

Bayesian Random Tomography of Particle Systems

Group Meeting

Yu-Hsiang Lien 連昱翔



Bayesian Random Tomography of Particle Systems

Nima Vakili¹ and Michael Habeck^{1,2*}

¹*Microscopic Image Analysis Group, Jena University Hospital, Jena, Germany*, ²*Statistical Inverse Problems in Biophysics, Max Planck Institute for Biophysical Chemistry, Göttingen, Germany*

The Reconstruction Problem in Random Tomography

- Aim to reconstruct a 3D volume $f(r)$ for $r \in \mathbb{R}^3$ and $f: \mathbb{R}^3 \mapsto \mathbb{R}_+$

- The projection images $g(u)$ can be described as:

$$g(u) = \int f(R^T r) dz = \mathcal{X}_R[f](u) \quad \rightarrow \text{X-ray transform}$$

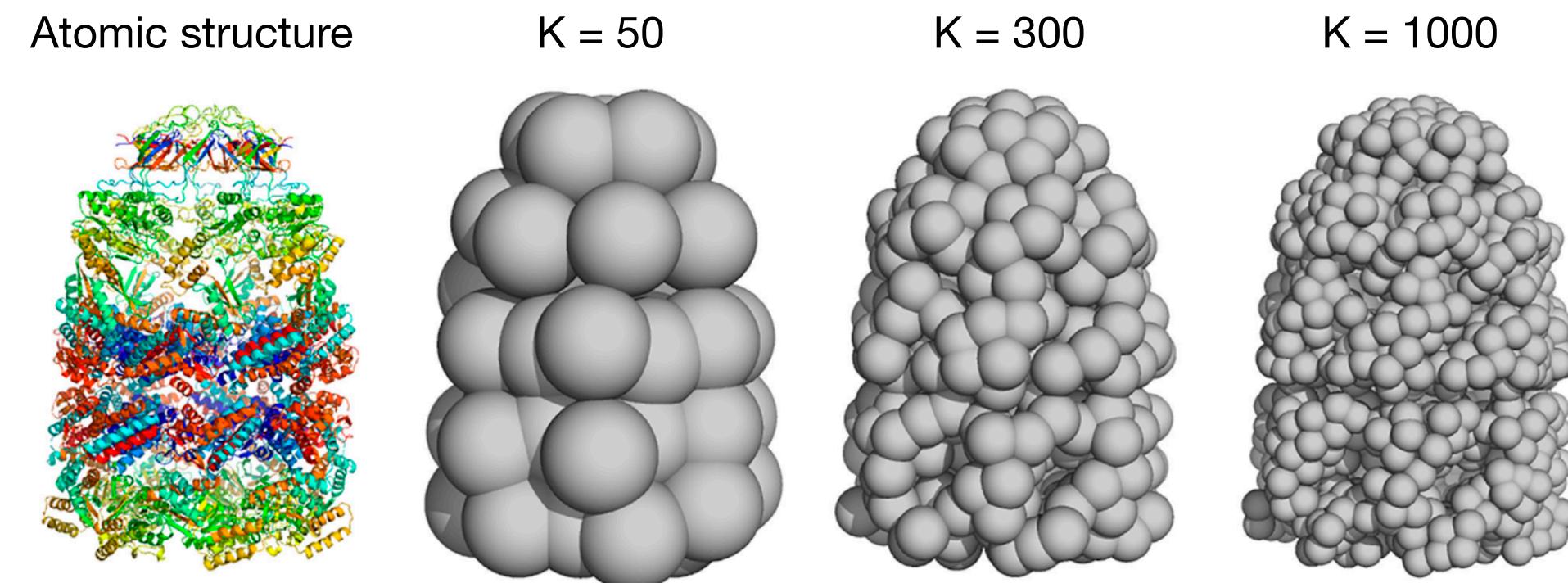
- ▶ $R \in \text{SO}(3)$ is a 3D rotation matrix, $u \in \mathbb{R}^2$ denotes a position in the projection image.
- The reconstruction problem in random tomography is to estimate $f(r)$ from N random projection directions R_n , s.t.:

$$g_n(u) = \mathcal{X}_{R_n}[f](u) + n(u), \quad n = 1, \dots, N \quad n(u): \text{Noise}$$

Kernel Expansion of Images and Volumes

- The discretization of images and volumes:

- ▶ **Regular 2D and 3D grids** → pixels and voxels
- ▶ **Sums of basis functions** → Radial Basis Function (RBF) kernels



resolution depends on the number of particles K

- The kernel representation of volume:

$$f(r) = \sum_{k=1}^K w_k \phi(r - x_k)$$

- In physical interpretation, the model can be interpreted as **the blurred version of a particle system**: (weighted point cloud)

$$f(r) = \left(\phi * \sum_{k=1}^K w_k \delta_{x_k} \right)(r)$$

- Collection of K particles at positions x_k with mass $w_k > 0$.
- Particle density $\sum_k w_k \delta_{x_k}$ is blurred by a convolution with the RBF kernel.

