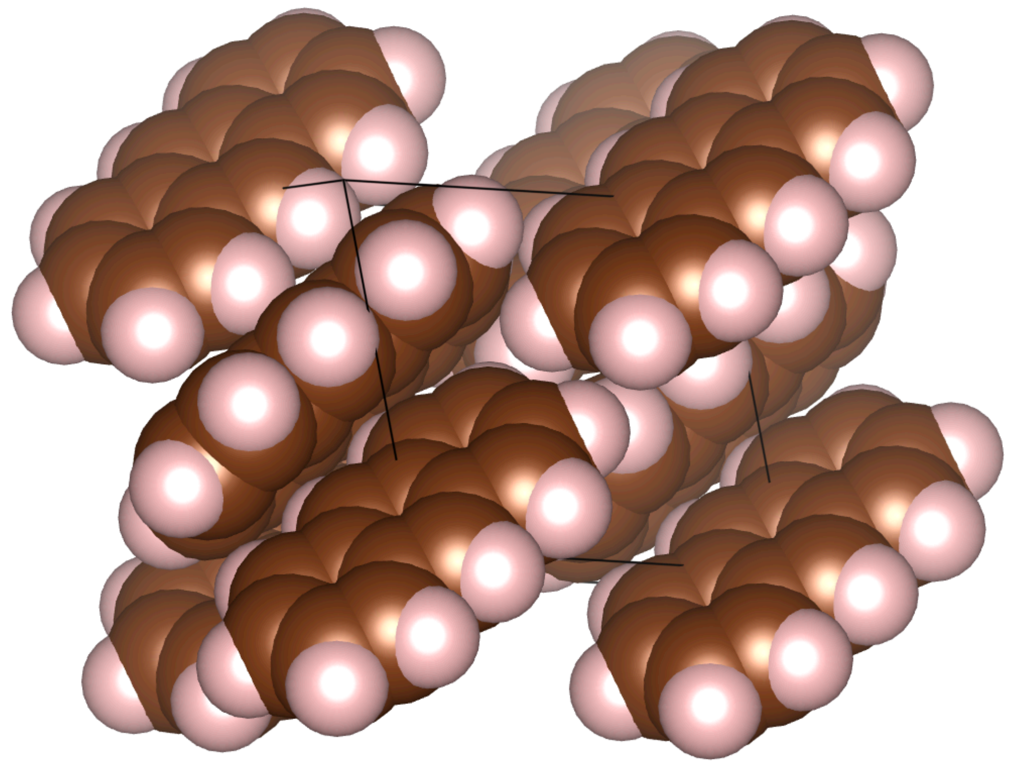






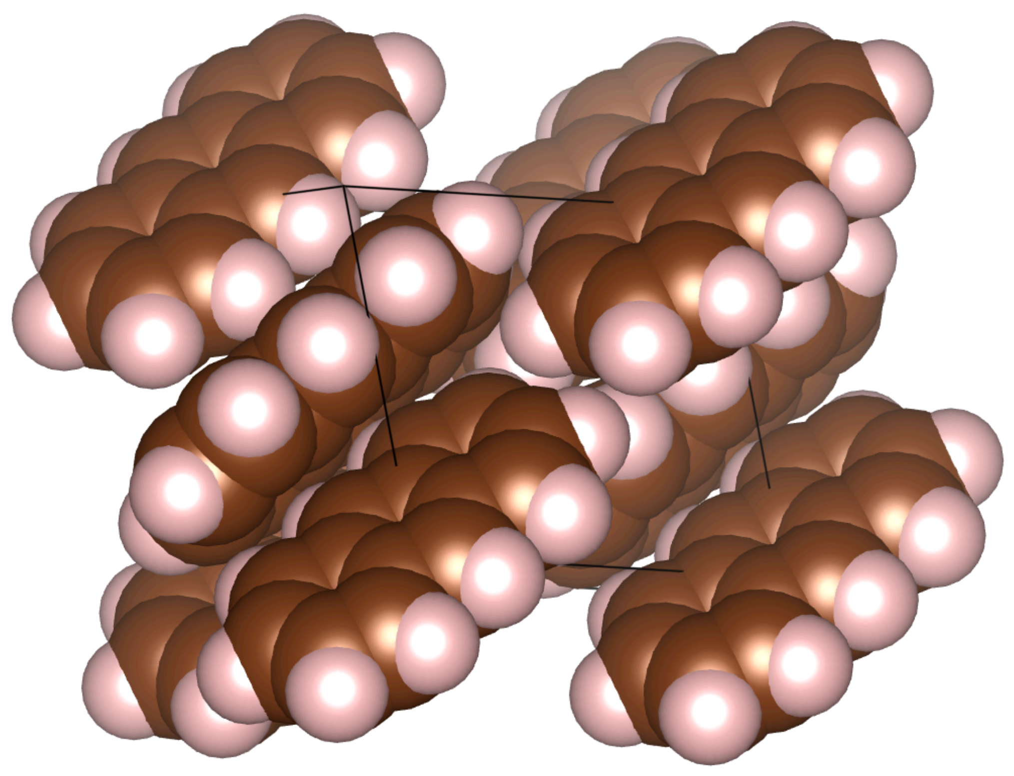
Alternative: Monte-Carlo Estimation

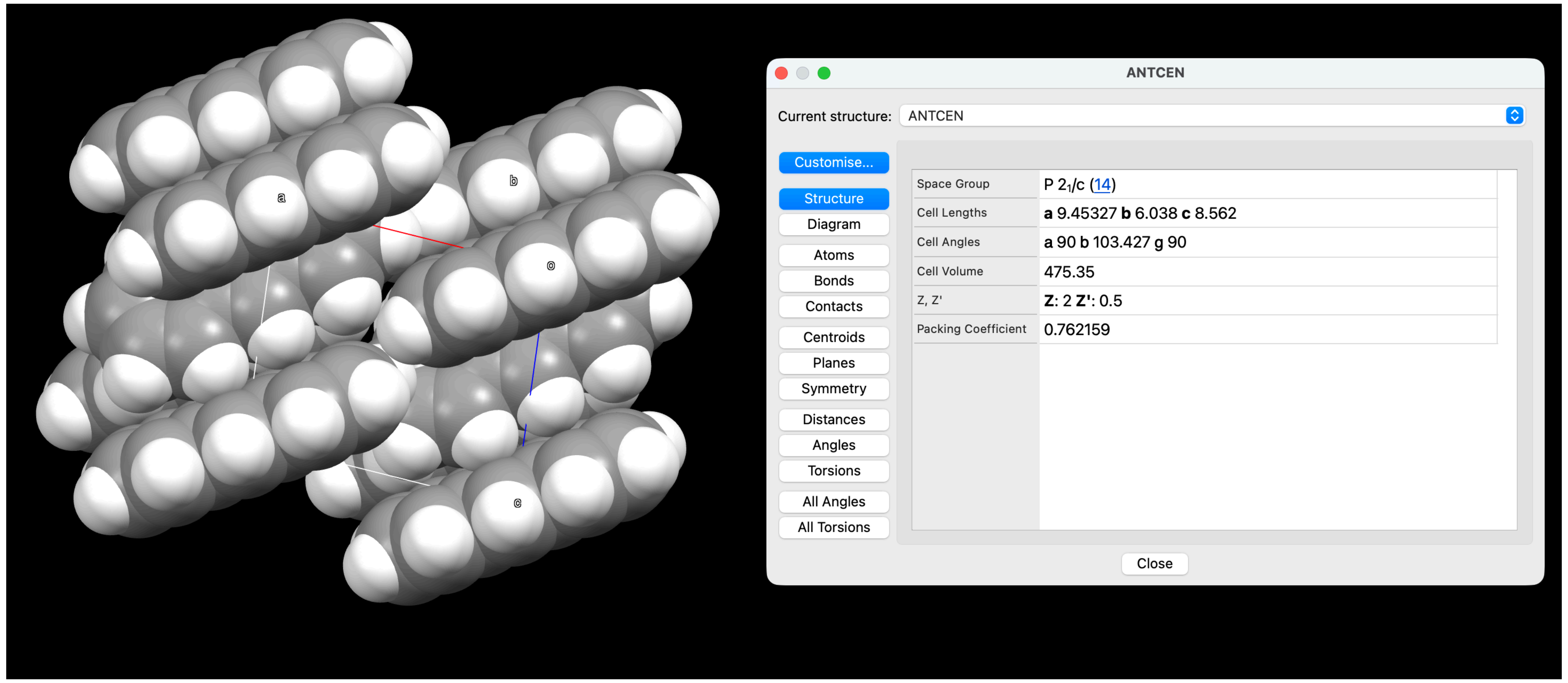
- N=100000 $\#\left\{X_i \in U^{-1}O\right\} = 76812$
- $\hat{\rho} = 0.76812$

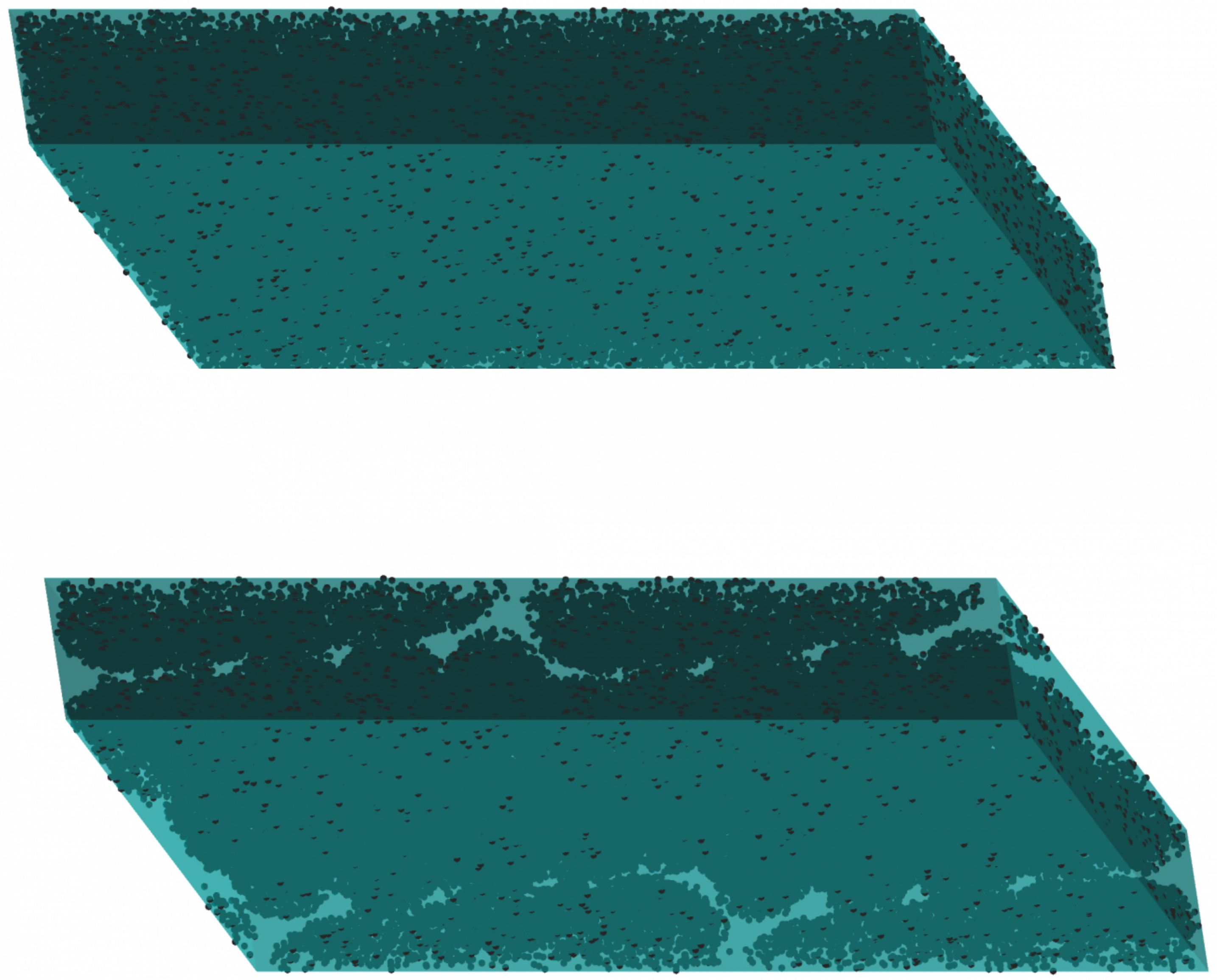


Alternative: Monte-Carlo Estimation

- N=100000 $\#\{X_i \in U^{-1}O\} = 76812$
- $\hat{\rho} = 0.76812$
- CCDC Mercury: 0.762159

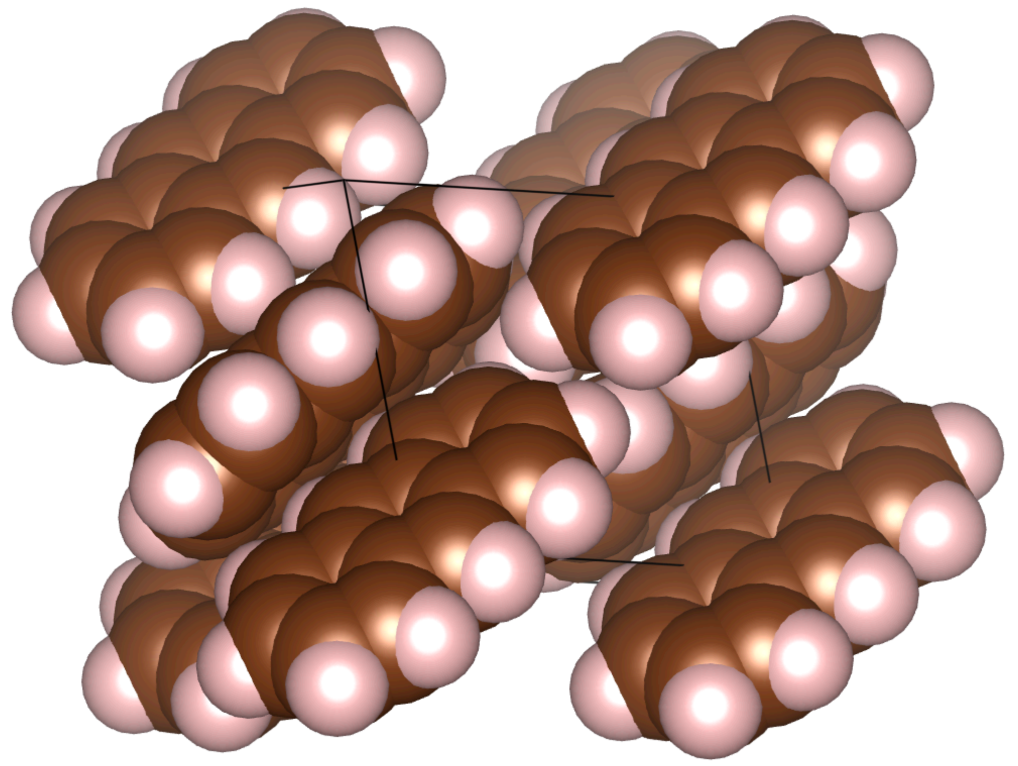


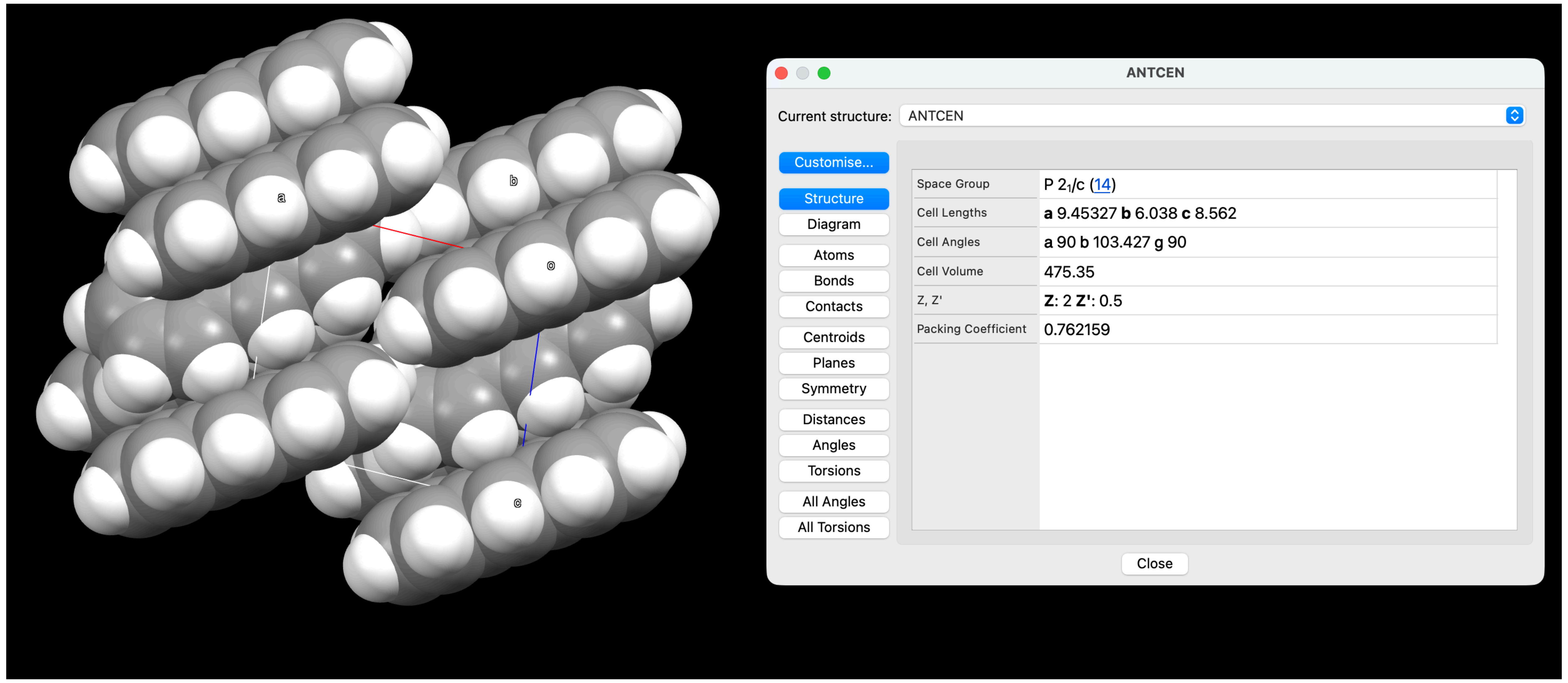


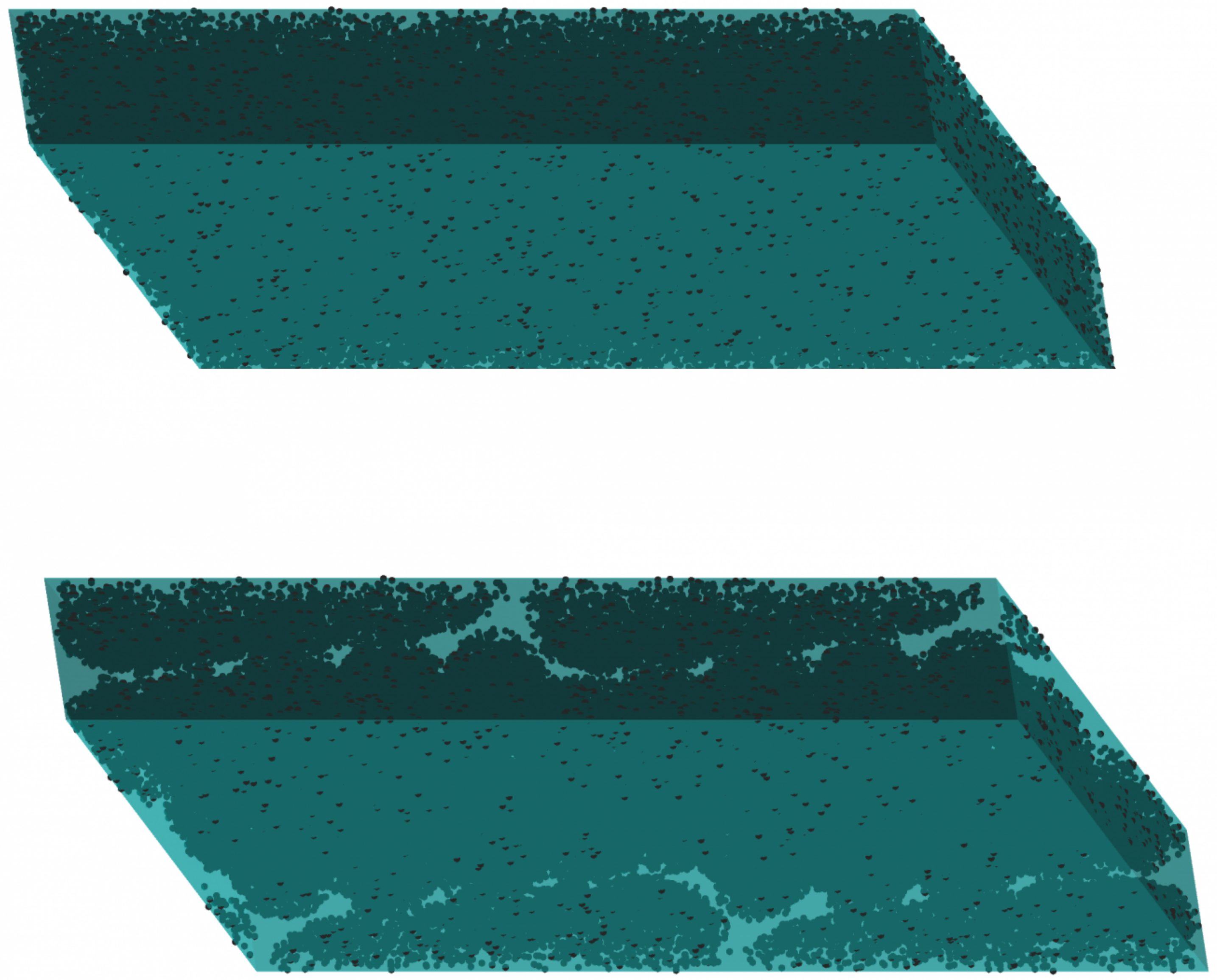


Alternative: Monte-Carlo Estimation

- N=100000 $\#\{X_i \in U^{-1}O\} = 76812$
- $\hat{\rho} = 0.76812$
- CCDC Mercury: 0.762159
- Which point estimate to trust?







Interval Estimate of the Packing Coefficient

- X_i - Bernoulli trial $\quad \mathbf{X} = \sum_{i=1}^{N} X_i \quad$ binomially distributed random variable

Expected value: $N\rho$

Variance: $N\rho\left(1-\rho\right)$

Interval Estimate of the Packing Coefficient

- X_i - Bernoulli trial $\quad \mathbf{X} = \sum_{i=1}^{N} X_i \quad$ binomially distributed random variable

Expected value: $N\rho$

Variance: $N\rho\left(1-\rho\right)$

- Central limit theorem: $\quad \mathbf{Z} = \dfrac{\sqrt{N}\left(\hat{\rho}-\rho\right)}{\sqrt{\rho\left(1-\rho\right)}} \xrightarrow{F} N(0,1)$

Z converges in distribution to a normally distributed random variable with a mean of 0 and a variance of 1

Interval Estimate of the Packing Coefficient

- X_i - Bernoulli trial $\quad \mathbf{X} = \sum_{i=1}^{N} X_i \quad$ binomially distributed random variable

Expected value: $N\rho$

Variance: $N\rho\left(1-\rho\right)$

- Central limit theorem: $\mathbf{Z} = \dfrac{\sqrt{N}\left(\hat{\rho}-\rho\right)}{\sqrt{\rho\left(1-\rho\right)}} \xrightarrow{F} N(0,1)$

Z converges in distribution to a normally distributed random variable with a mean of 0 and a variance of 1

- Confidence interval: $P\left(\underbrace{\left|\dfrac{(\rho-\hat{\rho})\sqrt{N}}{\sqrt{\hat{\rho}\left(1-\hat{\rho}\right)}}\right|}\geq 2z\right) = 1-\alpha$

Random quantity is the interval.

- CCDC Mercury: 0.762159

Interval Estimate of the Packing Coefficient

- X_i - Bernoulli trial $\quad \mathbf{X} = \sum_{i=1}^{N} X_i \quad$ binomially distributed random variable

Expected value: $N\rho$

Variance: $N\rho\left(1-\rho\right)$

- Central limit theorem: $\mathbf{Z} = \dfrac{\sqrt{N}\left(\hat{\rho}-\rho\right)}{\sqrt{\rho\left(1-\rho\right)}} \xrightarrow{F} N(0,1)$

Z converges in distribution to a normally distributed random variable with a mean of 0 and a variance of 1

- Confidence interval: $P\left(\underbrace{\left|\dfrac{(\rho-\hat{\rho})\sqrt{N}}{\sqrt{\hat{\rho}\left(1-\hat{\rho}\right)}}\right|}\geq 2z\right) = 1-\alpha$

Random quantity is the interval.

- Interval estimate: $\left[\hat{\rho}-\dfrac{\hat{\rho}(1-\hat{\rho})}{\sqrt{N}}z;\hat{\rho}+\dfrac{\hat{\rho}(1-\hat{\rho})}{\sqrt{N}}z\right]$

z **-** $1-\alpha$ **quantile of the standard normal distribution**

Interval Estimate of the Packing Coefficient

- X_i - Bernoulli trial $\mathbf{X} = \sum_{i=1}^{N} X_i$ binomially distributed random variable

Expected value: $N\rho$

Variance: $N\rho\left(1-\rho\right)$

- Central limit theorem: $\mathbf{Z} = \frac{\sqrt{N}\left(\hat{\rho}-\rho\right)}{\sqrt{\rho\left(1-\rho\right)}} \xrightarrow{F} N(0,1)$

Z converges in distribution to a normally distributed random variable with a mean of 0 and a variance of 1

- Confidence interval: $P\left(\underbrace{\left|\frac{(\rho-\hat{\rho})\sqrt{N}}{\sqrt{\hat{\rho}\left(1-\hat{\rho}\right)}}\right|}\geq 2z\right) = 1-\alpha$

Random quantity is the interval.

- Interval estimate: $\left[\hat{\rho} - \frac{\hat{\rho}(1-\hat{\rho})}{\sqrt{N}}z; \hat{\rho} + \frac{\hat{\rho}(1-\hat{\rho})}{\sqrt{N}}z\right]$

z **-** $1-\alpha$ **quantile of the standard normal distribution**

- A way to express confidence in the packing coefficient estimate $\hat{\rho}$

Interval Estimate of the Packing Coefficient

- X_i - Bernoulli trial $\quad \mathbf{X} = \sum_{i=1}^{N} X_i \quad$ binomially distributed random variable

Expected value: $N\rho$

Variance: $N\rho\left(1-\rho\right)$

- Central limit theorem: $$\mathbf{Z} = \frac{\sqrt{N}\left(\hat{\rho}-\rho\right)}{\sqrt{\rho\left(1-\rho\right)}} \xrightarrow{F} N(0,1)$$

Z converges in distribution to a normally distributed random variable with a mean of 0 and a variance of 1

- Confidence interval: $$P\left(\underbrace{\left|\frac{(\rho-\hat{\rho})\sqrt{N}}{\sqrt{\hat{\rho}\left(1-\hat{\rho}\right)}}\right|}\geq 2z\right) = 1-\alpha$$

Random quantity is the interval.

- Interval estimate: $$\left[\hat{\rho}-\frac{\hat{\rho}(1-\hat{\rho})}{\sqrt{N}}z;\hat{\rho}+\frac{\hat{\rho}(1-\hat{\rho})}{\sqrt{N}}z\right]$$

z **-** $1-\alpha$ **quantile of the standard normal distribution**

- A way to express confidence in the packing coefficient estimate $\hat{\rho}$
  - **The tighter the interval, the higher the confidence in the estimate.**

Interval Estimate of the Packing Coefficient

- X_i - Bernoulli trial $\quad \mathbf{X} = \sum_{i=1}^{N} X_i$ binomially distributed random variable

Expected value: $N\rho$

Variance: $N\rho\left(1-\rho\right)$

- Central limit theorem: $\mathbf{Z} = \frac{\sqrt{N}\left(\hat{\rho}-\rho\right)}{\sqrt{\rho\left(1-\rho\right)}} \xrightarrow{F} N(0,1)$

Z converges in distribution to a normally distributed random variable with a mean of 0 and a variance of 1

- Confidence interval: $$P\left(\underbrace{\left|\frac{(\rho-\hat{\rho})\sqrt{N}}{\sqrt{\hat{\rho}\left(1-\hat{\rho}\right)}}\right|}\geq 2z\right) = 1-\alpha$$

Random quantity is the interval.

- Interval estimate: $$\left[\hat{\rho}-\frac{\hat{\rho}(1-\hat{\rho})}{\sqrt{N}}z; \hat{\rho}+\frac{\hat{\rho}(1-\hat{\rho})}{\sqrt{N}}z\right]$$

z **-** $1-\alpha$ **quantile of the standard normal distribution**

- A way to express confidence in the packing coefficient estimate $\hat{\rho}$
    - **The tighter the interval, the higher the confidence in the estimate.**
- Even better, sample until the size of this interval to be less than some constant c with probability p

$$P\left(|\hat{\rho}-\rho| < 2z\sqrt{\frac{\hat{\rho}\left(1-\hat{\rho}\right)}{N}}\right) = p$$

Interval Estimate of the Packing Coefficient

- X_i - Bernoulli trial $\quad \mathbf{X} = \sum_{i=1}^{N} X_i \quad$ binomially distributed random variable

Expected value: $N\rho$

Variance: $N\rho\left(1-\rho\right)$

- Central limit theorem: $\quad \mathbf{Z} = \dfrac{\sqrt{N}\left(\hat{\rho}-\rho\right)}{\sqrt{\rho\left(1-\rho\right)}} \xrightarrow{F} N(0,1)$

Z converges in distribution to a normally distributed random variable with a mean of 0 and a variance of 1

- Confidence interval: $\quad P\left(\underbrace{\left|\dfrac{(\rho-\hat{\rho})\sqrt{N}}{\sqrt{\hat{\rho}\left(1-\hat{\rho}\right)}}\right|}\geq 2z\right) = 1-\alpha$

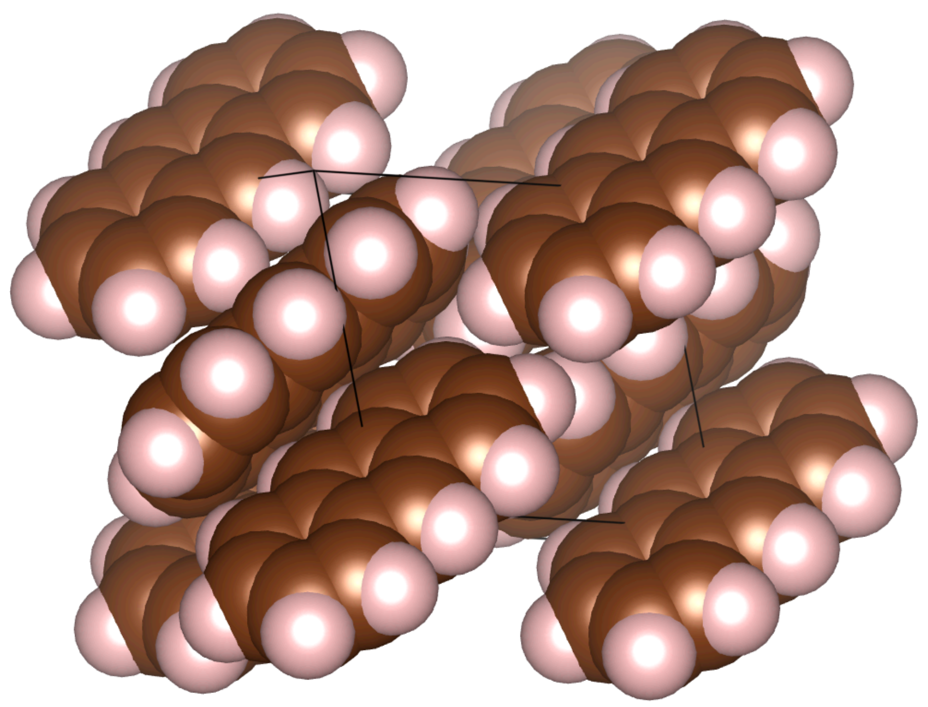
Random quantity is the interval.

- Interval estimate: $\quad \left[\hat{\rho} - \dfrac{\hat{\rho}(1-\hat{\rho})}{\sqrt{N}}z; \hat{\rho} + \dfrac{\hat{\rho}(1-\hat{\rho})}{\sqrt{N}}z\right]$

z **-** $1-\alpha$ **quantile of the standard normal distribution**

- A way to express confidence in the packing coefficient estimate $\hat{\rho}$
  - **The tighter the interval, the higher the confidence in the estimate.**
- Even better, sample until the size of this interval to be less than some constant c with probability p

$$P\left(|\hat{\rho}-\rho| < 2z\sqrt{\frac{\hat{\rho}\left(1-\hat{\rho}\right)}{N}}\right) = p$$



  - $|\hat{\rho}-\rho| < 0.001$
  - $p = 0.999$

Interval Estimate of the Packing Coefficient

- X_i - Bernoulli trial $\mathbf{X} = \sum_{i=1}^{N} X_i$ binomially distributed random variable

Expected value: $N\rho$

Variance: $N\rho\left(1-\rho\right)$

- Central limit theorem: $\mathbf{Z} = \dfrac{\sqrt{N}\left(\hat{\rho}-\rho\right)}{\sqrt{\rho\left(1-\rho\right)}} \xrightarrow{F} N(0,1)$

Z converges in distribution to a normally distributed random variable with a mean of 0 and a variance of 1

- Confidence interval: $P\left(\underbrace{\left|\dfrac{(\rho-\hat{\rho})\sqrt{N}}{\sqrt{\hat{\rho}\left(1-\hat{\rho}\right)}}\right|}\geq 2z\right) = 1-\alpha$

Random quantity is the interval.

- Interval estimate: $\left[\hat{\rho} - \dfrac{\hat{\rho}(1-\hat{\rho})}{\sqrt{N}}z; \hat{\rho} + \dfrac{\hat{\rho}(1-\hat{\rho})}{\sqrt{N}}z\right]$

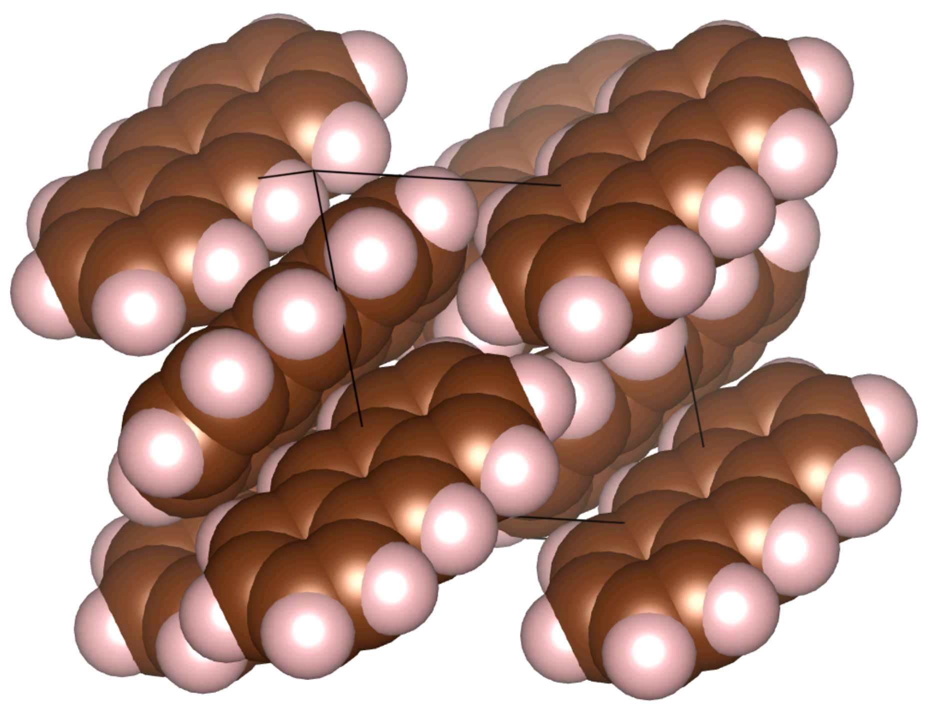
z **-** $1-\alpha$ **quantile of the standard normal distribution**

- A way to express confidence in the packing coefficient estimate $\hat{\rho}$
  - **The tighter the interval, the higher the confidence in the estimate.**
- Even better, sample until the size of this interval to be less than some constant c with probability p

$$P\left(|\hat{\rho}-\rho| < 2z\sqrt{\frac{\hat{\rho}\left(1-\hat{\rho}\right)}{N}}\right) = p$$

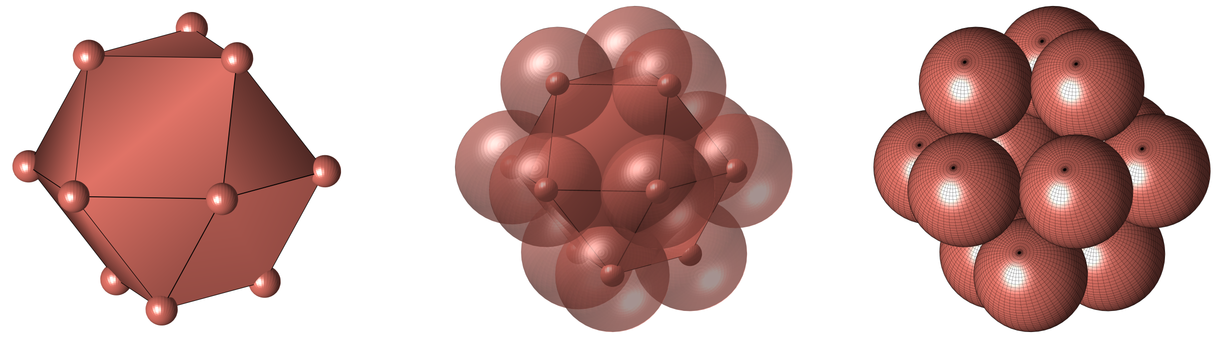
- $|\hat{\rho}-\rho| < 0.001$
- $p = 0.999$

- $\hat{\rho} \approx 0.767827$ N ~ 9,000,000
- N=100,000: 0.76812
- CCDC Mercury: 0.762159



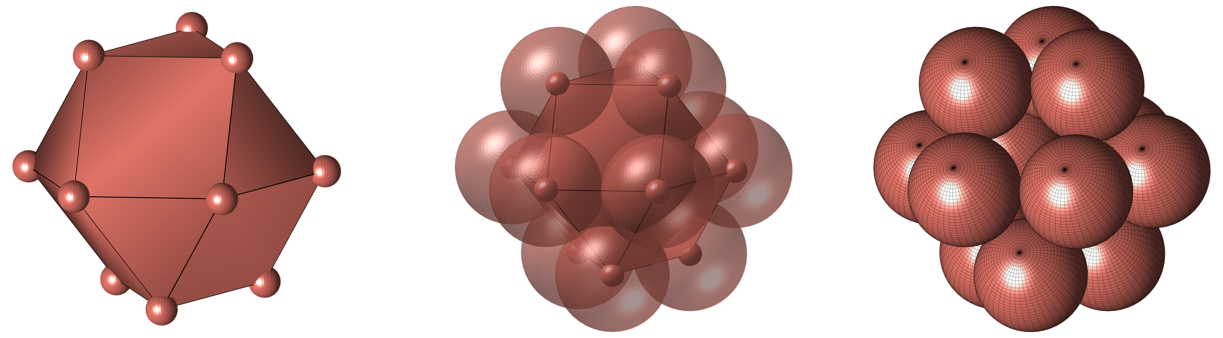
Estimating Packing Coefficient of a Model Molecule

- A collection of equal spheres with centres on the vertices of a cuboctahedron.



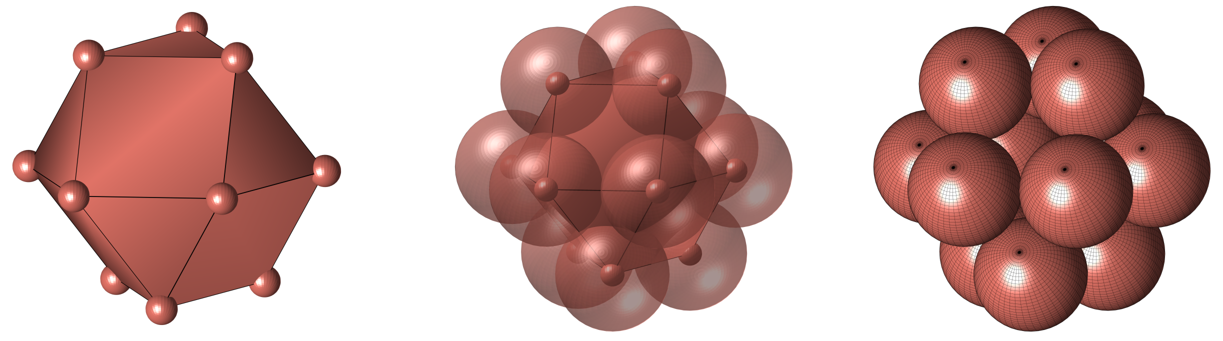
Estimating Packing Coefficient of a Model Molecule

- A collection of equal spheres with centres on the vertices of a cuboctahedron.
    - **Number of vertices: 12**



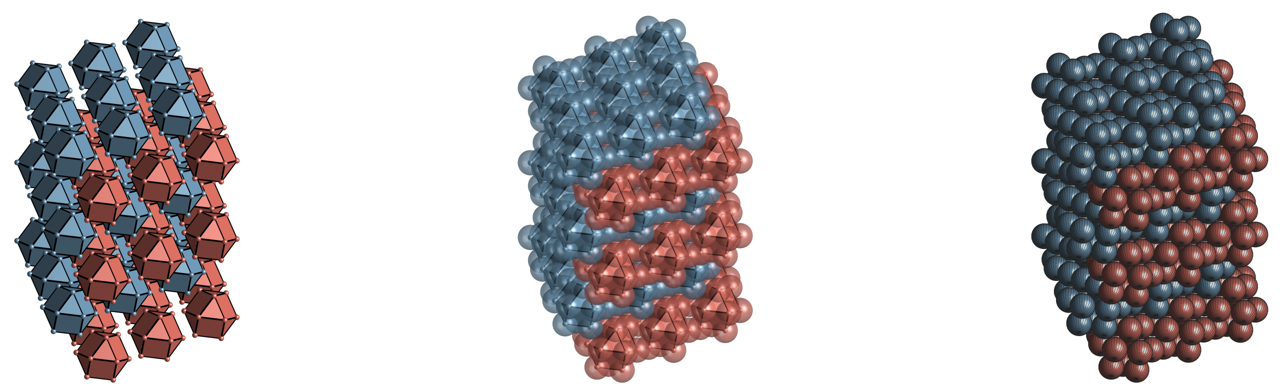
Estimating Packing Coefficient of a Model Molecule

- A collection of equal spheres with centres on the vertices of a cuboctahedron.
    - **Number of vertices: 12**
    - **Radius of the spheres set to the van der Waals radius of carbon.**



- Densest packing of cuboctahedral molecules

$$\rho = \frac{12}{13}\frac{\pi}{\sqrt{18}} \approx 0.68352...$$

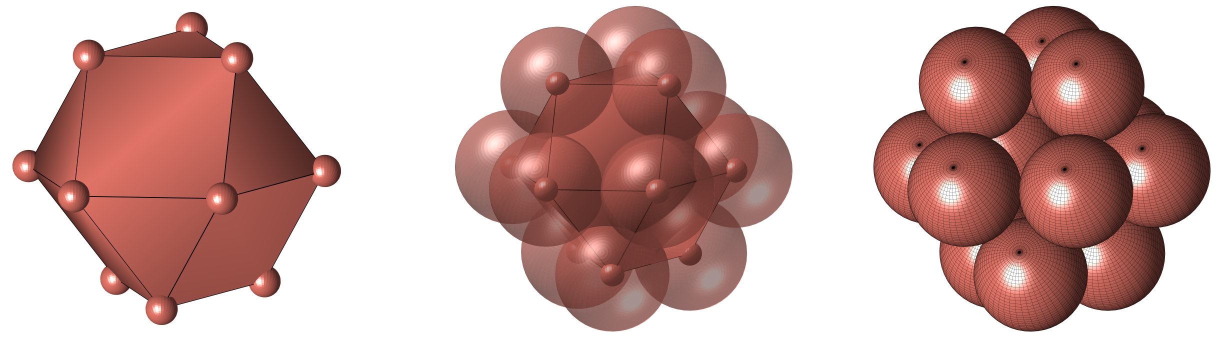


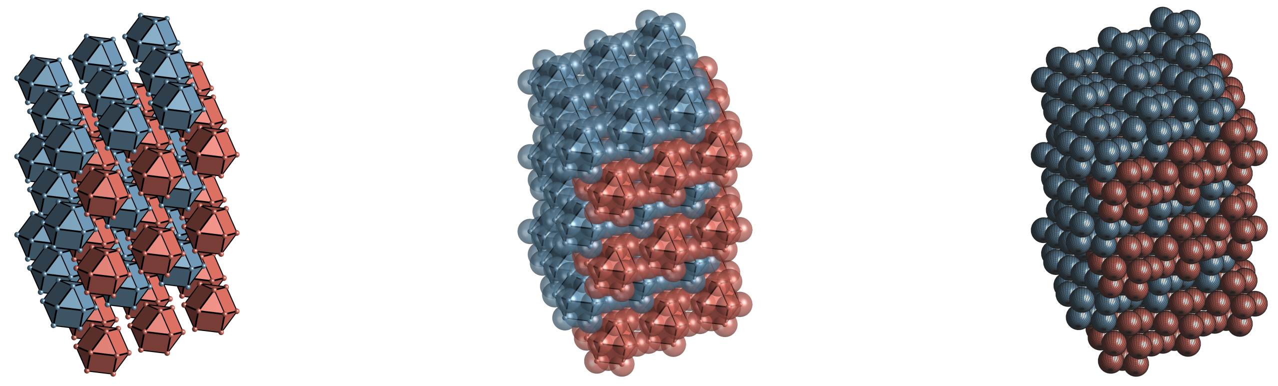
Estimating Packing Coefficient of a Model Molecule