Kernel Expansion of Images and Volumes

- The RBF kernel representation has the efficient transformation properties:

Transformation of $f(r) \rightarrow$ Transformation of particle positions x_k

$$f(r) \xrightarrow{R,t} f(R^T(r-t)) = \sum_k w_k \phi(r - Rx_k - t) = \sum_k w_k \phi(r - x'_k)$$

- There are many options for $\phi(r)$, here we choose **Gaussian RBF kernels**:

$$\phi_d(r; x, \sigma^2) = \frac{1}{(2\pi\sigma^2)^{d/2}} \exp\left\{-\frac{\|r - x\|^2}{2\sigma^2}\right\}$$

→ The d -dimensional Gaussian

- The volume representation become **a mixture of K spherical Gaussian**:

$$f(r) = \sum_{k=1}^K w_k \phi_3(r; x_k, \sigma^2)$$

The Spherical Gaussian Kernel

- A convenient property of the spherical Gaussian kernel is its behavior under the X-ray transform:

$$\mathcal{X}_R [\phi_d] (u) = \phi_{d-1} (u; PRx, \sigma^2)$$

- Spherical Gaussians are closed under the X-ray transform, and the projected volume is a K component mixture of spherical Gaussians:

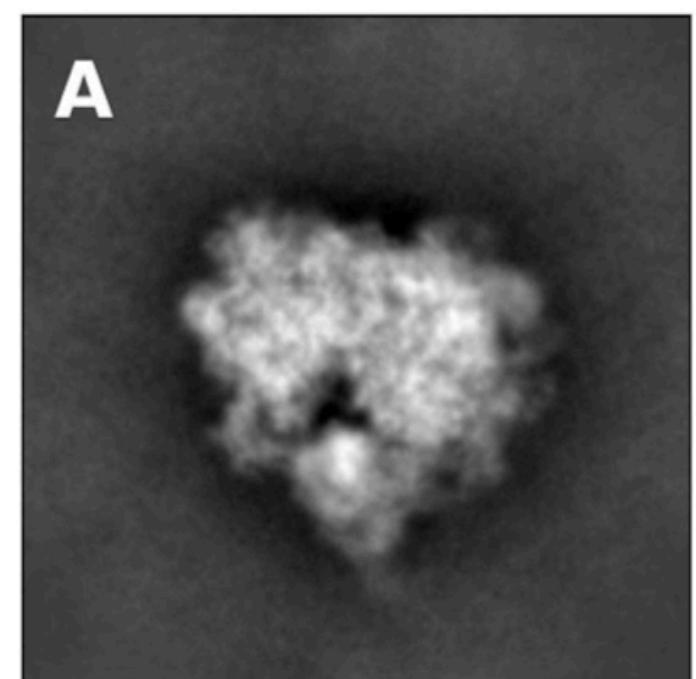
$$\mathcal{X}_R [f] (u) = \sum_{k=1}^K w_k \phi_2 (u; PRx_k, \sigma^2)$$

Probabilistic Model

- Unknown parameters of model:
 - ▶ Positions x_k
 - ▶ Rotation matrices R_n
 - ▶ Weights $w_k \rightarrow$ interpret the Gaussian components as particles of equal mass. Fixed the weights: $w_k = K^{-1}$
- Two probabilistic models for the input data:
 - ▶ **Directly used input images** $\{g_n ; n = 1, \dots, N\}$
 - ▶ **Kernel expansion of input images** (mixture of 2D spherical Gaussians) $Y_n = \{y_{nm} \in \mathbb{R}^2 ; m = 1, \dots, M_n\}$

Representation of Projection Images by Point Clouds

Original image



Thresholding



Clustering



D



E



Picking

2D Point Cloud
(EM algorithm)

Model image

- The projection properties of mixtures of spherical Gaussians suggest to also represent the projection image as a mixture of Gaussians (**2D point cloud**).
- A particle-based representation of the central connected component (signal) is obtained by the **EM algorithm**.
- The cross-correlation coefficient between the model and the original image w/o masking (w/ masking) is **95.8% (99.6%)**.

Point Clouds to Images w/ E-M Algorithm

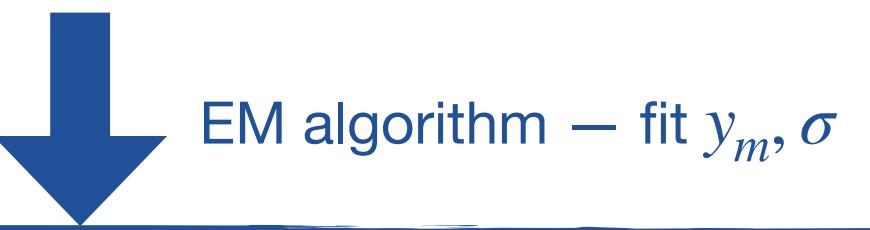
- Given a collection of pixels u_i with associated intensities g_i ($i = 1 \dots I$), aim to represent this by a 2D point cloud $\{y_m; m = 1, \dots, M\}$ (mixture of Gaussians):

$$g_i \approx \alpha + \gamma \sum_{m=1}^M \phi_2(u_i; y_m, \sigma^2)$$

→ correct for the background and let $\alpha = 0, \gamma = 1$ →

$$g_i \approx \sum_{m=1}^M \phi_2(u_i; y_m, \sigma^2)$$

α : suitable background parameter
 γ : scaling factor



- Soft assignment: Each pixel u_i with associated intensity g_i is assigned to a 2D point y_m with probability p_{im} :