- A collection of equal spheres with centres on the vertices of a cuboctahedron.
    - **Number of vertices: 12**
    - **Radius of the spheres set to the van der Waals radius of carbon.**

- Densest packing of cuboctahedral molecules $$\rho = \frac{12}{13}\frac{\pi}{\sqrt{18}} \approx 0.68352...$$

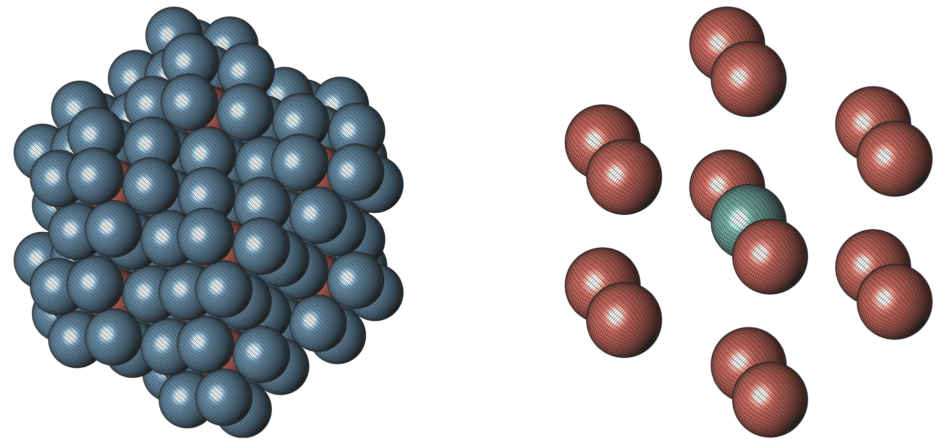
- Complementary packing $$\rho = \frac{1}{13}\frac{\pi}{\sqrt{18}} \approx 0.05696...$$





Packing of 15 cuboctahedral molecules

Complementary packing



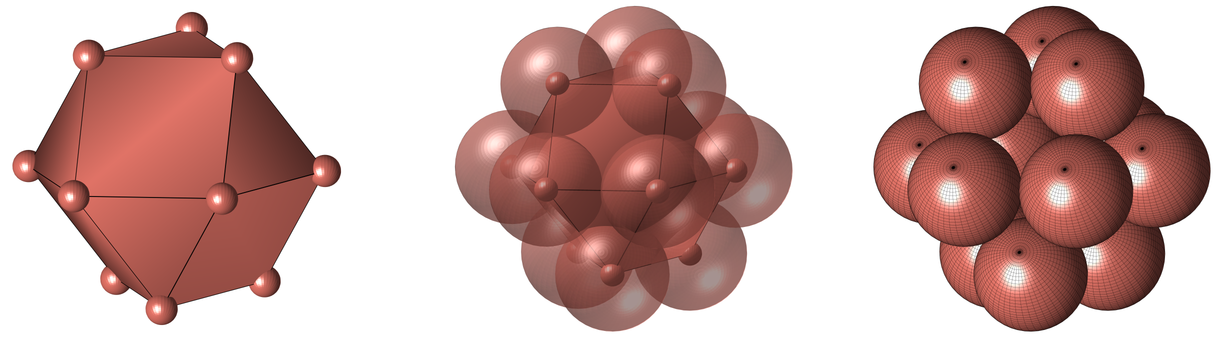
Estimating Packing Coefficient of a Model Molecule

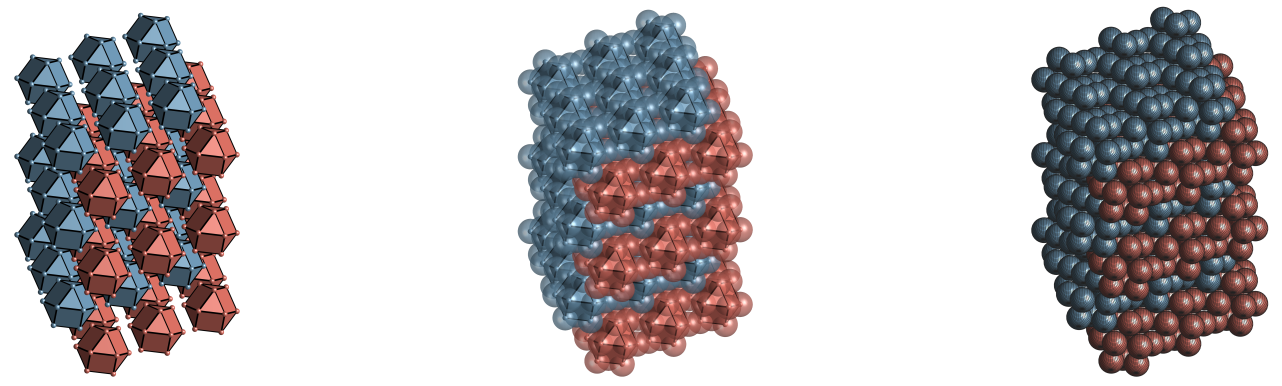
- A collection of equal spheres with centres on the vertices of a cuboctahedron.
    - **Number of vertices: 12**
    - **Radius of the spheres set to the van der Waals radius of carbon.**

- Densest packing of cuboctahedral molecules $$\rho = \frac{12}{13}\frac{\pi}{\sqrt{18}} \approx 0.68352...$$

- Complementary packing $$\rho = \frac{1}{13}\frac{\pi}{\sqrt{18}} \approx 0.05696...$$

- Face-Center Cubic (FCC) close-packing of equal spheres $$\rho = \frac{\pi}{\sqrt{18}} \approx 0.74048...$$

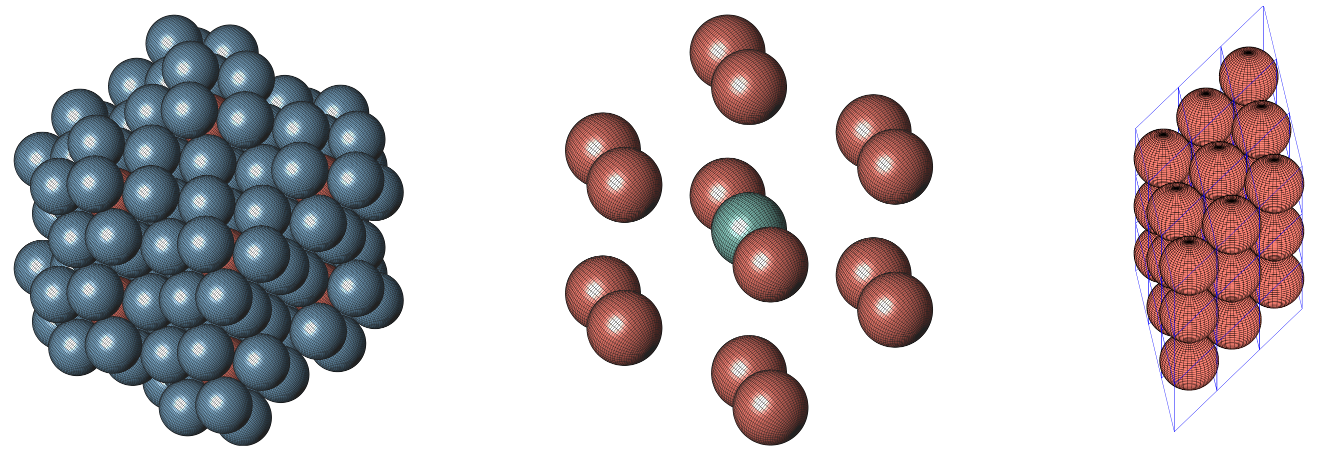




Packing of 15 cuboctahedral molecules

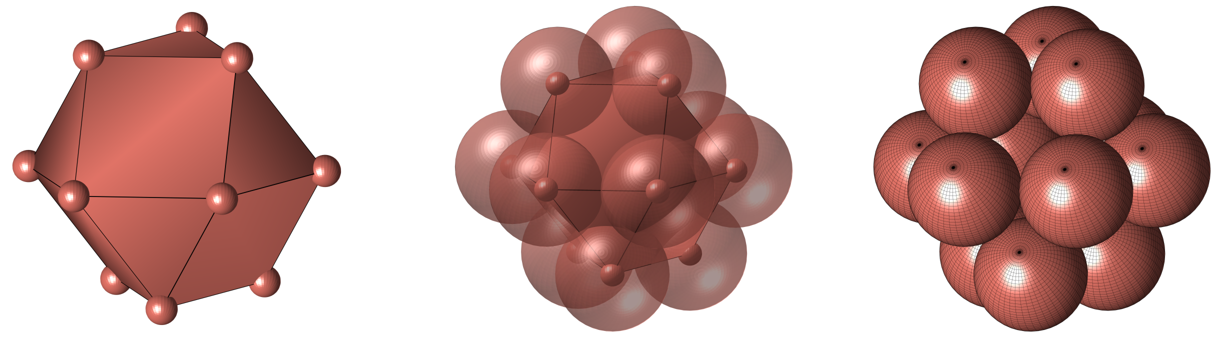
Complementary packing

FCC close-packing of equal spheres

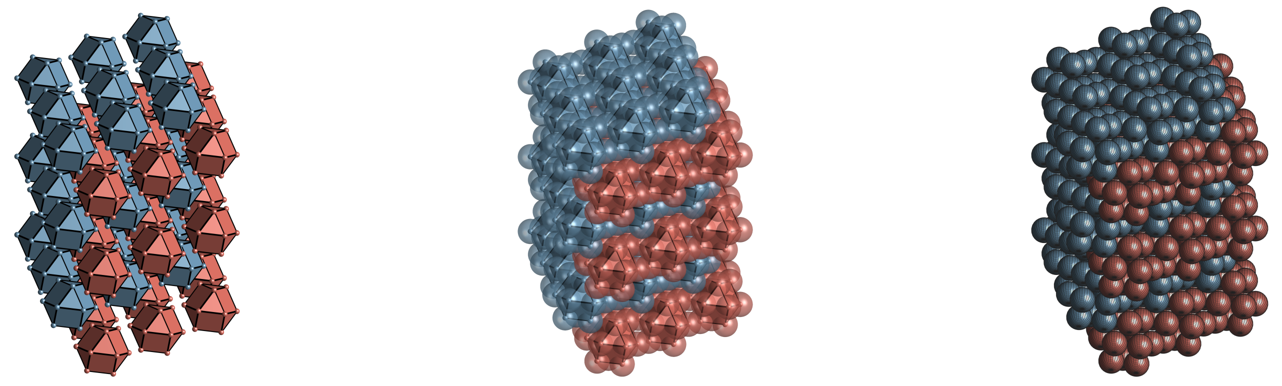


Estimating Packing Coefficient of a Model Molecule

- A collection of equal spheres with centres on the vertices of a cuboctahedron.
    - **Number of vertices: 12**
    - **Radius of the spheres set to the van der Waals radius of carbon.**



- Densest packing of cuboctahedral molecules $$\rho = \frac{12}{13}\frac{\pi}{\sqrt{18}} \approx 0.68352...$$

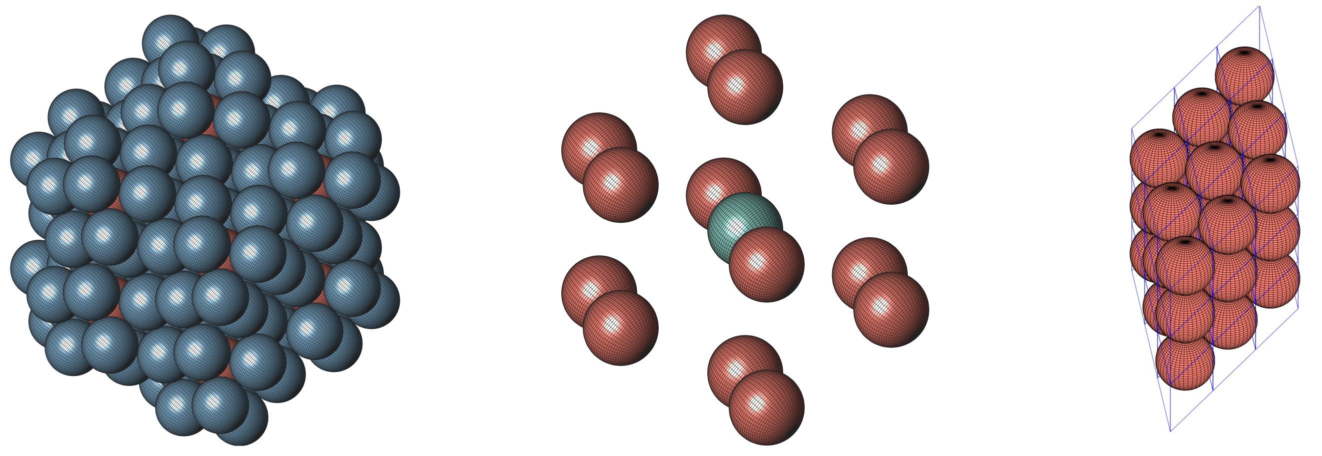


- Complementary packing $$\rho = \frac{1}{13}\frac{\pi}{\sqrt{18}} \approx 0.05696...$$

- Face-Center Cubic (FCC) close-packing of equal spheres $$\rho = \frac{\pi}{\sqrt{18}} \approx 0.74048...$$

Packing of 15 cuboctahedral molecules | Complementary packing | FCC close-packing of equal spheres

- **Conjectured in 1611 by Johannes Kepler to be densest possible packing in his essey *'The Six-Cornered Snowflake'***



Estimating Packing Coefficient of a Model Molecule

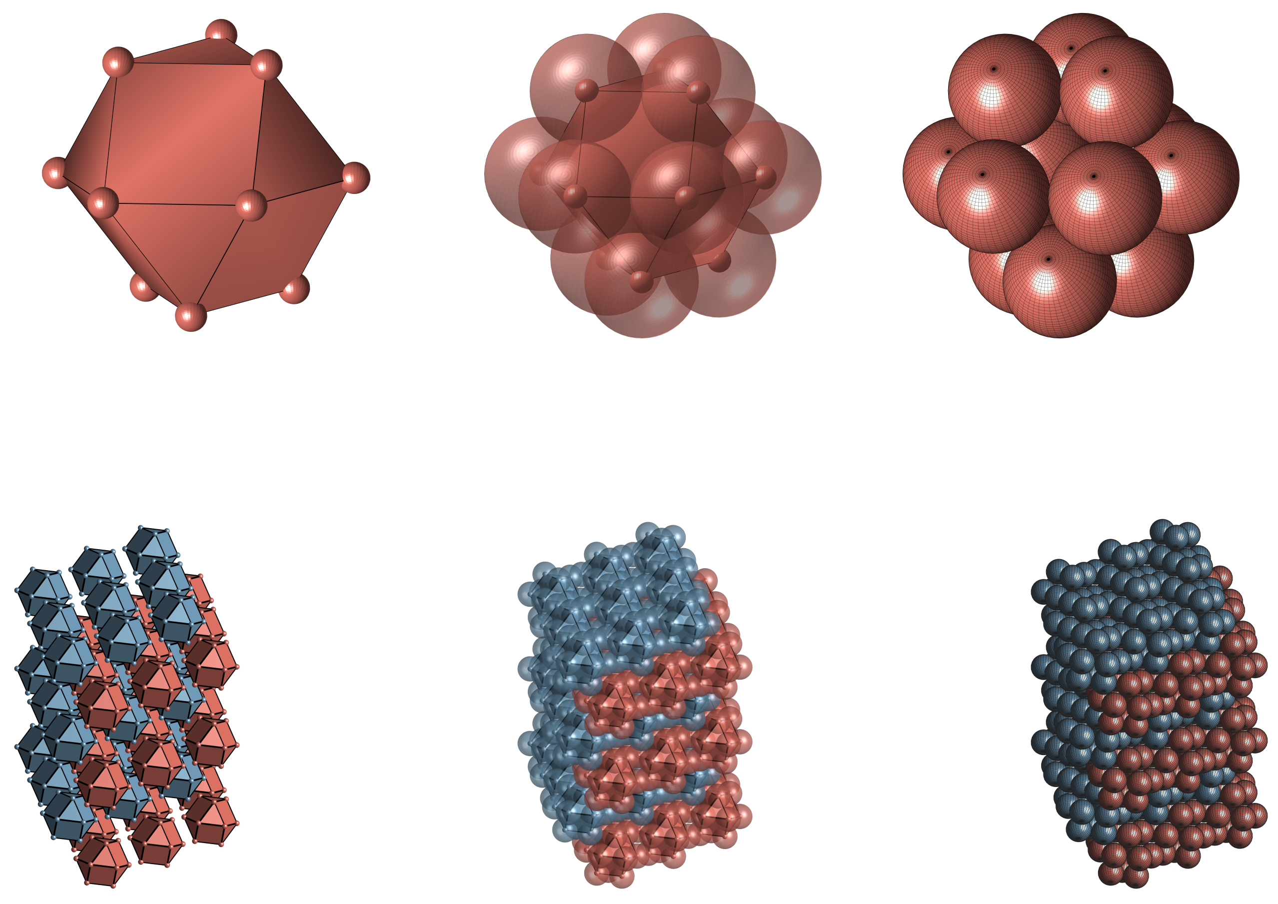
- A collection of equal spheres with centres on the vertices of a cuboctahedron.
    - **Number of vertices: 12**
    - **Radius of the spheres set to the van der Waals radius of carbon.**

- Densest packing of cuboctahedral molecules $\rho = \frac{12}{13}\frac{\pi}{\sqrt{18}} \approx 0.68352...$

- Complementary packing $\rho = \frac{1}{13}\frac{\pi}{\sqrt{18}} \approx 0.05696...$

- Face-Center Cubic (FCC) close-packing of equal spheres $\rho = \frac{\pi}{\sqrt{18}} \approx 0.74048...$

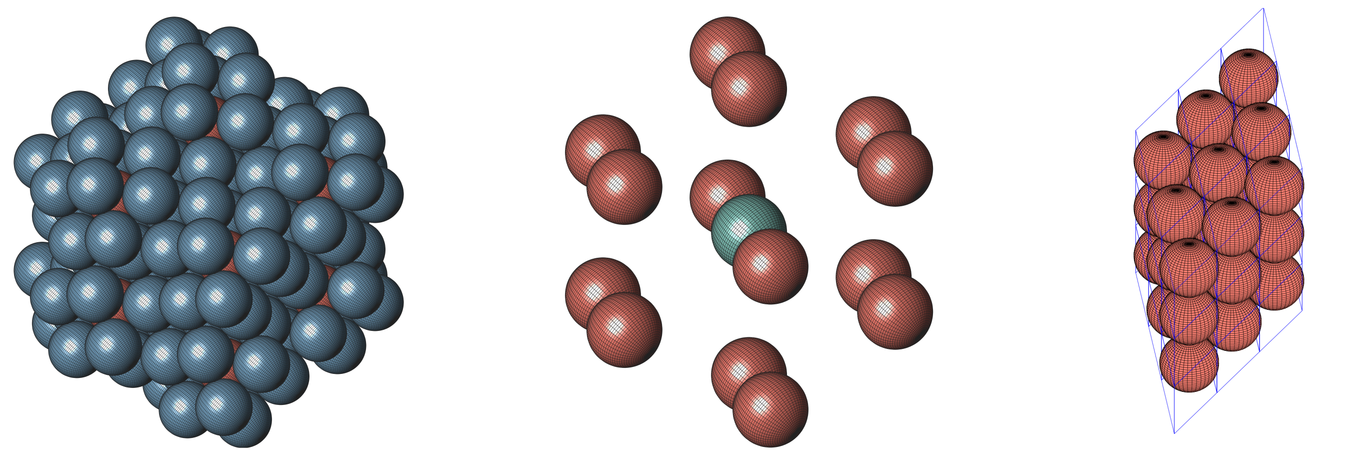
- **Conjectured in 1611 by Johannes Kepler to be densest possible packing in his essey *'The Six-Cornered Snowflake'***
- **Proved to be correct by Thomas Hales in 2005**



Packing of 15 cuboctahedral molecules

Complementary packing

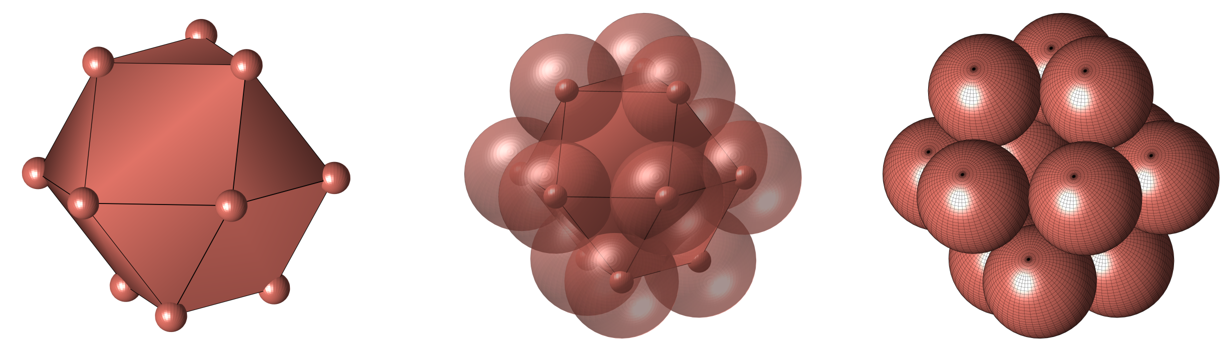
FCC close-packing of equal spheres



Hales, T. C. (2005). A proof of the Kepler conjecture. *Annals of mathematics*, 1065-1185.

Estimating Packing Coefficient of a Model Molecule

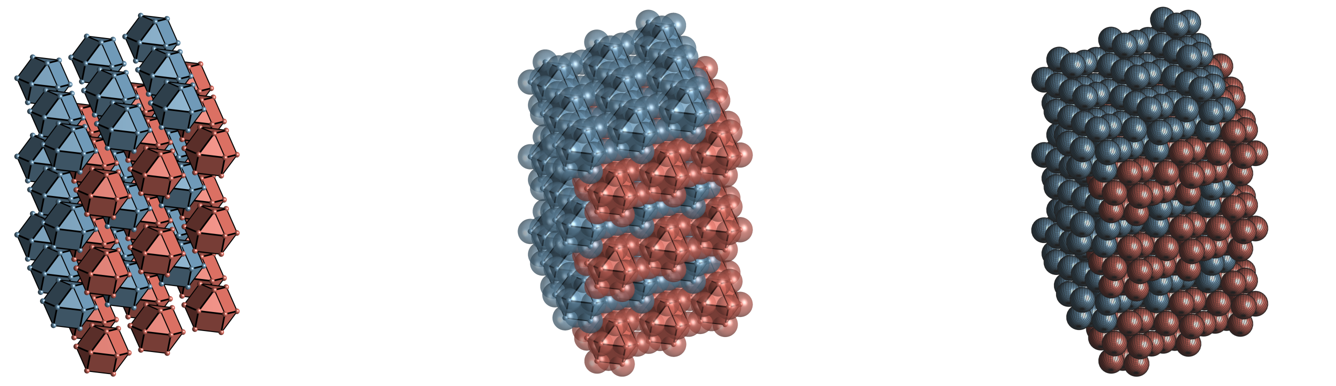
- A collection of equal spheres with centres on the vertices of a cuboctahedron.
    - **Number of vertices: 12**
    - **Radius of the spheres set to the van der Waals radius of carbon.**



- Densest packing of cuboctahedral molecules $\rho = \frac{12}{13}\frac{\pi}{\sqrt{18}} \approx 0.68352...$

- Complementary packing $\rho = \frac{1}{13}\frac{\pi}{\sqrt{18}} \approx 0.05696...$

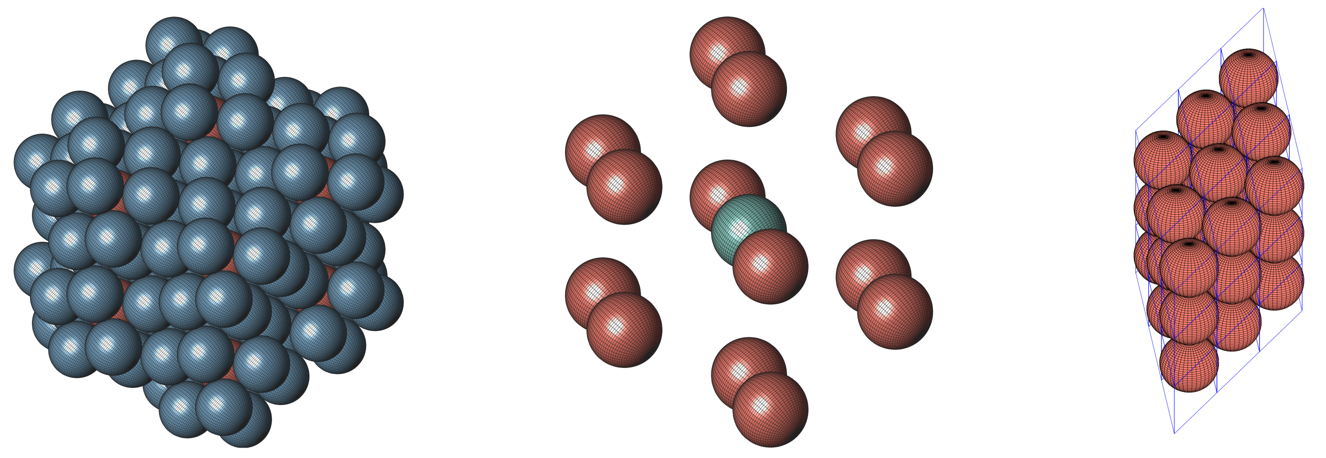
- Face-Center Cubic (FCC) close-packing of equal spheres $\rho = \frac{\pi}{\sqrt{18}} \approx 0.74048...$
    - **Conjectured in 1611 by Johannes Kepler to be densest possible packing in his essey *'The Six-Cornered Snowflake'***
    - **Proved to be correct by Thomas Hales in 2005**



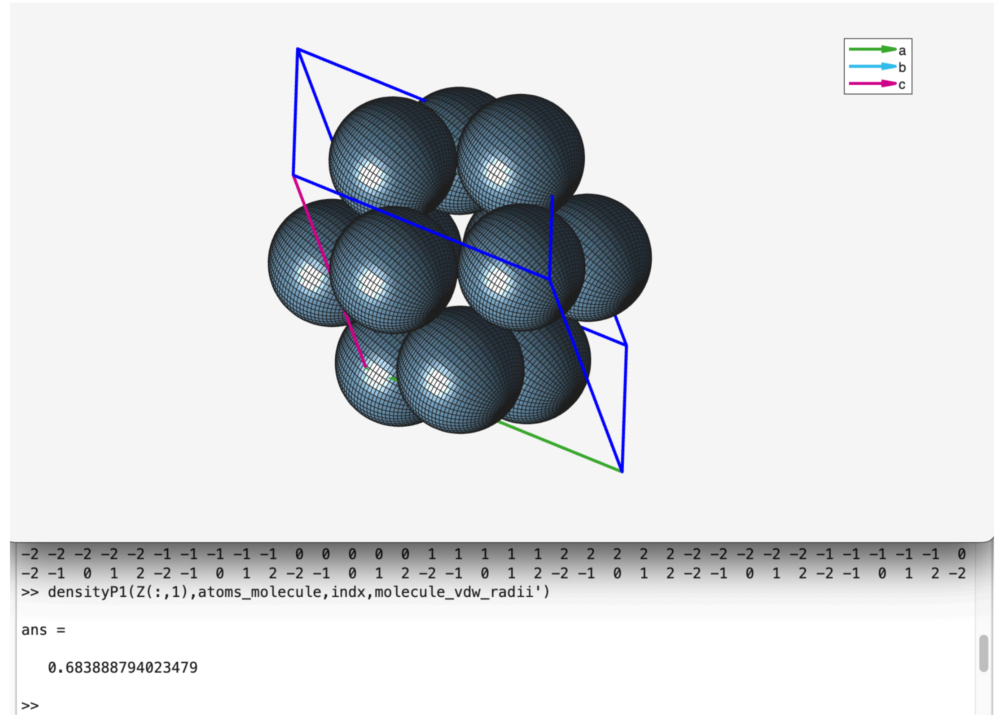
Packing of 15 cuboctahedral molecules

Complementary packing

FCC close-packing of equal spheres



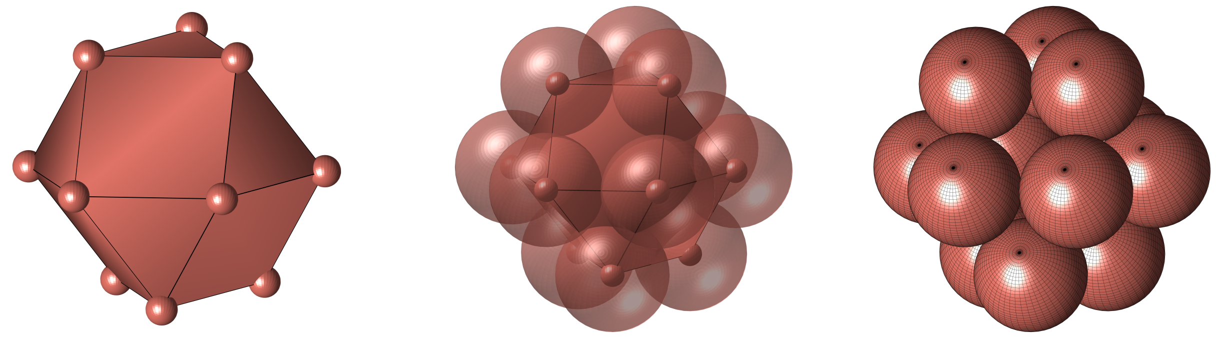
- $\hat{\rho} \approx 0.683888$ $\epsilon = 0.00036$



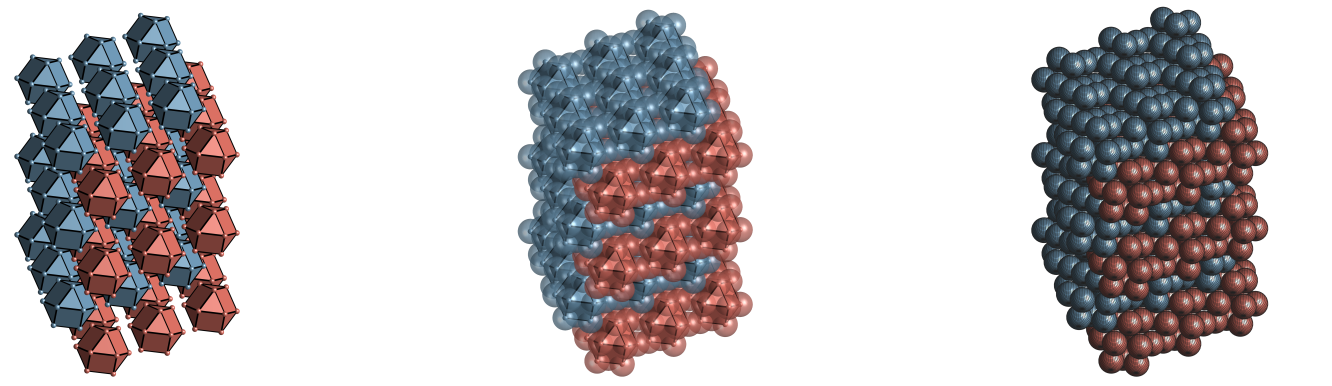
Hales, T. C. (2005). A proof of the Kepler conjecture. *Annals of mathematics*, 1065-1185.

Estimating Packing Coefficient of a Model Molecule

- A collection of equal spheres with centres on the vertices of a cuboctahedron.
    - **Number of vertices: 12**
    - **Radius of the spheres set to the van der Waals radius of carbon.**



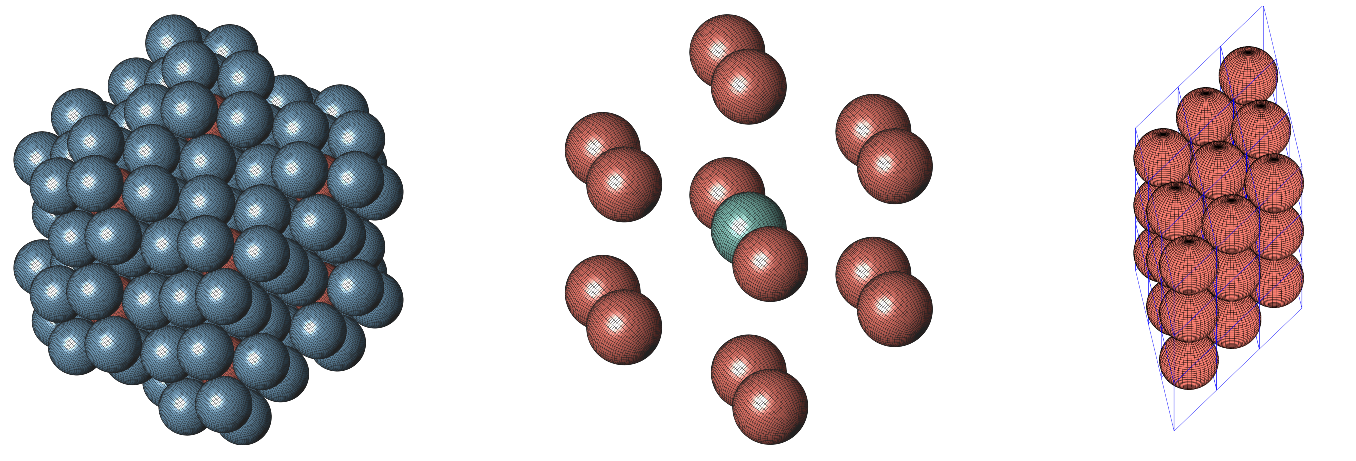
- Densest packing of cuboctahedral molecules $\rho = \frac{12}{13}\frac{\pi}{\sqrt{18}} \approx 0.68352...$
- Complementary packing $\rho = \frac{1}{13}\frac{\pi}{\sqrt{18}} \approx 0.05696...$
- Face-Center Cubic (FCC) close-packing of equal spheres $\rho = \frac{\pi}{\sqrt{18}} \approx 0.74048...$
    - **Conjectured in 1611 by Johannes Kepler to be densest possible packing in his essey *'The Six-Cornered Snowflake'***
    - **Proved to be correct by Thomas Hales in 2005**



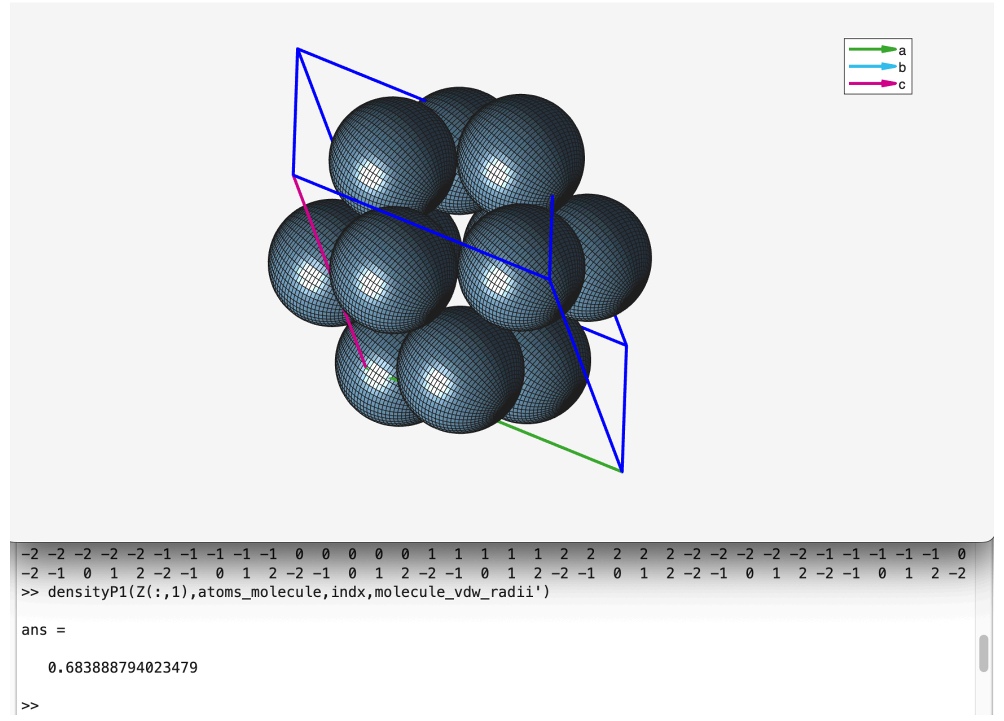
Packing of 15 cuboctahedral molecules

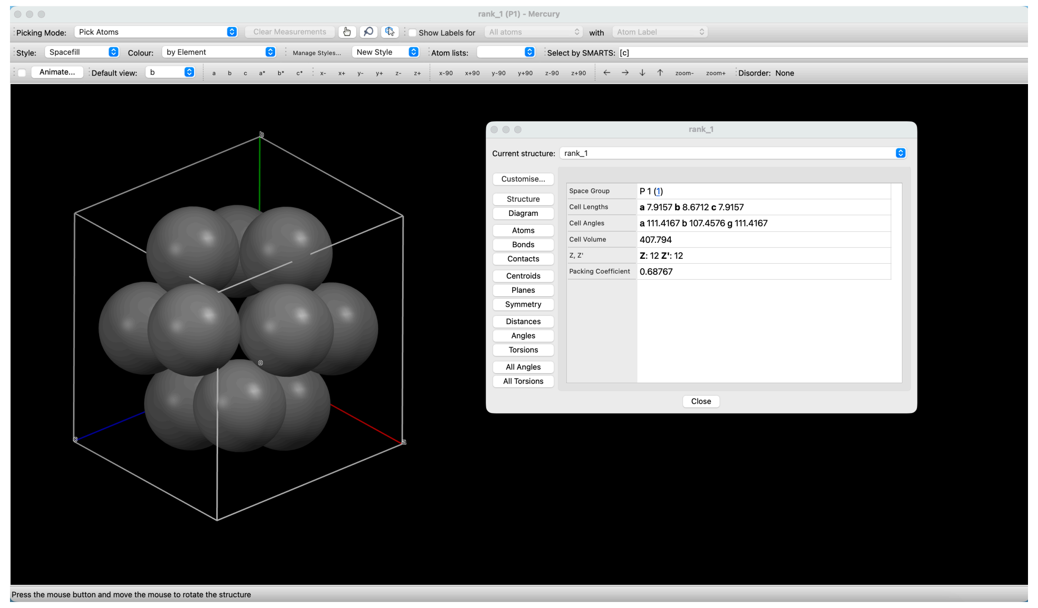
Complementary packing

FCC close-packing of equal spheres



- $\hat{\rho} \approx 0.683888$ $\qquad \epsilon = 0.00036$
- CCDC Mercury: 0.68767 $\qquad \epsilon = 0.00415$

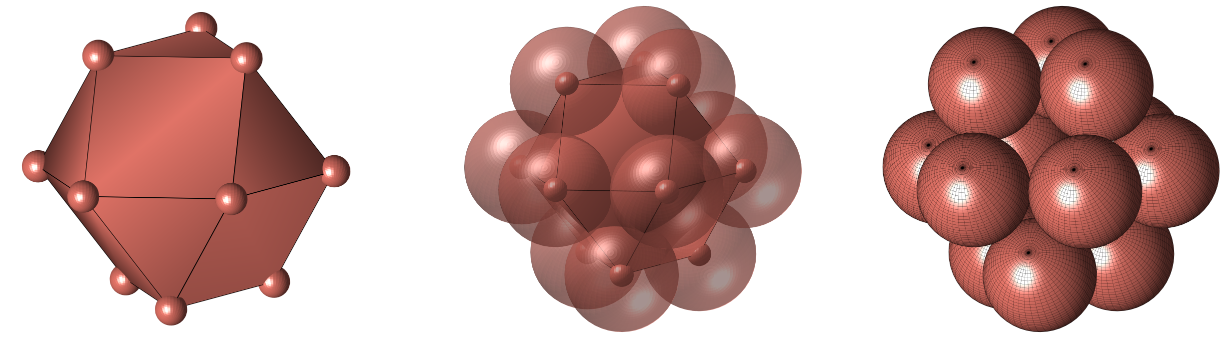




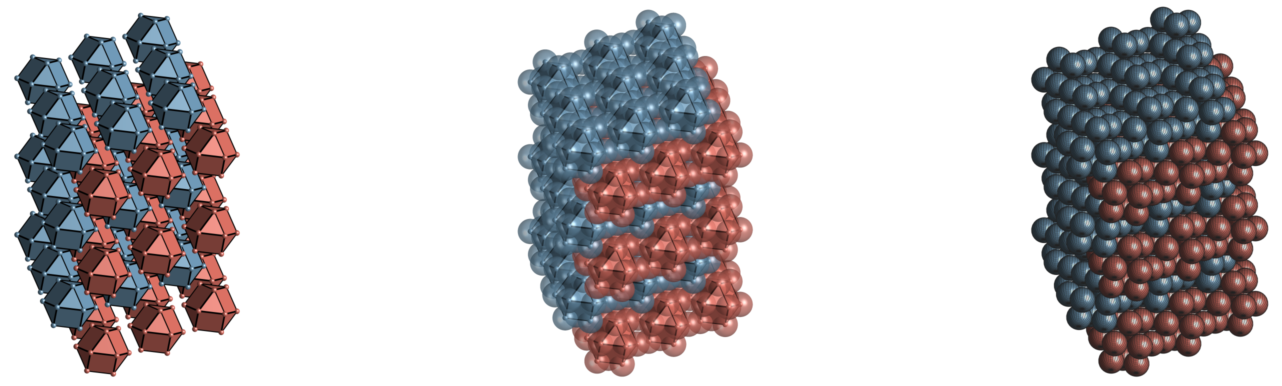
Hales, T. C. (2005). A proof of the Kepler conjecture. *Annals of mathematics*, 1065-1185.

Estimating Packing Coefficient of a Model Molecule

- A collection of equal spheres with centres on the vertices of a cuboctahedron.
    - **Number of vertices: 12**
    - **Radius of the spheres set to the van der Waals radius of carbon.**



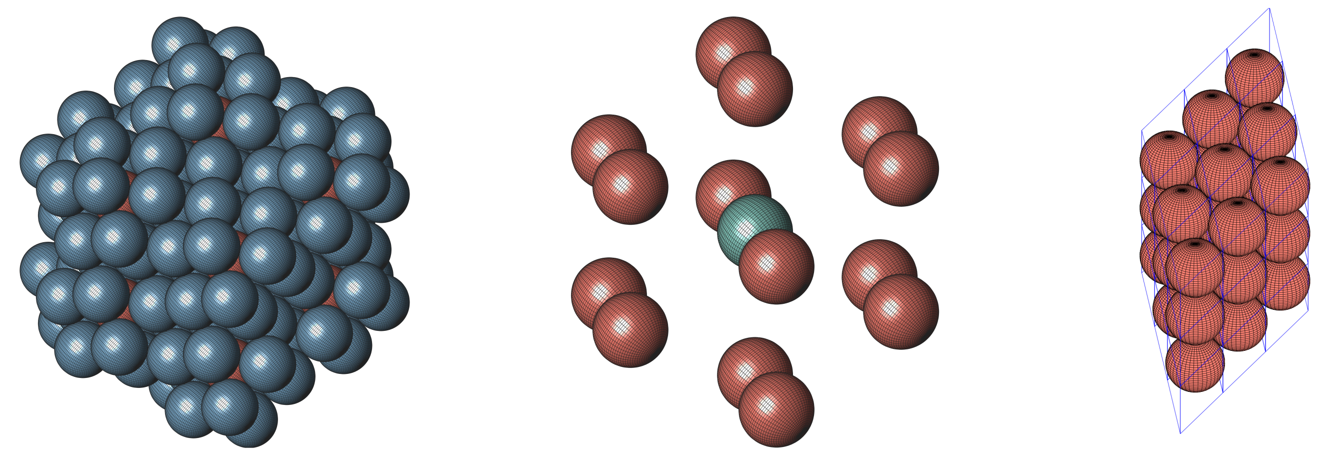
- Densest packing of cuboctahedral molecules $\rho = \frac{12}{13}\frac{\pi}{\sqrt{18}} \approx 0.68352...$
- Complementary packing $\rho = \frac{1}{13}\frac{\pi}{\sqrt{18}} \approx 0.05696...$
- Face-Center Cubic (FCC) close-packing of equal spheres $\rho = \frac{\pi}{\sqrt{18}} \approx 0.74048...$
    - **Conjectured in 1611 by Johannes Kepler to be densest possible packing in his essey *'The Six-Cornered Snowflake'***
    - **Proved to be correct by Thomas Hales in 2005**



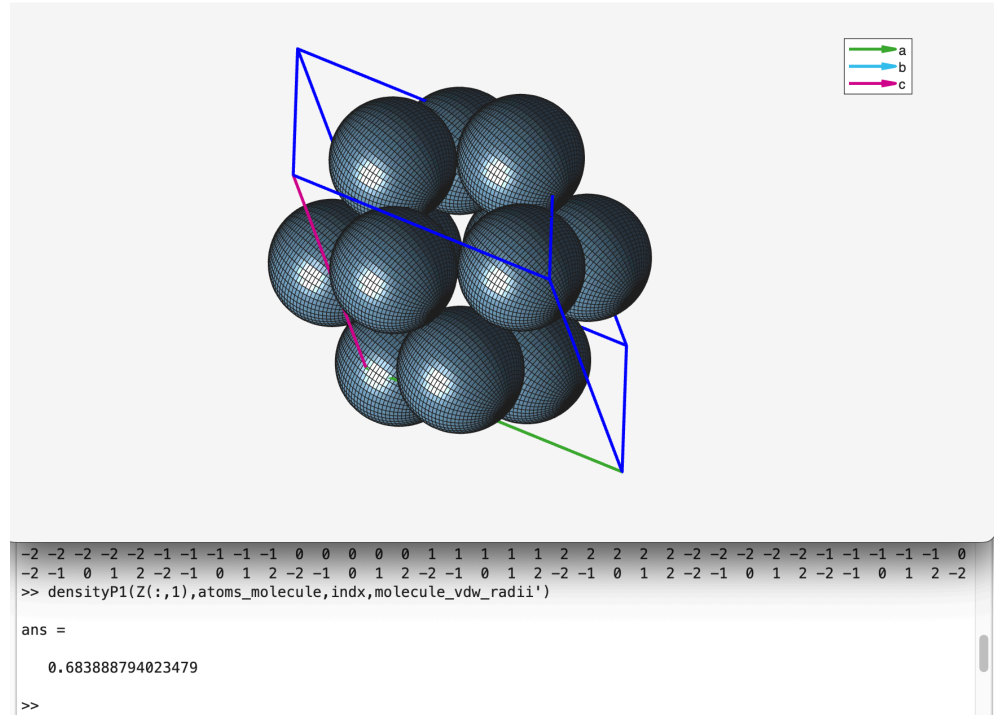
Packing of 15 cuboctahedral molecules

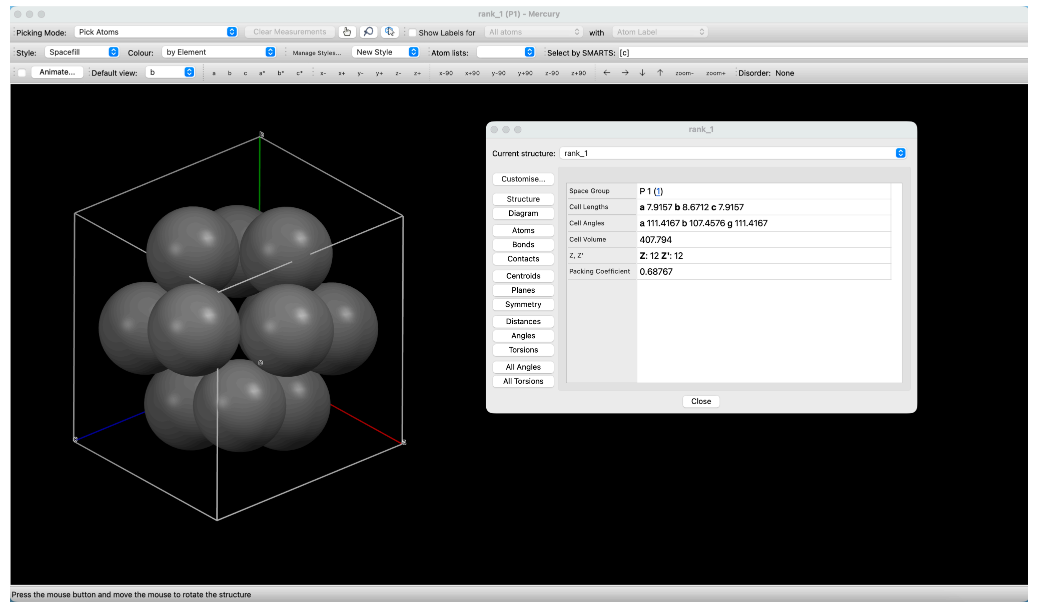
Complementary packing

FCC close-packing of equal spheres



- $\hat{\rho} \approx 0.683888$ $\epsilon = 0.00036$ $N = 9{,}370{,}000$
- CCDC Mercury: 0.68767 $\epsilon = 0.00415$





Hales, T. C. (2005). A proof of the Kepler conjecture. *Annals of mathematics*, 1065-1185.