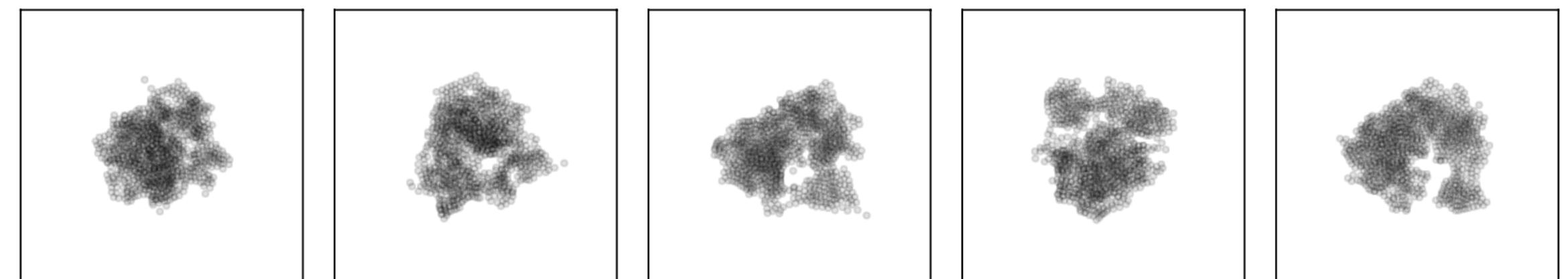
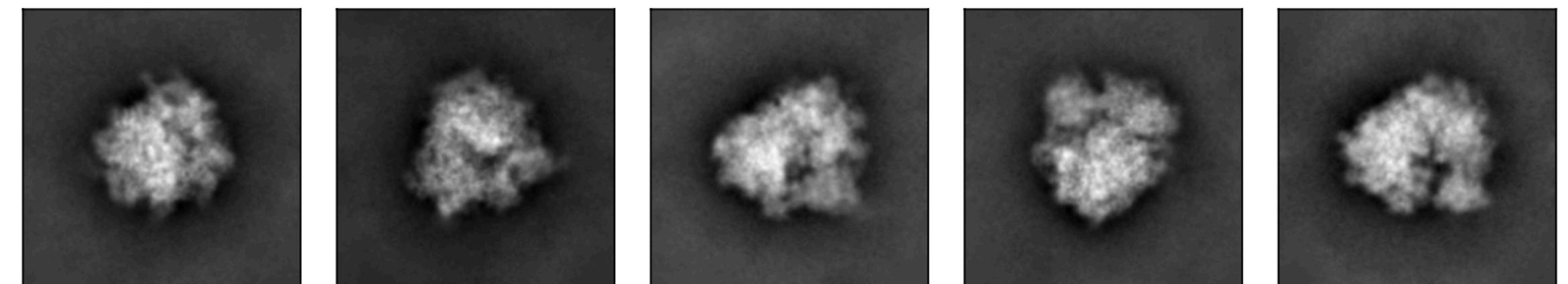
$$p_{im} = \frac{\phi_2(u_i; y_m, \sigma^2)}{\sum_{m'=1}^M \phi_2(u_i; y_{m'}, \sigma^2)}$$

- Particle positions y_m are the centers of mass of the assigned pixels weighted by the image intensity g_i and the assignment probabilities p_{im} computed in the previous step:

$$y_m = \frac{\sum_{i=1}^I p_{im} g_i u_i}{\sum_{i=1}^I p_{im} g_i}$$

- The width of the Gaussians is estimated by:

$$\sigma = \sqrt{\frac{1}{\sum_{i=1}^I g_i} \sum_{i=1}^I \sum_{m=1}^M p_{im} g_i \|u_i - y_m\|^2}$$



Class averages (top) and point cloud representation (bottom) of 80S ribosome

Likelihoods of Images

- The likelihood of the n -th image (**idealized image formation model**):

$$\Pr(g_n | x, R_n, t_n, \gamma_n, \alpha_n, \tau_n) = \left(\frac{\tau_n}{2\pi}\right)^{M_n/2} \exp \left\{ -\frac{\tau_n}{2} \sum_{m=1}^{M_n} \left[g_{nm} - \alpha_n - \gamma_n \sum_k \phi_2(u_{nm}; PR_n x_k + t_n, \sigma^2) \right]^2 \right\}$$