Optimisation Under Uncertainty

- Stochastic optimisation problems

Optimisation Under Uncertainty

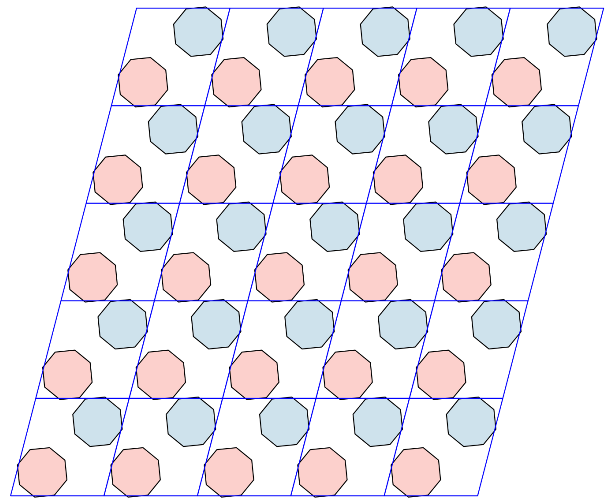
- Stochastic optimisation problems
    - **Uses the tools of probability theory and mathematical statistics.**

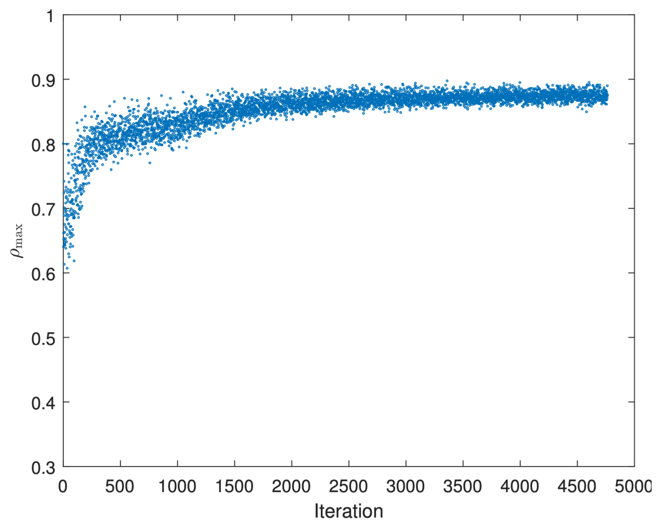
Optimisation Under Uncertainty

- Stochastic optimisation problems
    - **Uses the tools of probability theory and mathematical statistics.**
    - **Situations where the optimum is unknown.**

Optimisation Under Uncertainty

- Stochastic optimisation problems
    - **Uses the tools of probability theory and mathematical statistics.**
    - **Situations where the optimum is unknown.**

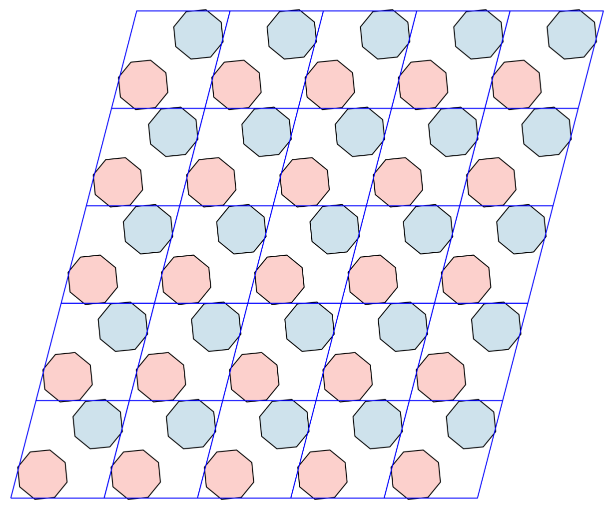


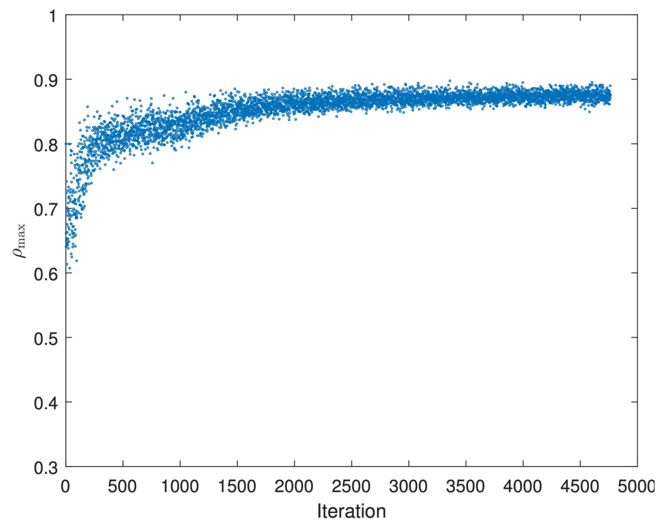


- **Example**: Packing regular octagons in the plane group p2.

Optimisation Under Uncertainty

- Stochastic optimisation problems
  - **Uses the tools of probability theory and mathematical statistics.**
  - **Situations where the optimum is unknown.**



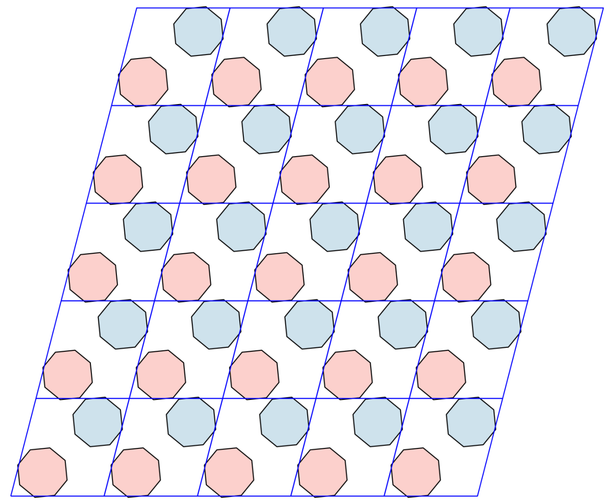


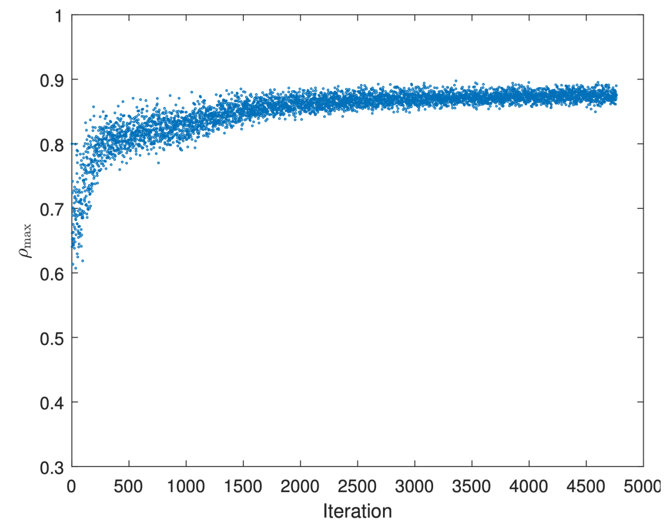
- **Example**: Packing regular octagons in the plane group p2.

- How to express confidence in the best solution found?

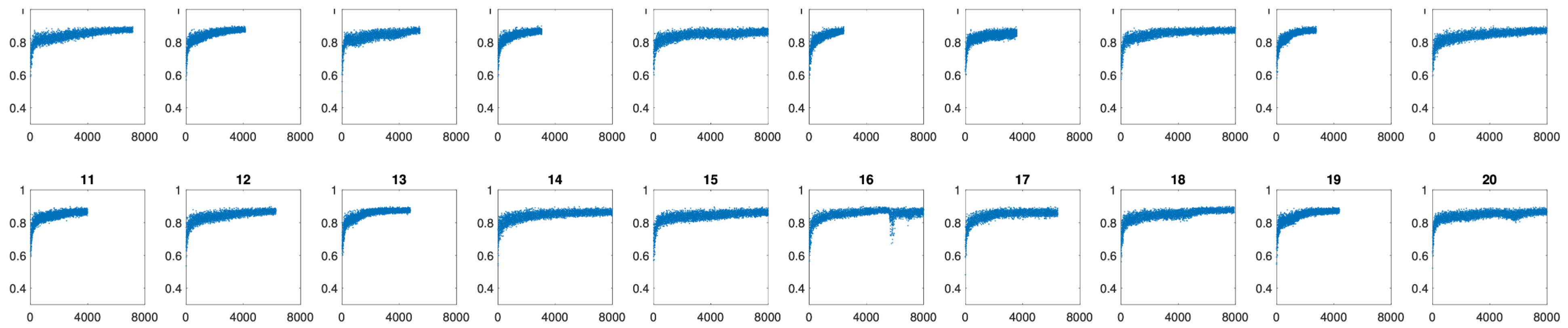
Optimisation Under Uncertainty

- Stochastic optimisation problems
    - **Uses the tools of probability theory and mathematical statistics.**
    - **Situations where the optimum is unknown.**





- **Example**: Packing regular octagons in the plane group p2.

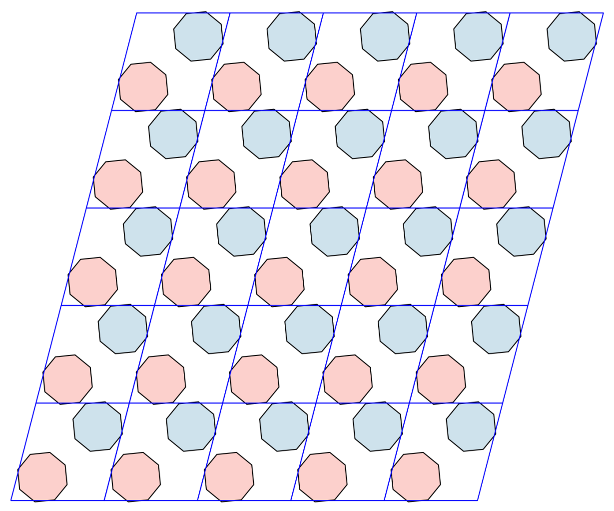


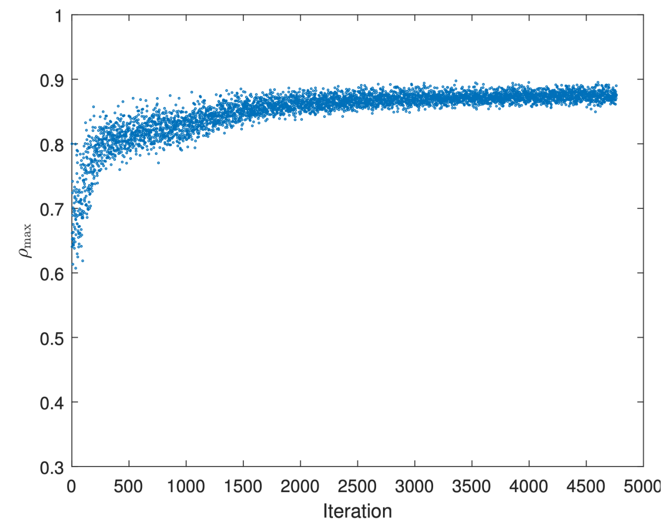
- How to express confidence in the best solution found?

- Run multiple simulations and compute an interval estimate.

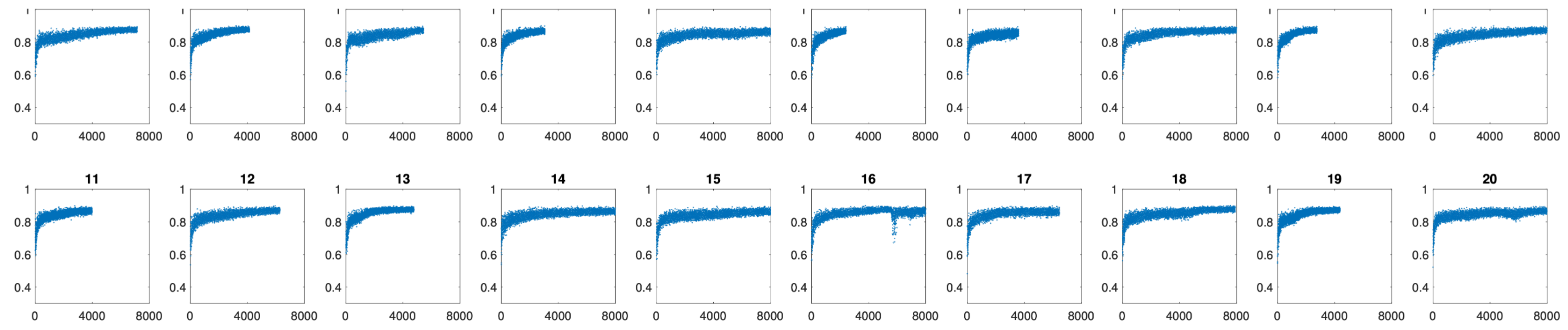
Optimisation Under Uncertainty

- Stochastic optimisation problems
  - **Uses the tools of probability theory and mathematical statistics.**
  - **Situations where the optimum is unknown.**



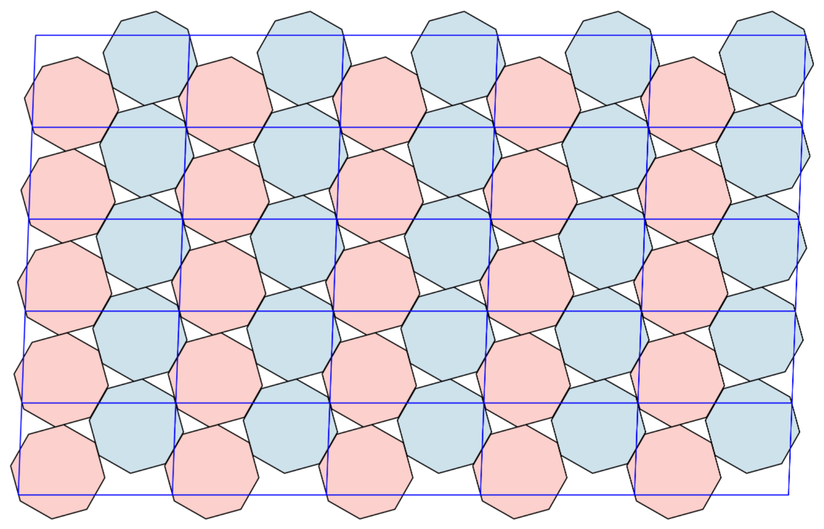


- **Example**: Packing regular octagons in the plane group p2.



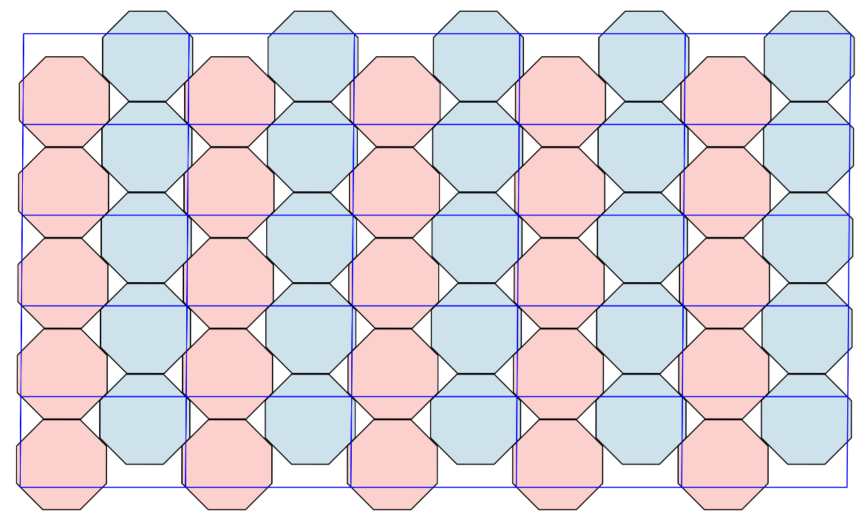
- How to express confidence in the best solution found?
- Run multiple simulations and compute an interval estimate.
- Hodges-Lehmann estimator of the pseudomedian of 20 runs

$$\hat{m} = 0.8970032$$



Highest maximal packing solution

$\rho = 0.89984$

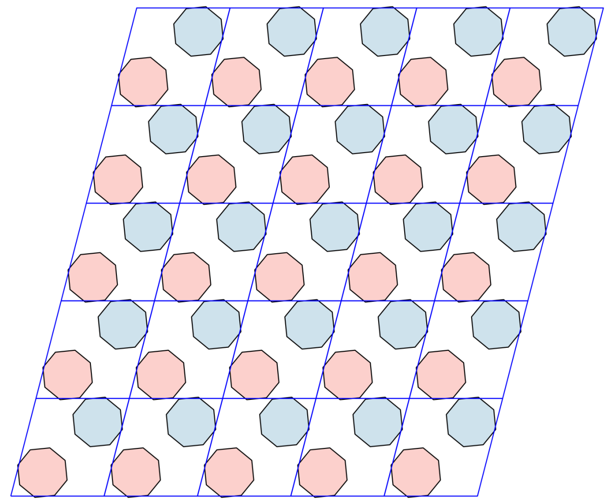


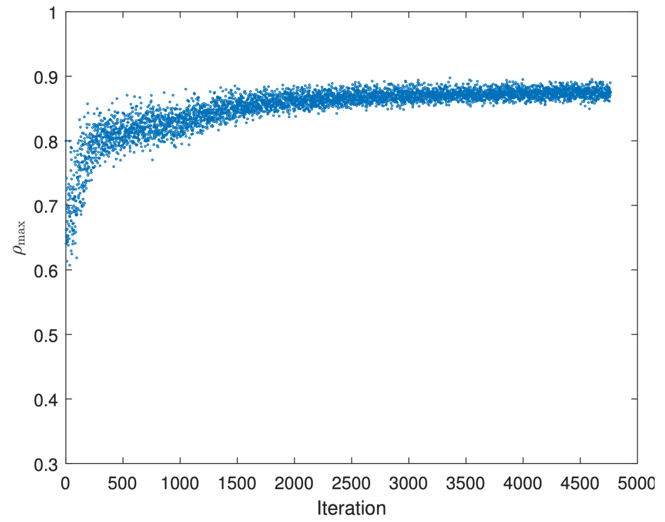
Lowest maximal packing solution

$\rho = 0.89336$

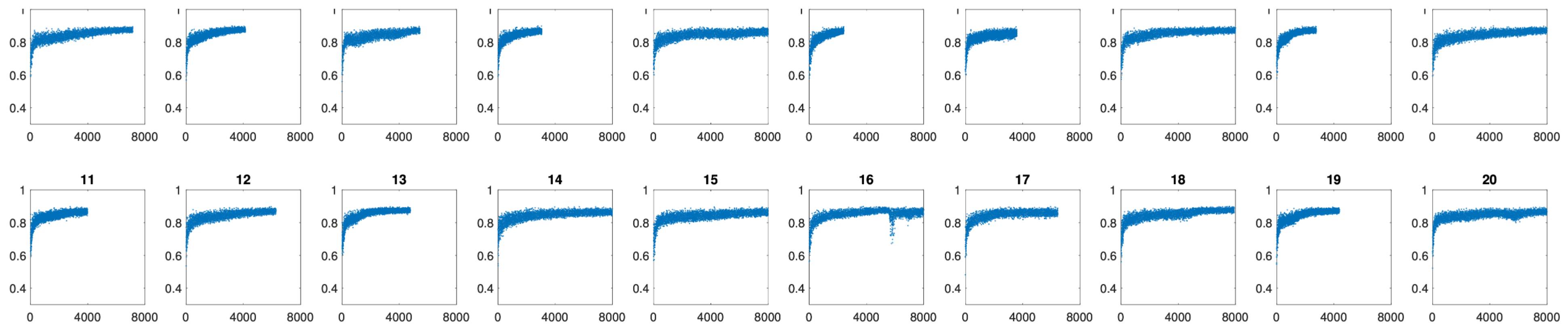
Optimisation Under Uncertainty

- Stochastic optimisation problems
    - **Uses the tools of probability theory and mathematical statistics.**
    - **Situations where the optimum is unknown.**





- **Example**: Packing regular octagons in the plane group p2.

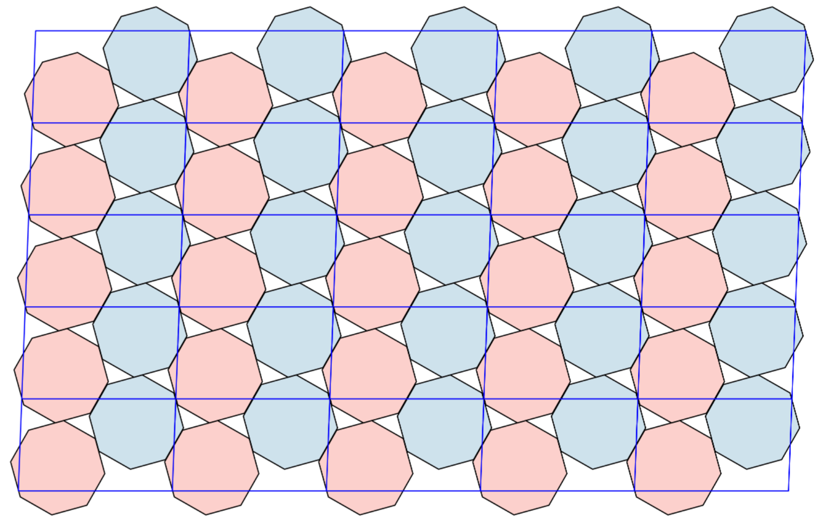


- How to express confidence in the best solution found?
- Run multiple simulations and compute an interval estimate.
- Hodges-Lehmann estimator of the pseudomedian of 20 runs

$$\hat{m} = 0.8970032$$

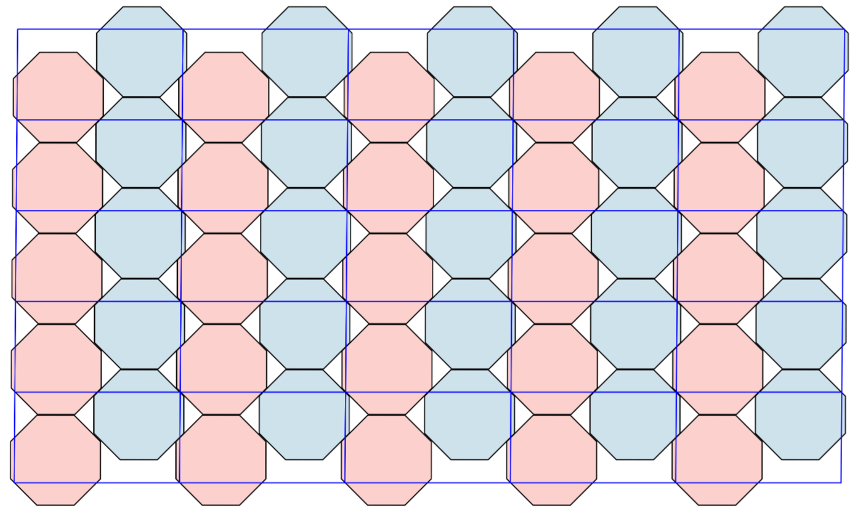
- 95% confidence interval based on Wilcoxon's signed rank test

$$(0.8959246, 0.8980686)$$



Highest maximal packing solution

$\rho = 0.89984$



Lowest maximal packing solution

$\rho = 0.89336$

- The universal phrase in global optimisation literature:

“Repeat until convergence.”

Optimisation Under Uncertainty

- The universal phrase in global optimisation literature:

"Repeat until convergence."

- Simulated annealing example

Optimisation Under Uncertainty

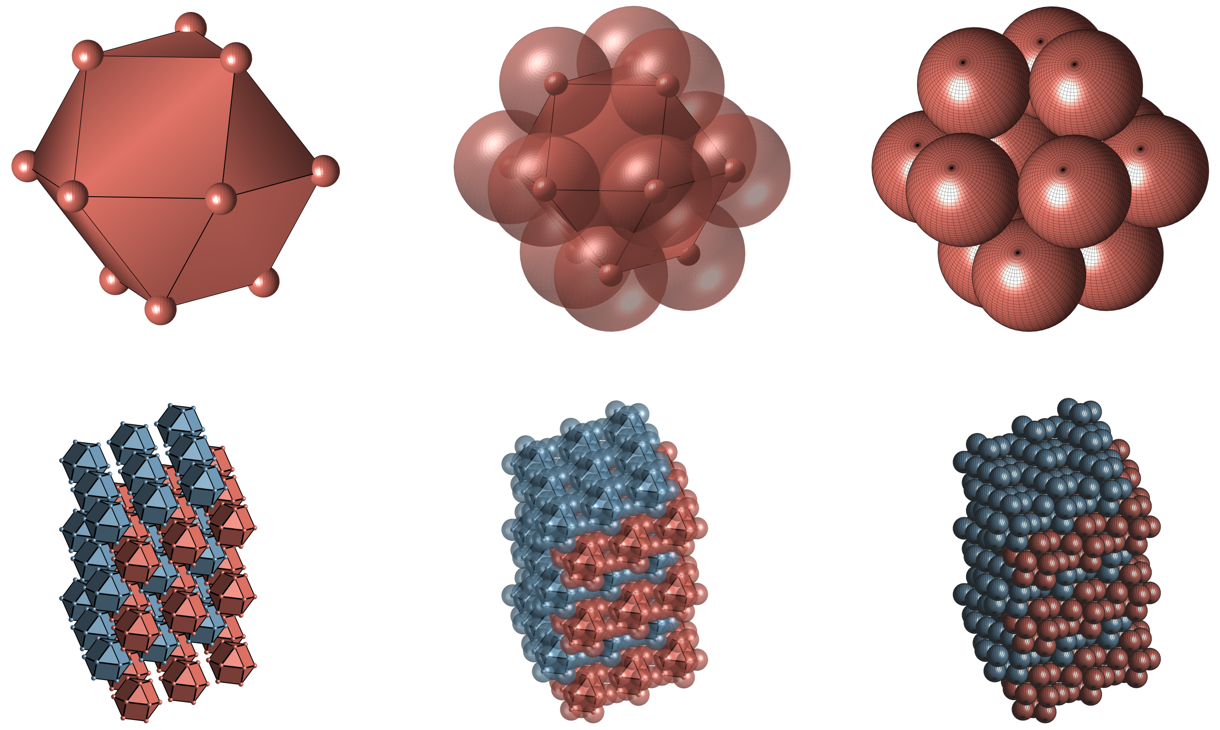
- The universal phrase in global optimisation literature:

“Repeat until convergence.”

- Simulated annealing example
    - **Converge in probability to the set of globally optimal solutions**

$$\lim_{k\to\infty} P(X_k \in S_{\min}) = 1$$

Optimisation Under Uncertainty



- The universal phrase in global optimisation literature:

  "Repeat until convergence."

- Simulated annealing example

  - **Converge in probability to the set of globally optimal solutions**

$$\lim_{k\to\infty} P(X_k \in S_{\min}) = 1$$

  - **We don't know when it happens. It might take years.**

Optimisation Under Uncertainty

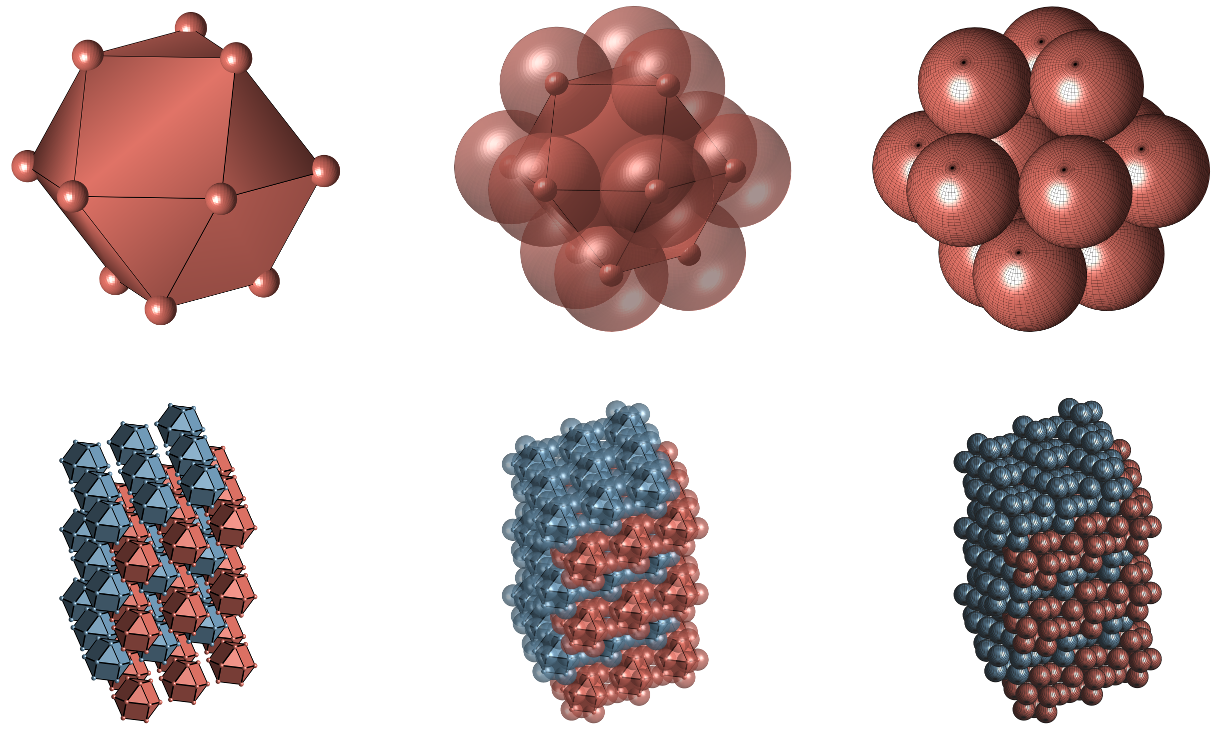
- The universal phrase in global optimisation literature:

  "Repeat until convergence."

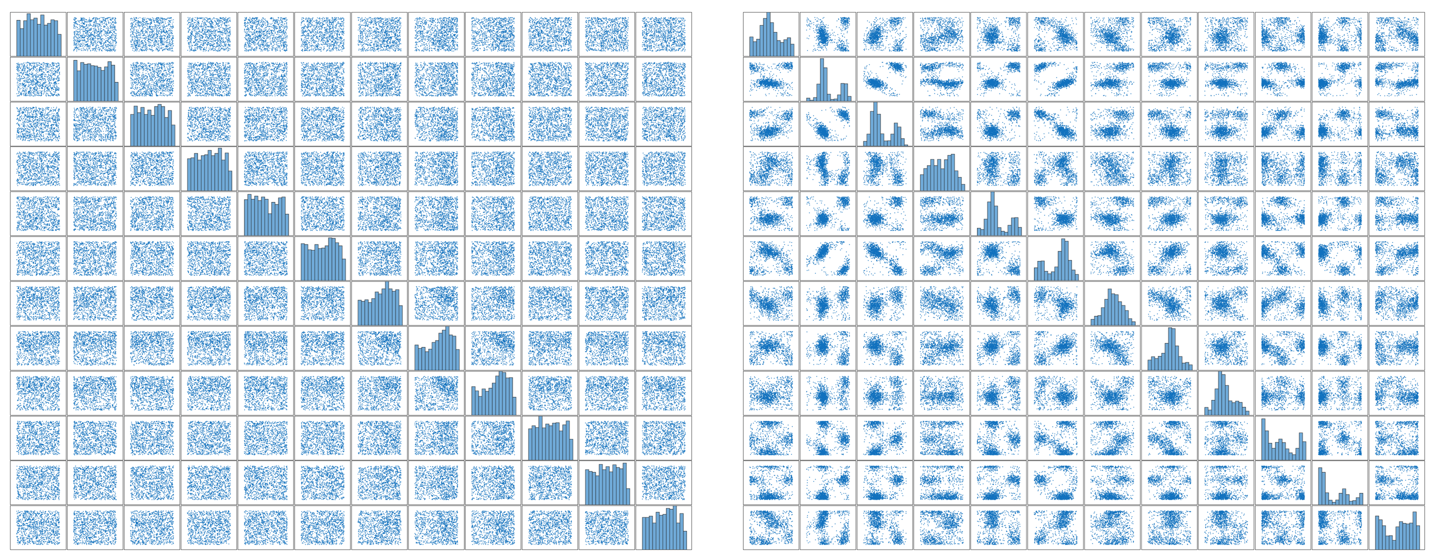
- Simulated annealing example
  - **Converge in probability to the set of globally optimal solutions**

    $$\lim_{k\to\infty} P(X_k \in S_{\min}) = 1$$

  - **We don't know when it happens. It might take years.**



1 **100**



Optimisation Under Uncertainty

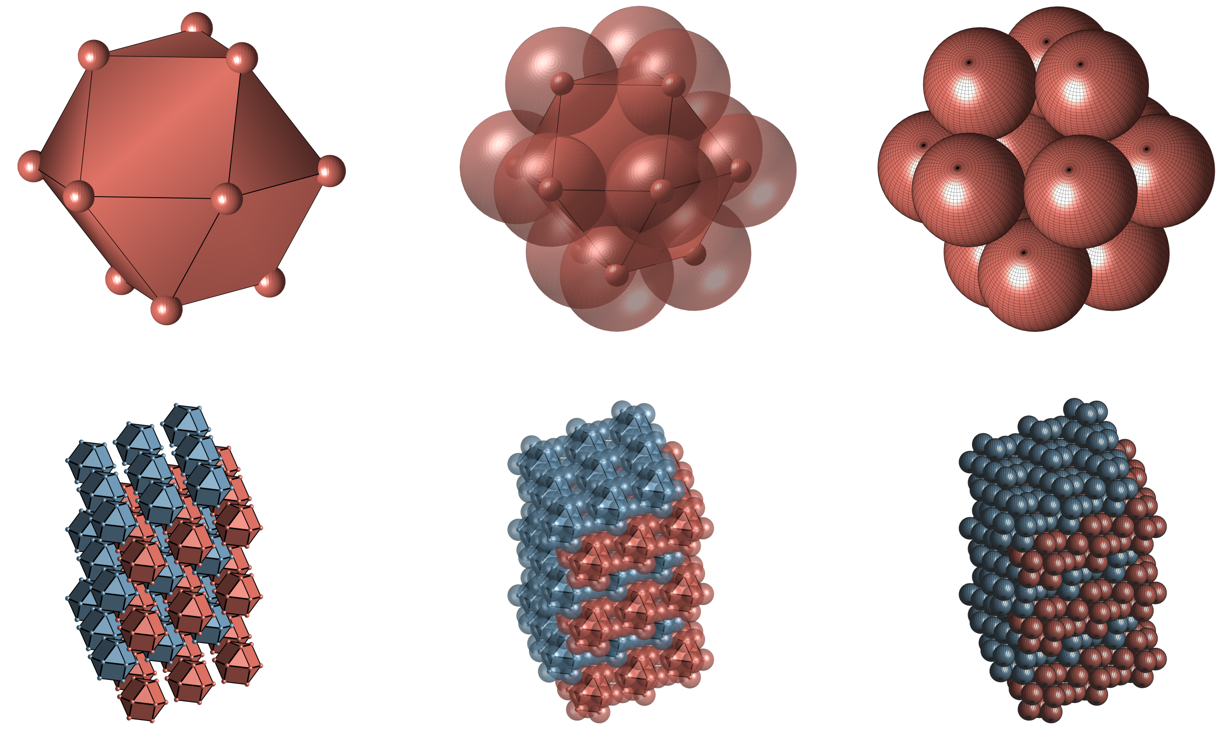
- The universal phrase in global optimisation literature:

"Repeat until convergence."

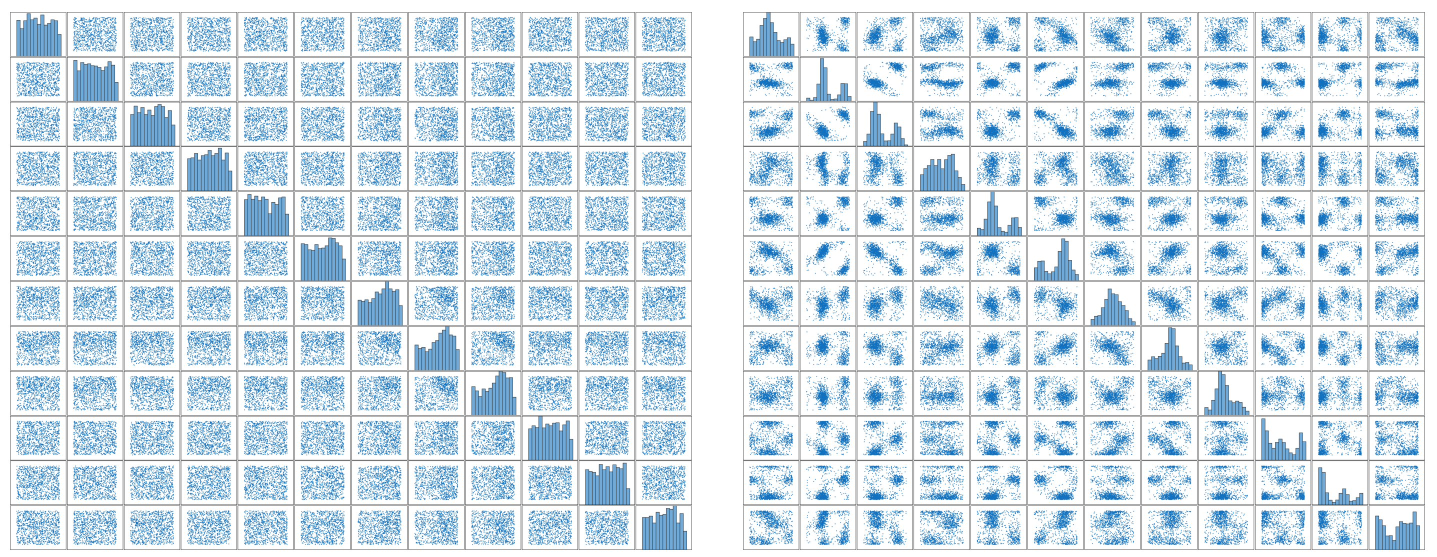
- Simulated annealing example
  - **Converge in probability to the set of globally optimal solutions**

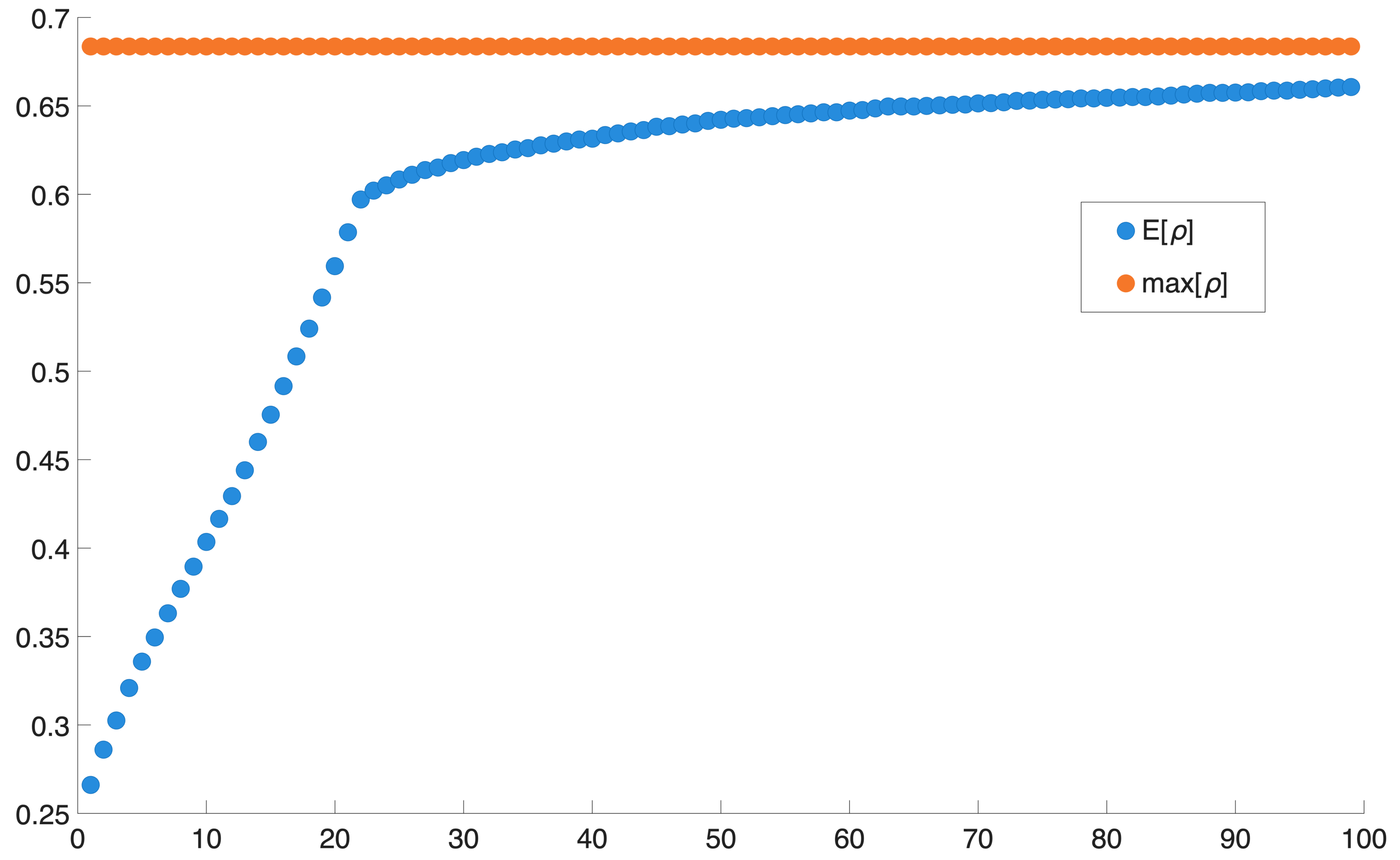
$$\lim_{k \to \infty} P(X_k \in S_{\min}) = 1$$

  - **We don't know when it happens. It might take years.**



1 **100**





Optimisation Under Uncertainty

- The universal phrase in global optimisation literature:

"Repeat until convergence."