- The likelihood of the n -th image (**mixture of spherical Gaussians**): $g_n(u) \approx \sum_{m=1}^{M_n} \phi_2(u; y_{nm}, \sigma_n^2)$

$$\Pr(Y_n | x, R_n, t_n, \sigma_n) = \prod_{m=1}^{M_n} \frac{1}{K} \sum_{k=1}^K \phi_2(y_{nm}; PR_n x_k + t_n, \sigma_n^2)$$

- Denote all nuisance parameters (all parameters except x and R_n) by ξ , and denote the data by D . s.t. we denote both likelihood as:

$$\Pr(D | x, R, \xi)$$

Posterior Distribution

- To derive the posterior distribution by invoking Bayes' theorem:

$$\Pr(x, R, \xi | D) = \frac{\Pr(D | x, R, \xi) \Pr(x, R, \xi)}{\Pr(D)}$$

- Assume all parameters are independent:

$$\Pr(x, R, \xi) = \Pr(x) \Pr(R) \Pr(\xi)$$

- ▶ The normalization factor $\Pr(D)$ is the model evidence, which can be ignored if we are only interested in parameter estimation.

Priors

- Priors for the nuisance parameters:
 - ▶ Precisions τ_n and $1/\sigma_n^2 \rightarrow$ Jeffreys priors.
 - ▶ Scaling factors γ_n and offsets $\alpha_n \rightarrow$ flat distribution.
 - ▶ Image shifts $t_n \rightarrow$ a zero-centered 2D Gaussian distribution.
- Priors for the rotations $R_n \rightarrow$ uniform distribution over $SO(3)$:

$$\Pr(R) = \prod_{n=1}^N \Pr(R_n) \propto 1$$

- Priors for the positions x_k :
 - ▶ Previous: zero-centered Gaussian distribution.

Priors for the Particle Positions

- Using **Boltzmann distribution** over the positions involving a soft repulsive interaction potential $E(x)$:

$$\Pr(x_1, \dots, x_K) \propto \exp\{-\beta E(x_1, \dots, x_K)\}$$

- Choose a **quartic repulsion** (commonly used in NMR structure calculation):

$$E(x_1, \dots, x_K) = \sum_{k < k'} [\|x_k - x_{k'}\| \leq 2R] \left(1 - \frac{\|x_k - x_{k'}\|}{2R}\right)^4$$

- The estimates for R and β are based on an analysis of several biomolecular structures at different levels of coarse graining.
- To avoid isolated particles, add a radius of **gyration term** such that the overall prior for particle positions is:

$$\Pr(x_1, \dots, x_K) \propto \exp\{-\beta E(x_1, \dots, x_K)\} \exp\{-\alpha R_g(x)\}$$

α doesn't have strong impact on the final structure.

Inference

- Bayesian random tomography employs **Markov Chain Monte Carlo (MCMC)** sampling from the posterior distribution.
- Use **Gibbs sampling strategy** where each group of parameters, the particle positions x , the rotations R and the nuisance parameters ξ , is updated separately while clamping the other parameters to their current values.
- The conditional posteriors of the particle positions x and the rotations R are not of a standard form and need to be updated with more sophisticated algorithms:
 - ▶ Particle position: **Hamiltonian Monte Carlo (HMC)**.
 - ▶ Rotation matrices: **Metropolis-Hastings (MH) algorithm**.
- To update the nuisance parameters, use standard samplers for generating Gamma variates and normally distributed random variables.

Metropolis-Hastings (MH) Algorithm

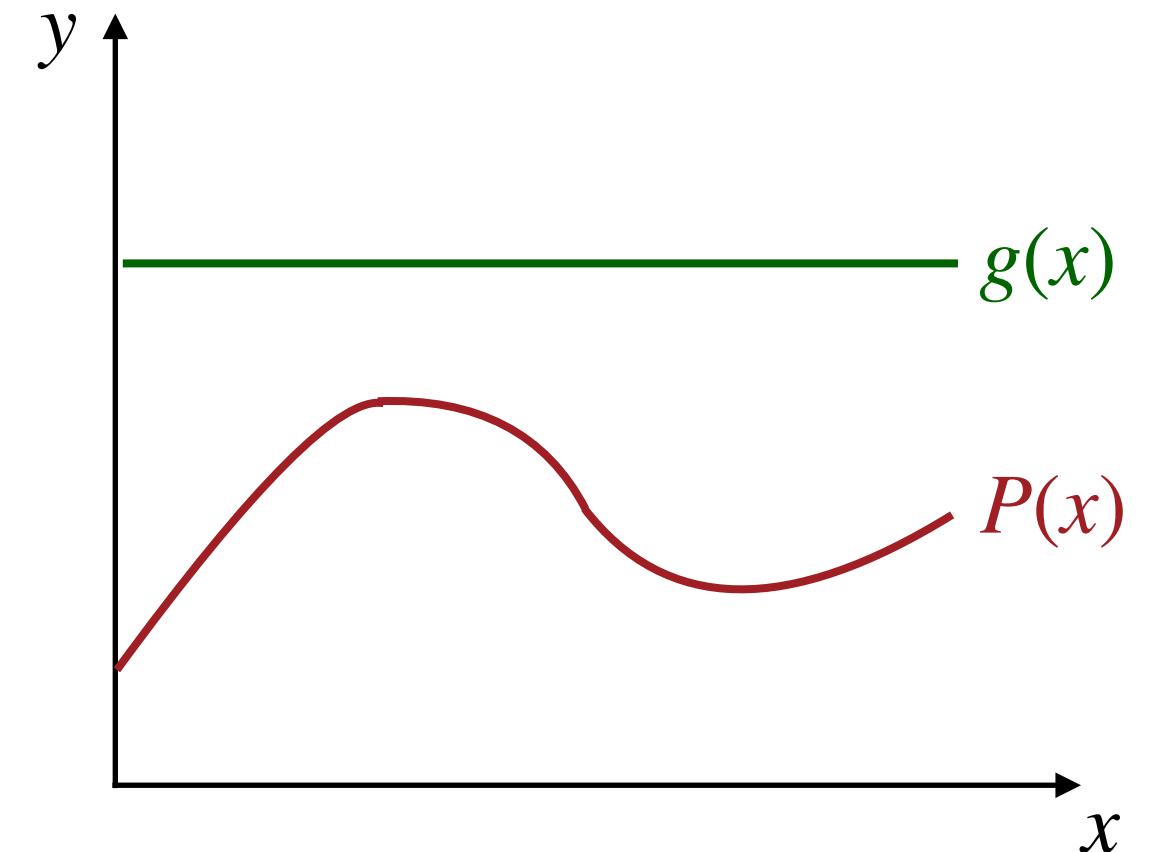
- Assume $P(x)$ is the desire distribution, the MH algorithm can be written as follows:

- Initialise:
 - Pick an initial state x_0
 - Set $t = 0$

- Iterate:
 - Generate a random candidate state x' according to $g(x'|x_t)$.

- Calculate the acceptance probability $A(x'|x_t) = \min\left(1, \frac{P(x')}{P(x_t)} \frac{g(x_t|x')}{g(x'|x_t)}\right)$.

- Accept or reject event base on acceptance.
 - Increment: set $t = t + 1$



Sampling Particle Positions w/ Hamiltonian Monte Carlo (HMC)

- The conditional posterior distribution over particle positions is:

$$\Pr(x|R, \xi, D) \propto \Pr(D|x, R, \xi) \Pr(x)$$

- In HMC, $-\log \Pr(x|R, \xi, D)$ defines a potential energy over configuration space that is composed of an attractive term $-\log \Pr(D|x, R, \xi)$ matching particle positions to the projection data, and a **repulsive contribution** $-\log \Pr(x)$ stemming from the excluded-volume term.
- For fixed rotations and nuisance parameters, the particle positions undergo Hamiltonian dynamics following the gradient of $-\Pr(x|R, \xi, D)$.

Sampling Rotational Parameters w/ Metropolis-Hastings

- The conditional posterior distribution over rotations is:

$$\Pr(R_n | x, \xi, D) \propto \exp \left\{ -\frac{\tau_n}{2} \sum_{m=1}^{M_n} \left[g_{nm} - \alpha_n - \gamma_n \sum_{k=1}^K \phi_2(u_{nm}; PR_n x_k + t_n, \sigma^2) \right]^2 \right\}$$

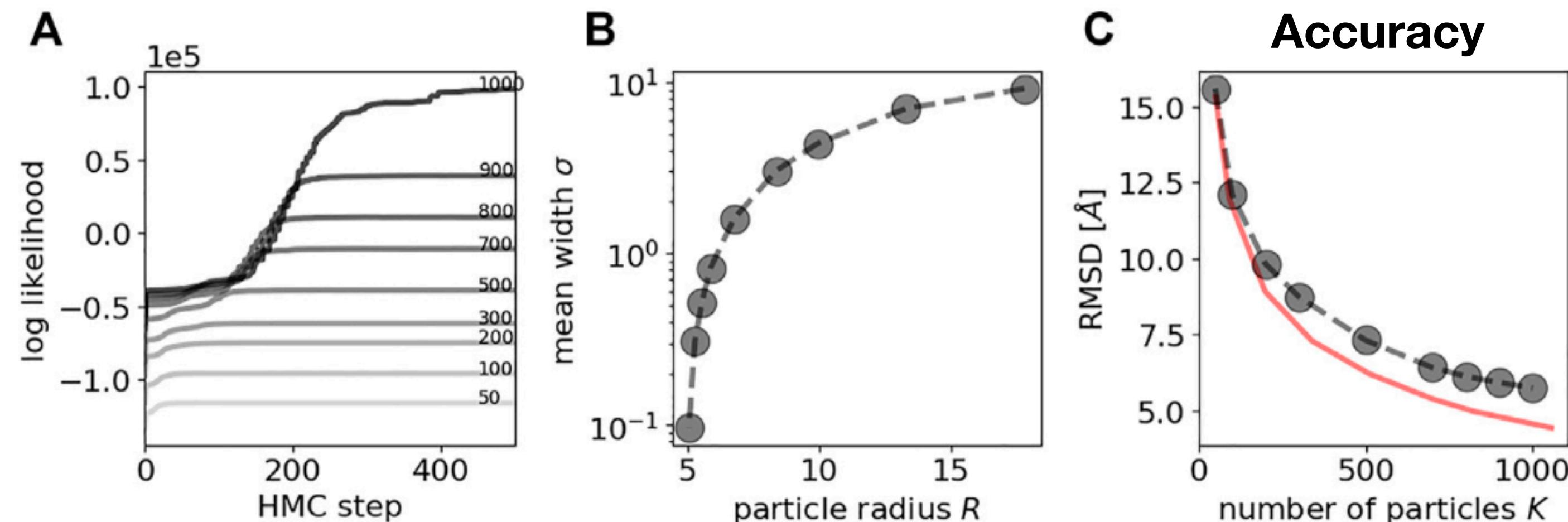
$$\Pr(R_n | x, \xi, D) \propto \prod_{m=1}^{M_n} \sum_{k=1}^K \phi_2(y_{nm}; PR_n x_k + t_n, \sigma^2)$$