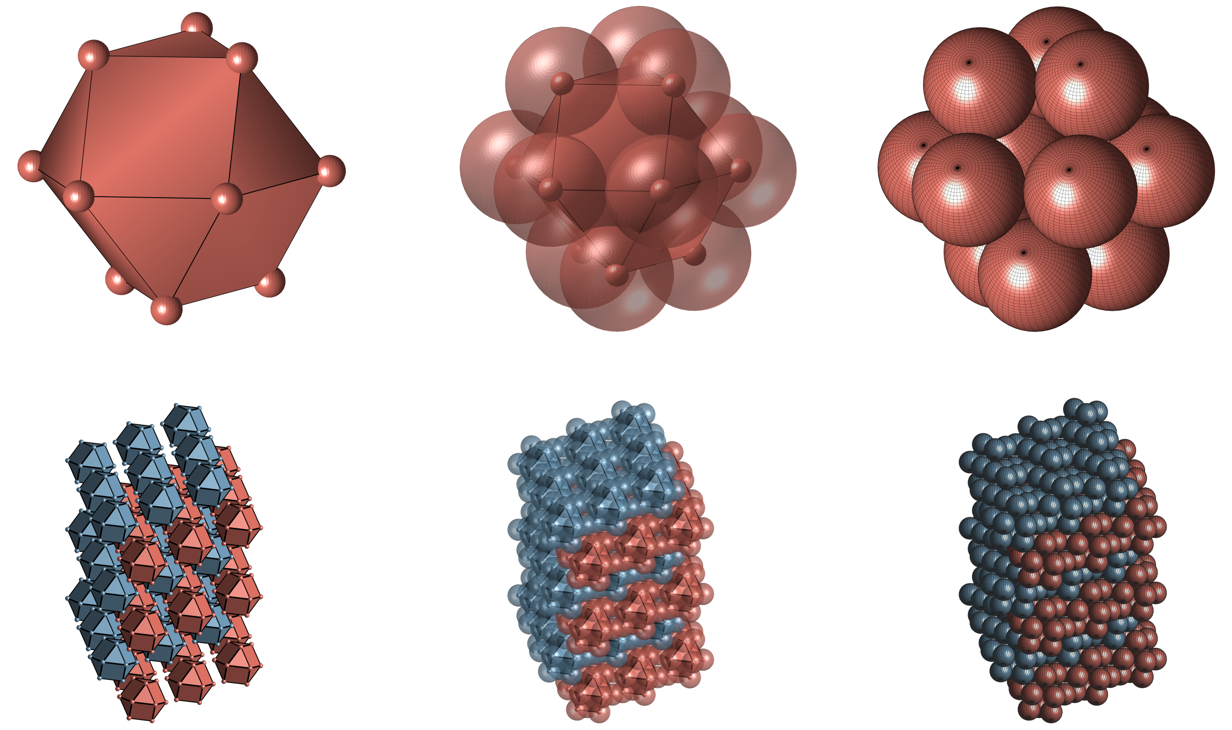
- Simulated annealing example
    - **Converge in probability to the set of globally optimal solutions**

$$\lim_{k\to\infty} P(X_k \in S_{\min}) = 1$$

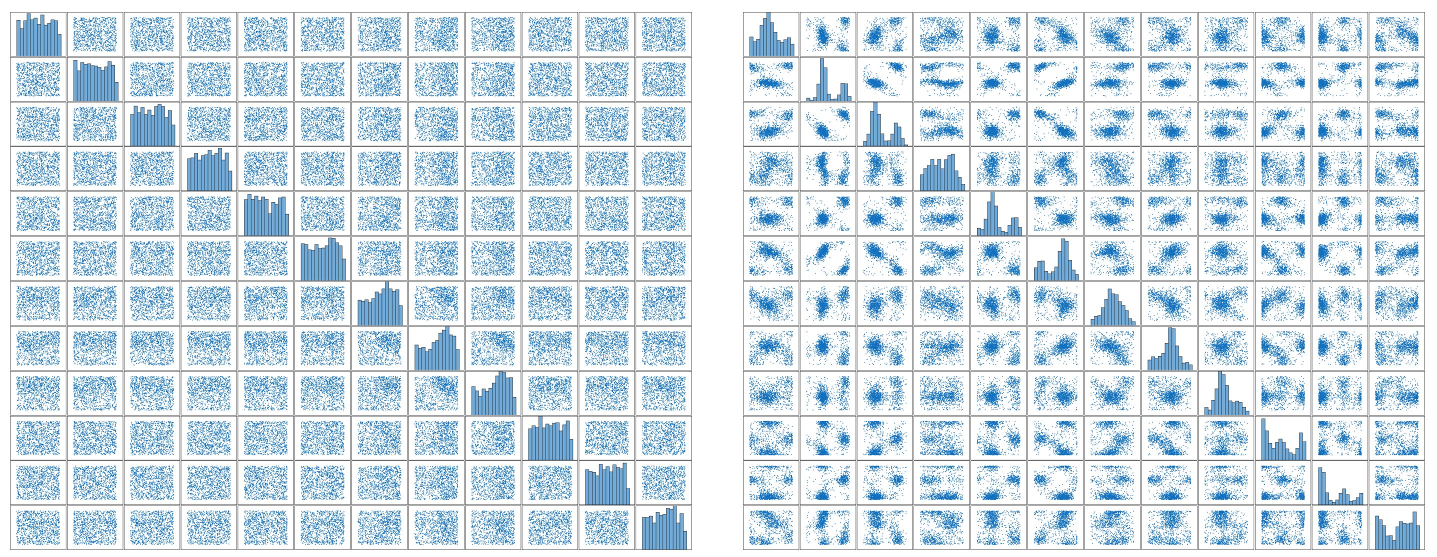
    - **We don't know when it happens. It might take years.**

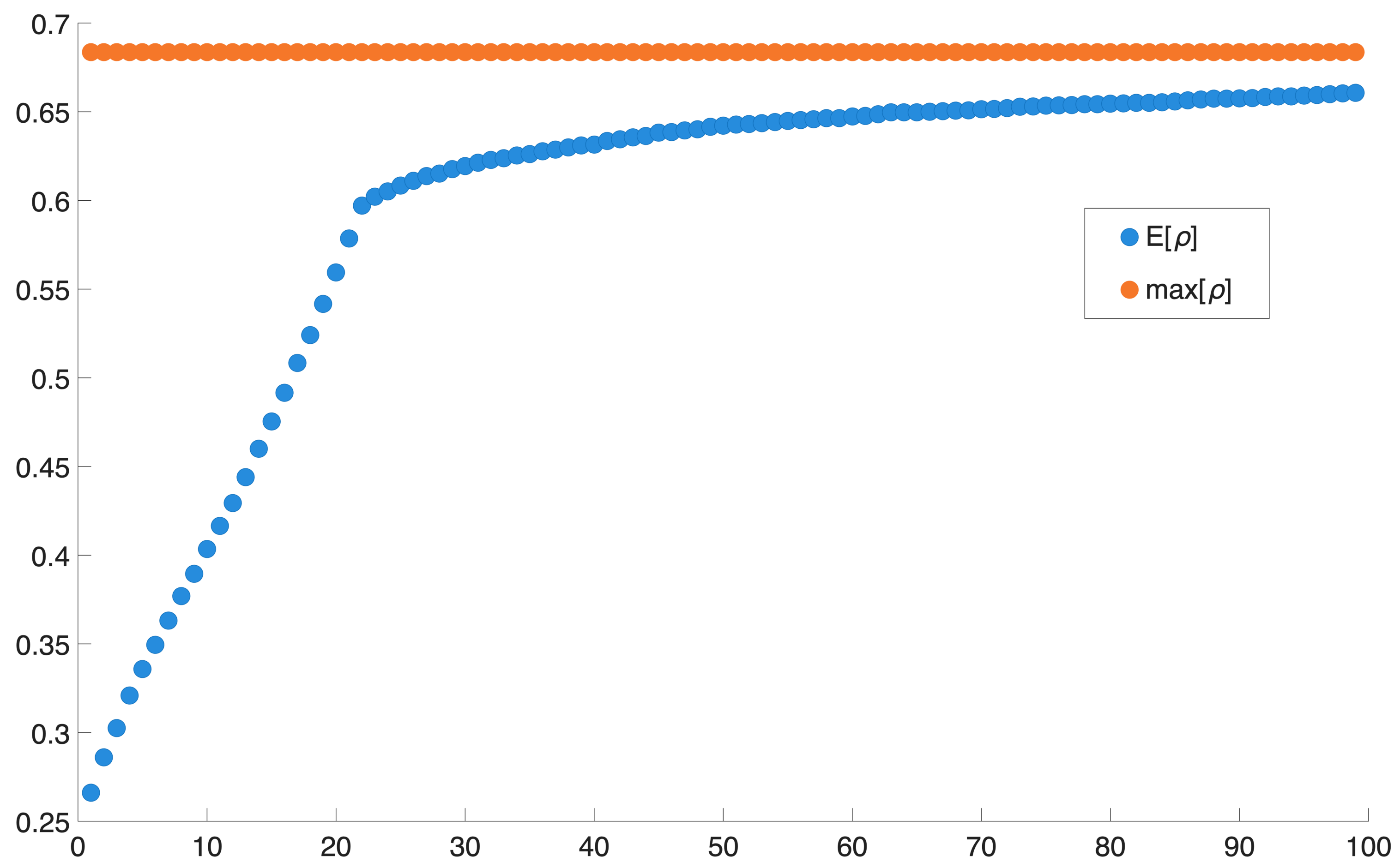
- Estimate optimality gaps

$$\max[\rho] - E[\rho]$$



1 **100**





Optimisation Under Uncertainty

- The universal phrase in global optimisation literature:

"Repeat until convergence."

- Simulated annealing example
  - **Converge in probability to the set of globally optimal solutions**

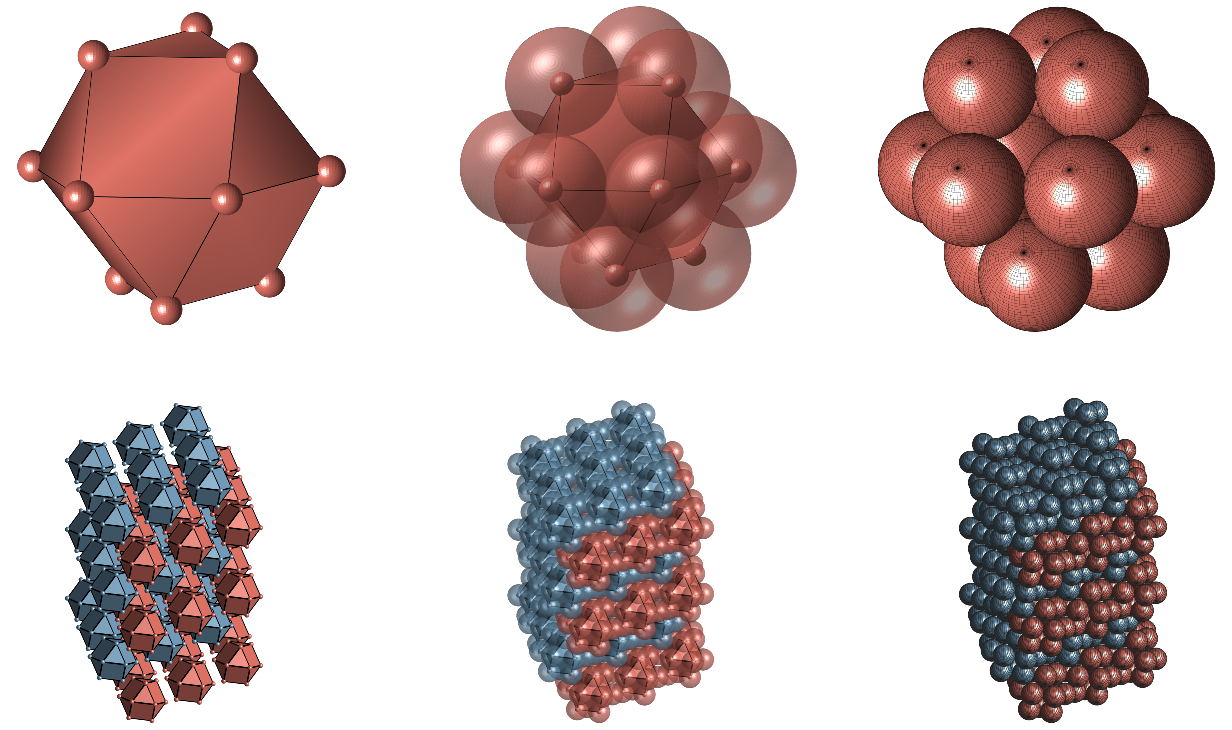
$$\lim_{k\to\infty} P(X_k \in S_{\min}) = 1$$