- Parameterize rotation matrices using **unit quaternions**.
- Use the **Metropolis-Hastings (MH) algorithm** to estimate the rotation matrices.

Sampling Tests

- To generate input 2D point clouds, use the crystal structure of **GroEL/GroES** (58,674 atom coordinates in total).
- The 2D point clouds are generated by projecting the 3D positions of every 10th Carbon-alpha atom (802 points in total) along 35 random directions into 2D.
- Also generated corresponding projection images by blurring the point clouds with a Gaussian filter of width 5 \AA .

Sampling Particle Positions and Precisions w/ Fixed Rotations



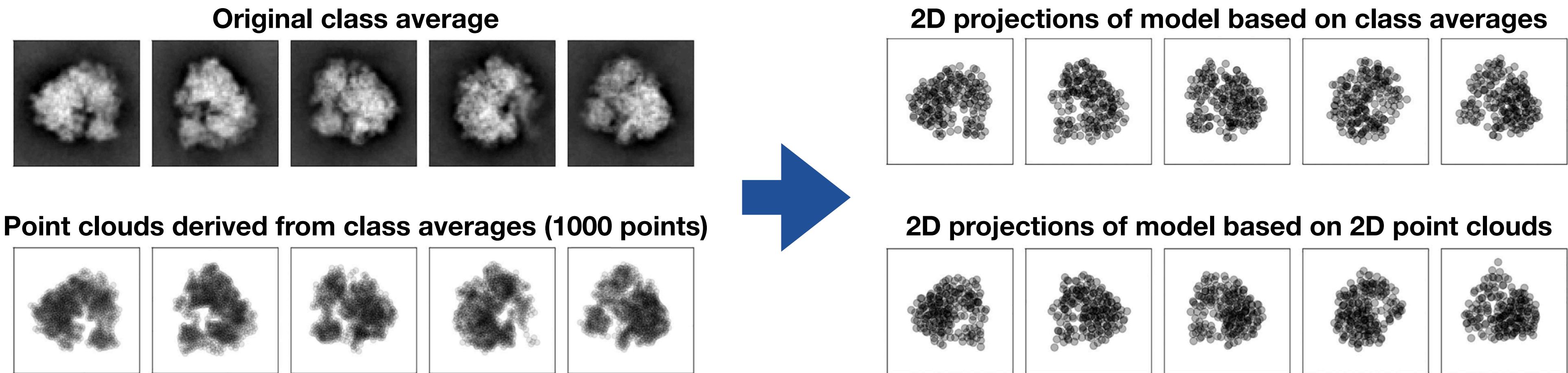
#particles	resolution (FSC at 0.143) [\AA]		RMSD [\AA]	
	point clouds	images	point clouds	images
50	12.2	12.2	15.6	15.5
100	10.5	10.5	12.1	11.7
200	7.8	9.3	9.8	9.1
300	8.3	8.3	8.7	7.8
500	7.1	6.5	7.3	6.4
700	6.5	6.2	6.4	5.6
800	6.0	5.8	6.1	5.3
900	6.0	4.8	5.9	5.0
1000	5.8	5.0	5.7	5.0

- Studied the performance of sampling particle positions by fixing the rotations to the correct values and sampling only the particle positions and the precisions of the projection data.
- A higher number of particles K results in more flexible models that result in a better goodness of fit and higher precision → **HMC is highly suited to sample particle configurations.**
- With increasing number of particles K , the gap in accuracy widens but is still similar to the maximum attainable value.

3D Reconstruction by Sampling the Full Posterior Distribution

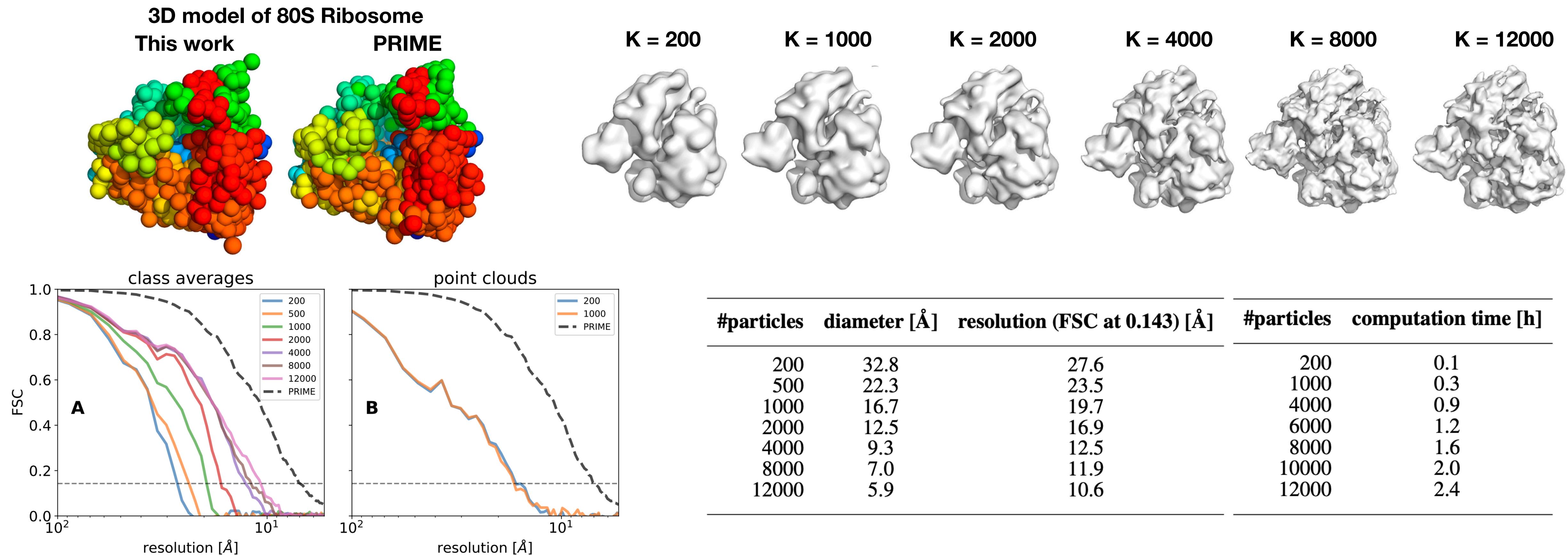
- Applied Bayesian random tomography to real cryo-EM datasets:
 - ▶ 400 of 2D class averages of the 80S ribosome → only first 50 class averages have been used in this test.
- Sampled the joint posterior distribution of all unknown parameters (x_k, R_n, ξ) with the MCMC techniques.
- The reconstruction simulations started from spherical random structures and random rotations and did not observe any dependence on the initial values.