  - **We don't know when it happens. It might take years.**

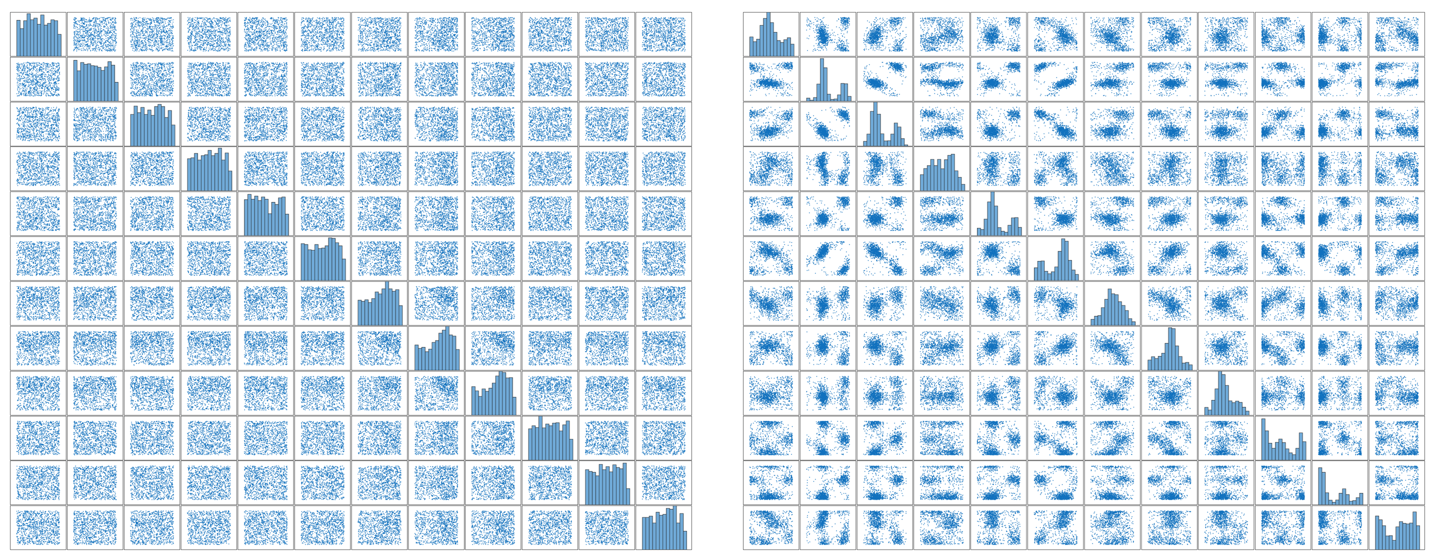
- Estimate optimality gaps

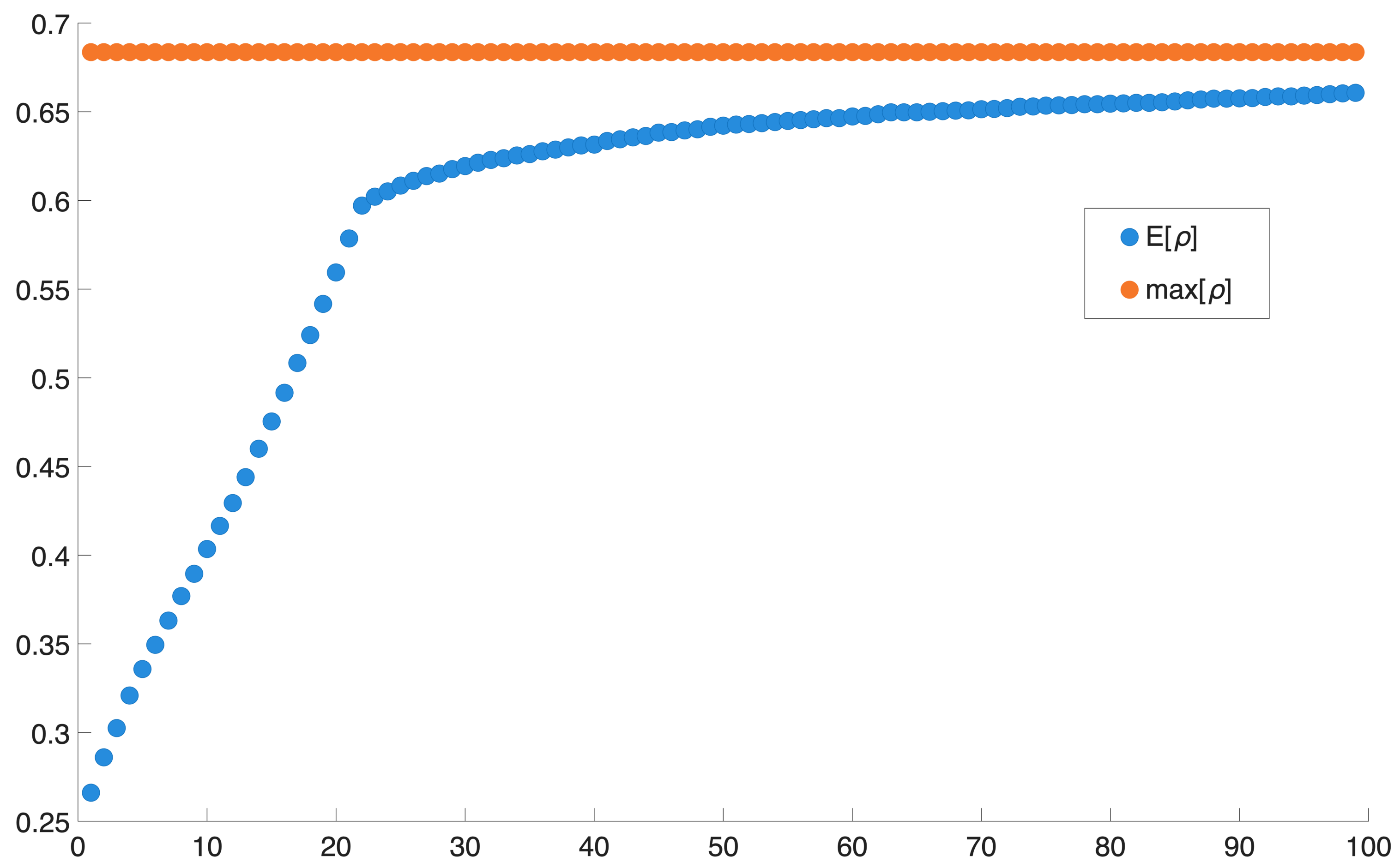
$$\max[\rho] - E[\rho]$$

- **Stopping Criterion**: stop when the optimality gap falls below 0.1%



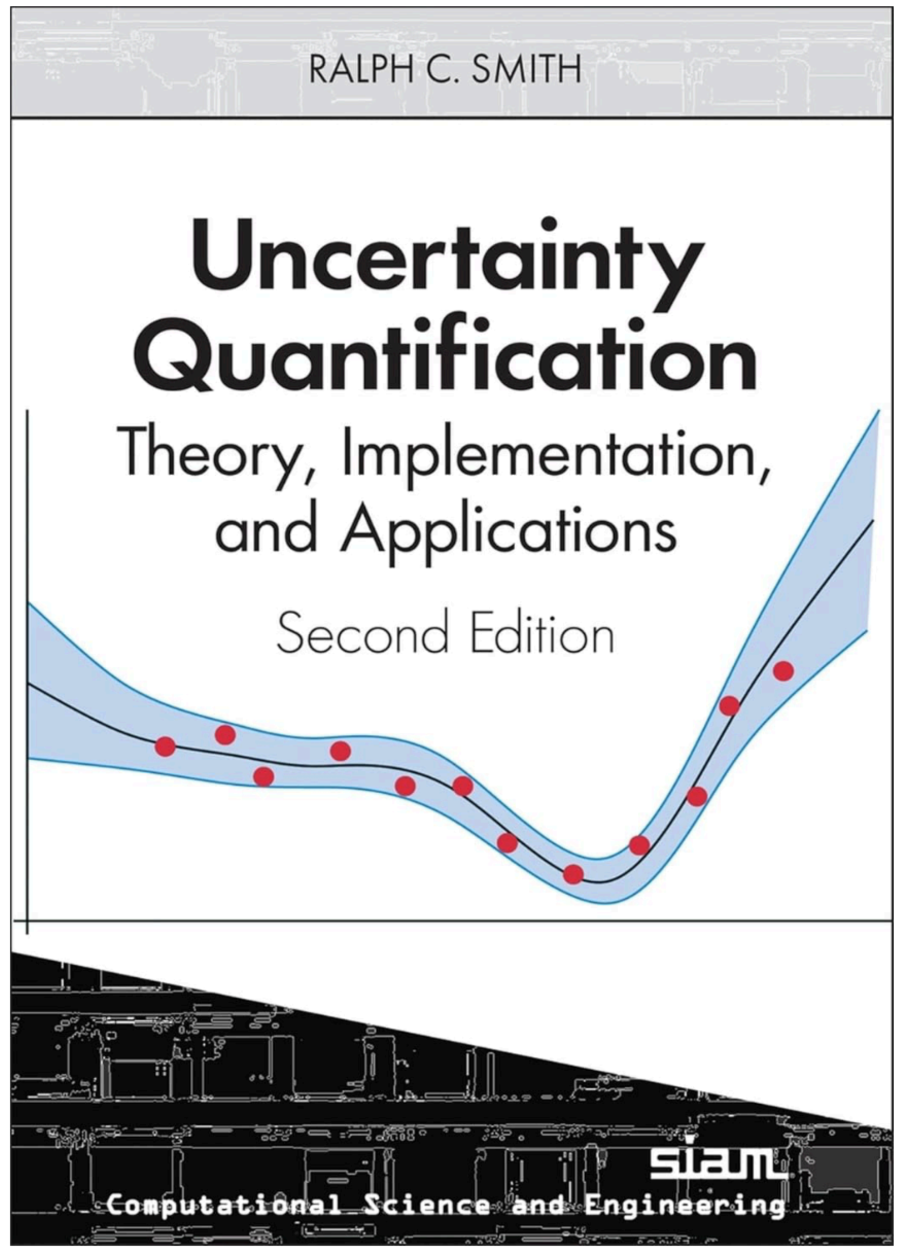
1 **100**





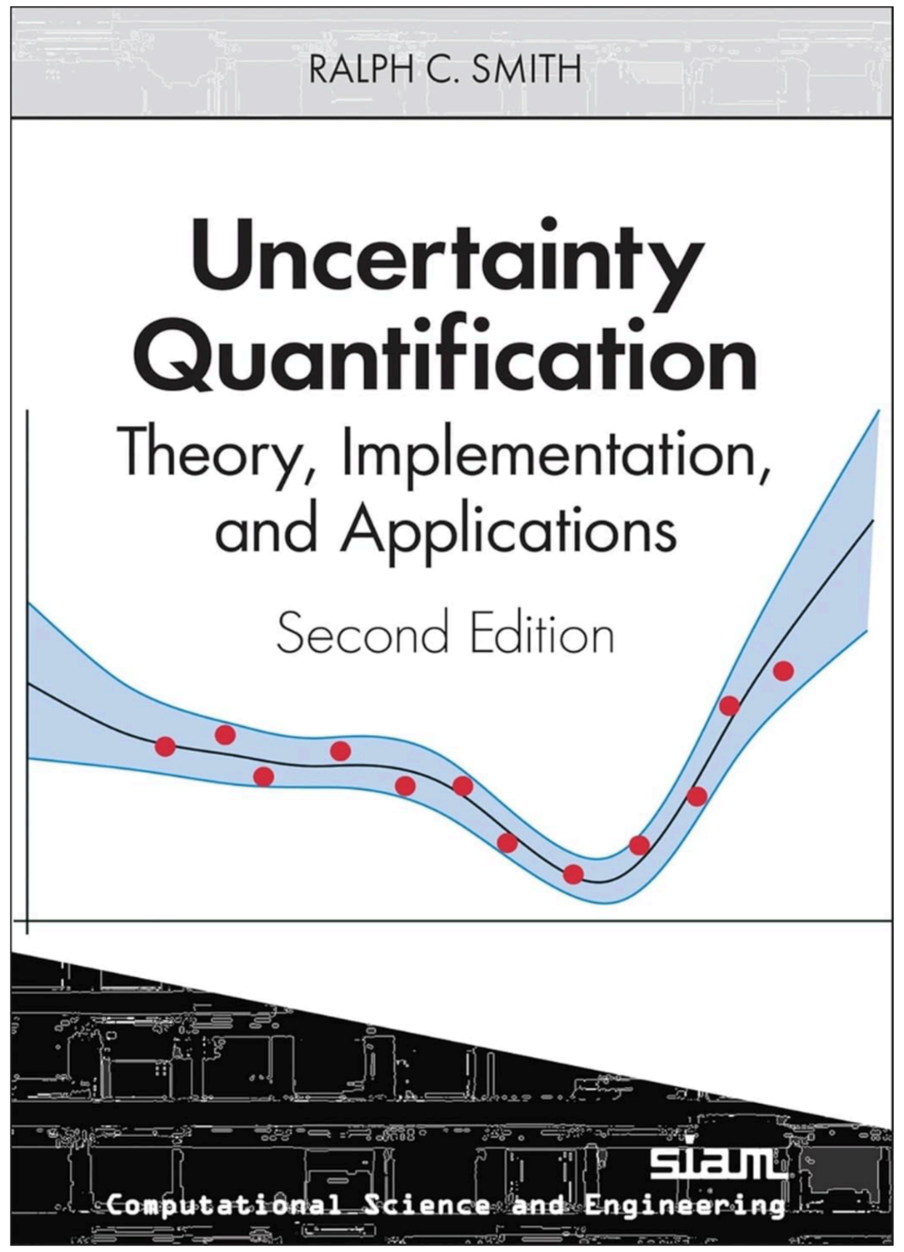
Uncertainties in Predictive Estimation

- Objective of predictive science



Uncertainties in Predictive Estimation

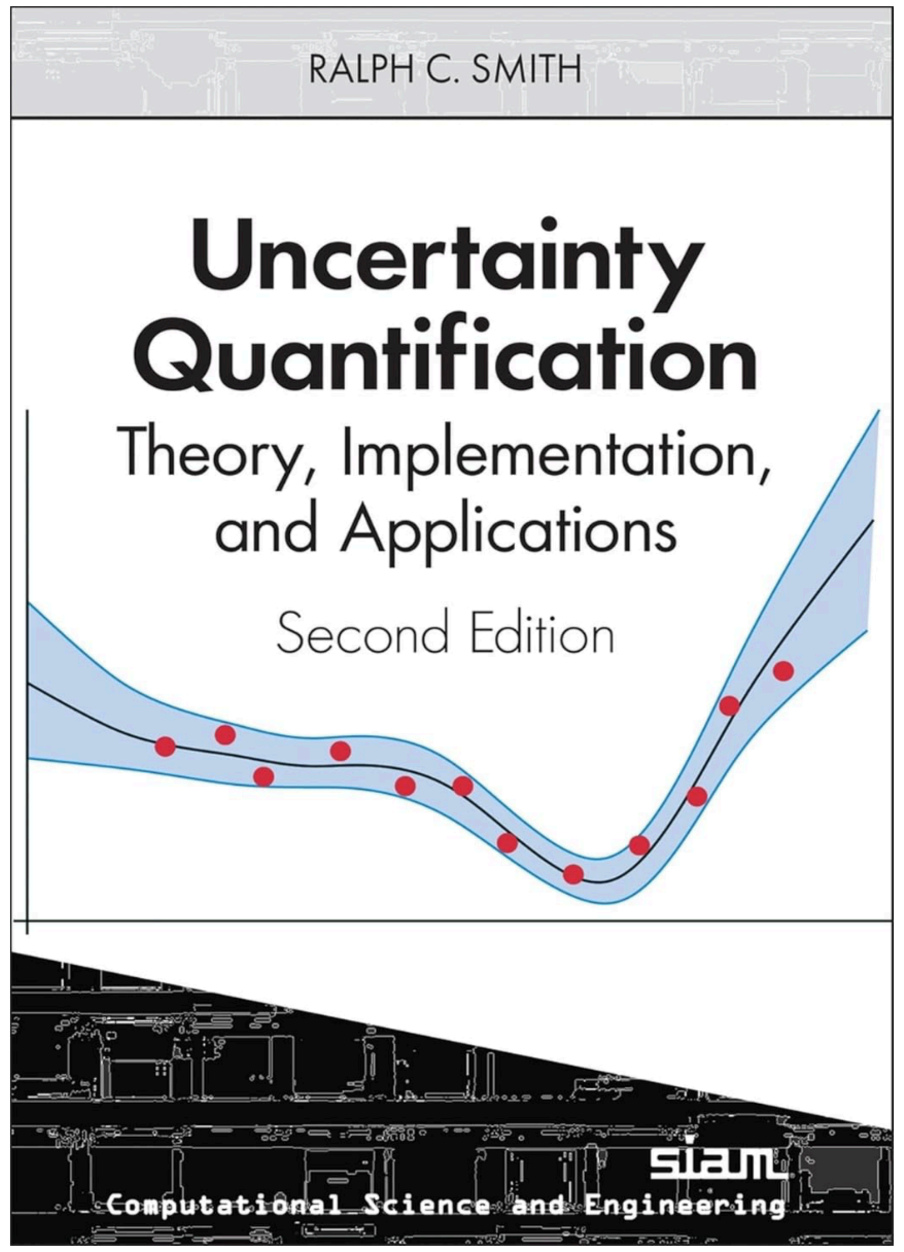
- Objective of predictive science
    - **is to use models, simulation codes, and experiments to predict system responses with quantified and reduced uncertainties**



Uncertainties in Predictive Estimation

- Objective of predictive science
  - **is to use models, simulation codes, and experiments to predict system responses with quantified and reduced uncertainties**
- Experimental Uncertainties and Limitations

  "Experimental results are believed by everyone, except for the person who ran the experiment"



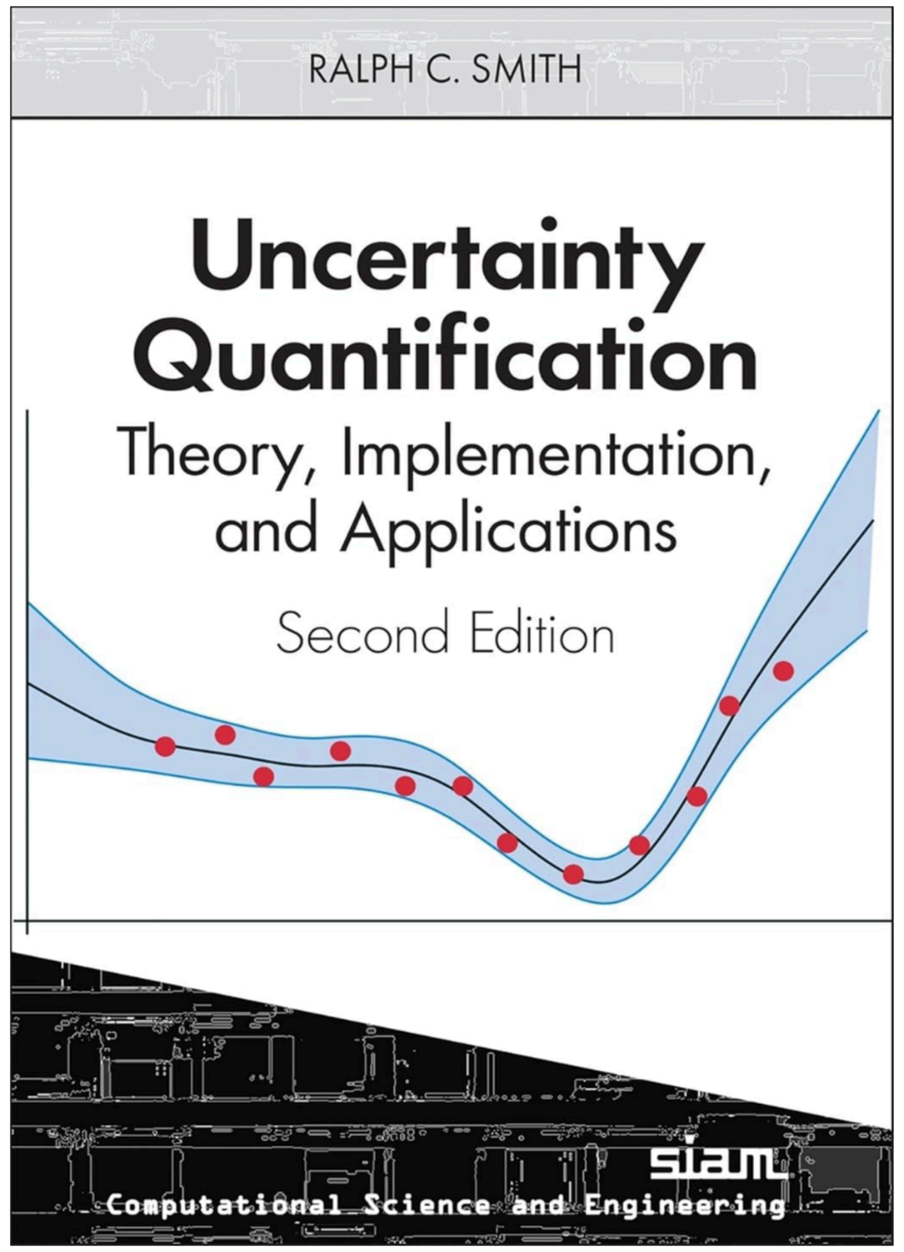
Uncertainties in Predictive Estimation

- Objective of predictive science
  - **is to use models, simulation codes, and experiments to predict system responses with quantified and reduced uncertainties**
- Experimental Uncertainties and Limitations

  "Experimental results are believed by everyone, except for the person who ran the experiment"

- Numerical Errors and Uncertainties

  "Computational results are believed by no one, except the person who wrote the code."



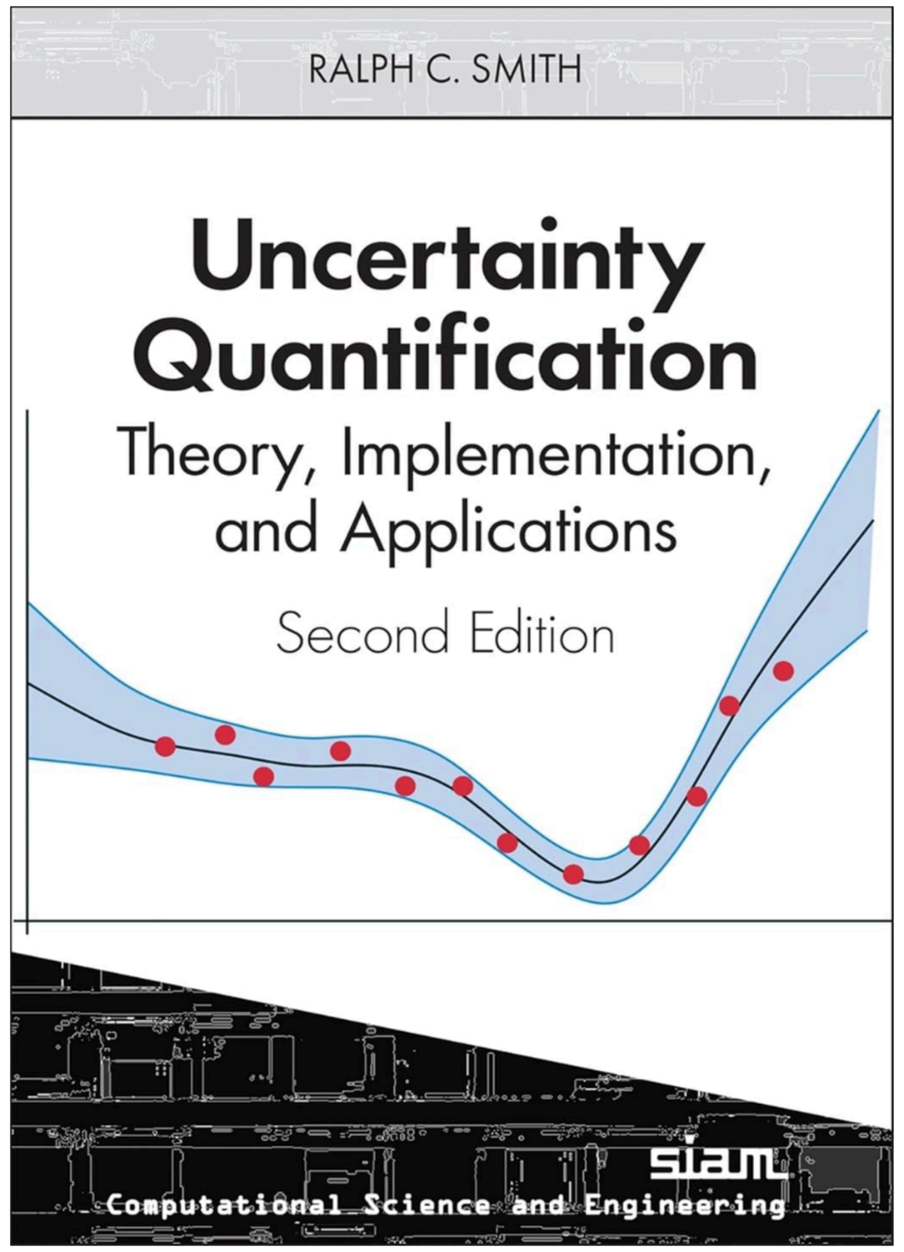
Uncertainties in Predictive Estimation

- Objective of predictive science
    - **is to use models, simulation codes, and experiments to predict system responses with quantified and reduced uncertainties**
- Experimental Uncertainties and Limitations

  “Experimental results are believed by everyone, except for the person who ran the experiment”
- Numerical Errors and Uncertainties

  "Computational results are believed by no one, except the person who wrote the code.”
- Model Discrepancy

  “Essentially, all models are wrong, but some are useful,“ G.E.P. Box



Uncertainties in Predictive Estimation

- Objective of predictive science
    - **is to use models, simulation codes, and experiments to predict system responses with quantified and reduced uncertainties**
- Experimental Uncertainties and Limitations

“Experimental results are believed by everyone, except for the person who ran the experiment”

- Numerical Errors and Uncertainties

"Computational results are believed by no one, except the person who wrote the code.”

- Model Discrepancy

“Essentially, all models are wrong, but some are useful,“ G.E.P. Box

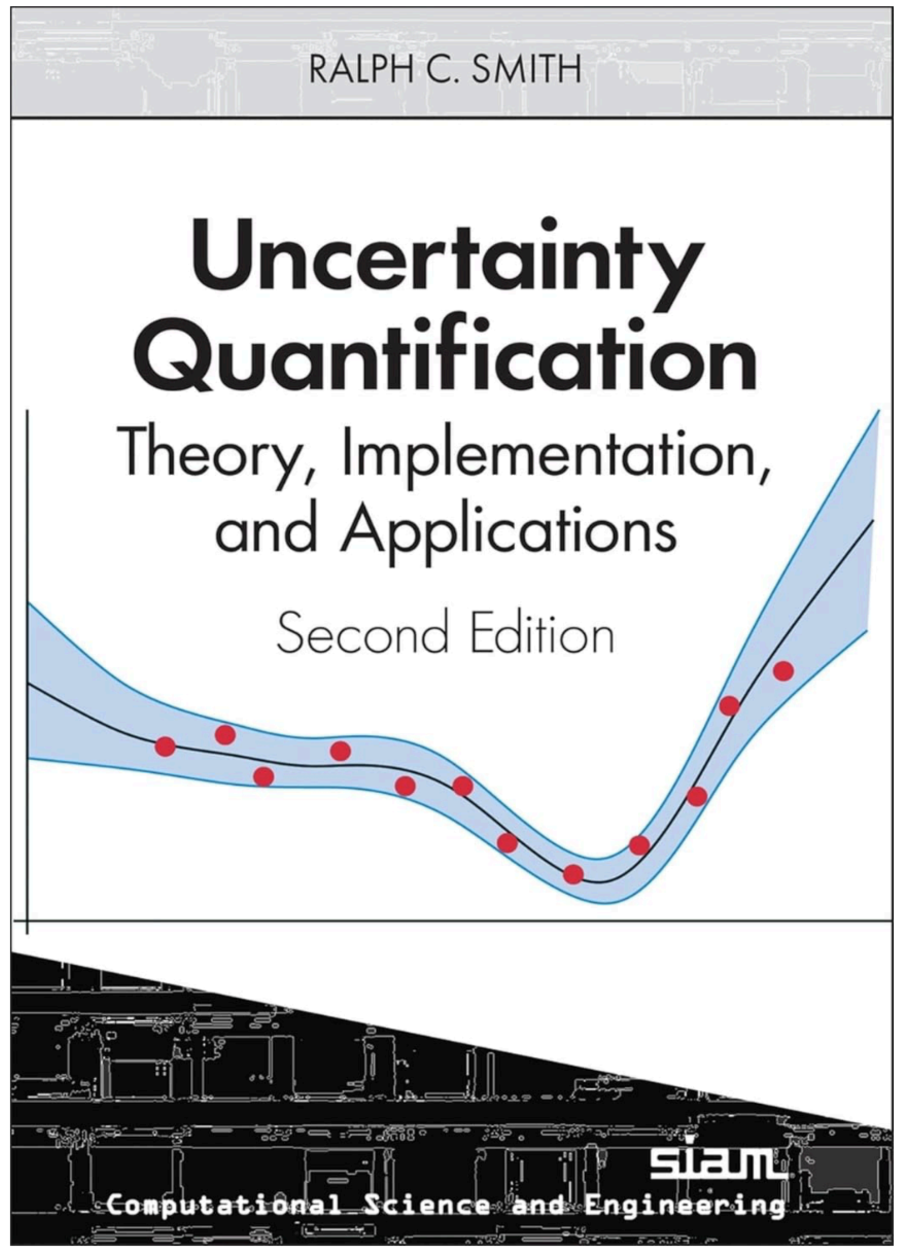
- **Model of i-th observation:** $y_i = \xi(x_i) + \epsilon_i$

$$\xi(x_i) = \eta(x_i, \theta) + \delta(x_i)$$

$\xi(x_i)$ **- true value of the physical system at** x_i

$\eta(x_i, \theta)$ **- value of the simulator at** x_i

$\delta(x_i)$ **- model discrepancy at** x_i



Uncertainties in Predictive Estimation

- Objective of predictive science
  - **is to use models, simulation codes, and experiments to predict system responses with quantified and reduced uncertainties**
- Experimental Uncertainties and Limitations

  “Experimental results are believed by everyone, except for the person who ran the experiment”
- Numerical Errors and Uncertainties

  "Computational results are believed by no one, except the person who wrote the code.”
- Model Discrepancy

  “Essentially, all models are wrong, but some are useful,“ G.E.P. Box

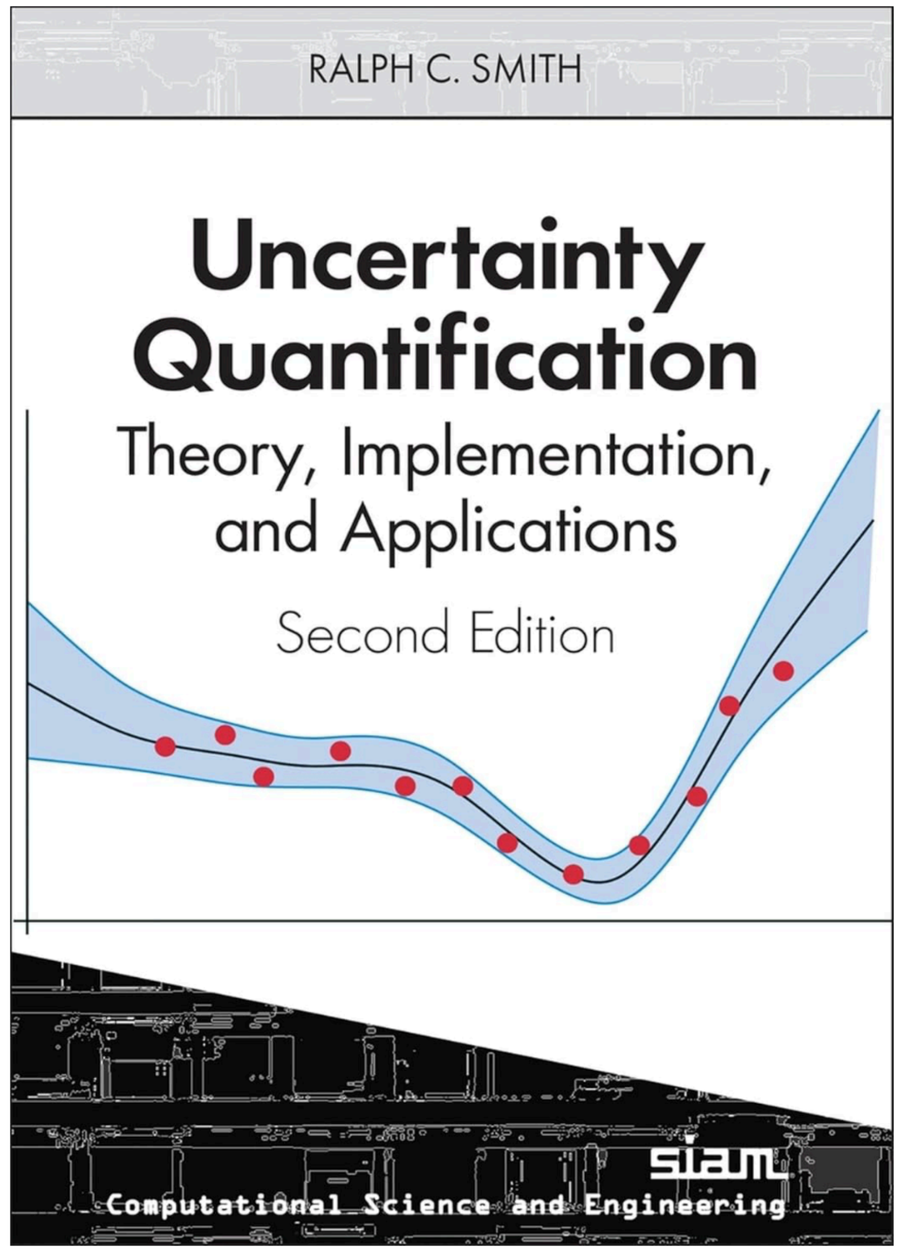
  - **Model of i-th observation:** $y_i = \xi(x_i) + \epsilon_i$

$$\xi(x_i) = \eta(x_i, \theta) + \delta(x_i)$$

$\xi(x_i)$ **- true value of the physical system at** x_i

$\eta(x_i, \theta)$ **- value of the simulator at** x_i

$\delta(x_i)$ **- model discrepancy at** x_i



- When **extrapolating** to contexts outside those for which we have observations of the physical system, it is essential to impose **strict prior information** on the permissible function space used to construct $\delta(x_i)$.

THANK YOU

The authors thank the Leverhulme Trust for funding this research via the Leverhulme Research Centre for Functional Materials Design