Reconstruction – 80S Ribosome



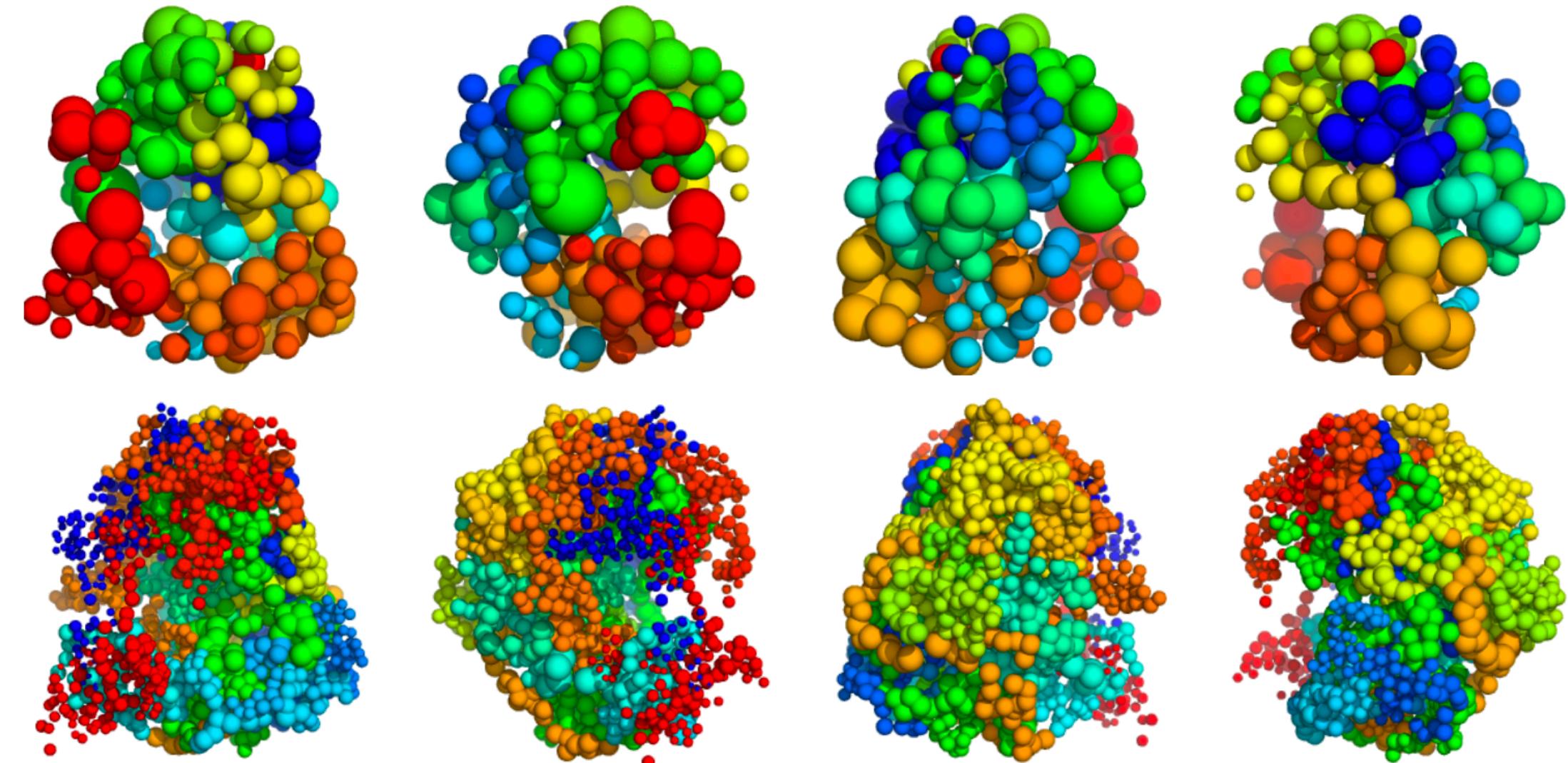
- 50 of 2D class averages of the 80S ribosome computed with SIMPLE2 — image size is 80×80 pixels, pixel size is 2.68 \AA .
- Class averages were converted to 2D point clouds each composed of 1,000 points.
- Used $K = 200$ and $K = 1000$ particles with a radius of $R = 16.4$ and $R = 8.4 \text{ \AA}$ to fit the ribosome 3D and 2D point clouds.
- Ran 500 iterations of Gibbs sampling for the rotational parameters and HMC for the particle positions.
- Observe a good agreement between the experimental point clouds and the model point clouds with an RMSD:
 - ▶ Base on point clouds: $\text{RMSD} = [6.4, 9.8] \text{ \AA}$ and an average of $7.7 \pm 0.7 \text{ \AA}$.
 - ▶ Base on class average: $\text{RMSD} = [6.1, 13.1] \text{ \AA}$ and an average of $8.3 \pm 3.0 \text{ \AA}$.

Reconstruction – 80S Ribosome



- The resolution of the particle models ranges from 27.6 Å (200 particles) to 10.6 Å (12,000 particles).
- The reconstruction obtained with SIMPLE reaches a resolution of 6.2 Å based on 200 class averages.

Uncertainty Quantification



Model based on 200 particles

Model based on 2000 particles

- The posterior samples can be used to assess the uncertainty of the particle models in the form of structural error bars.
- To carry out uncertainty quantification, the particle models first need to be superimposed and a correspondence between particles across different samples has to be established.
- Using the Iterative Closed Point (ICP) method followed by a linear assignment step where particle distances between superimpose clouds are used as a cost.
- The size of the spheres in uncertainty quantification of the ribosome model is proportional to the standard deviation of particle positions after superposition and assignment.

Summary and Discussion

- This paper outlined a Bayesian approach to random tomography, the problem of reconstructing a 3D structure from 2D views along unknown random directions.
- The core of this approach is a representation of 3D volumes using a radial basis function kernel whose centers are the main inference parameters.
- This paper demonstrated that coarse-grained models can be inferred from projection data (images or point clouds) with MCMC algorithms such as HMC and global sampling of the rotations.
- The current application only limited on class averages (high SNR) without CTF effect, to be checked the application on raw images and also including CTF effect.
- In all applications discussed in this paper, the number of particles K was fixed. An interesting question for future research is to estimate the number of particles based on the projection data.

Back Up

Gibbs Sampling

- Gibbs sampling is applicable when the joint distribution is not known explicitly or is difficult to sample from directly, but the conditional distribution of each variable is known and is easier to sample from.
- Suppose we want to obtain k samples of $X = (x_1, \dots, x_n)$ from a joint distribution $p(x_1, \dots, x_n)$. Denote the i -th sample by $X^{(i)} = (x_1^{(i)}, \dots, x_n^{(i)})$, the Gibbs Sampling proceed as following steps:
 - ▶ Begin with initial value $X^{(0)}$.
 - ▶ To sample $x_j^{(i+1)}$, we update it according to the distribution specified by $p(x_j^{(i+1)} | x_1^{(i+1)}, \dots, x_{j-1}^{(i+1)}, x_{j+1}^{(i)}, \dots, x_n^{(i)})$.
 - ▶ Repeat the above step k times.