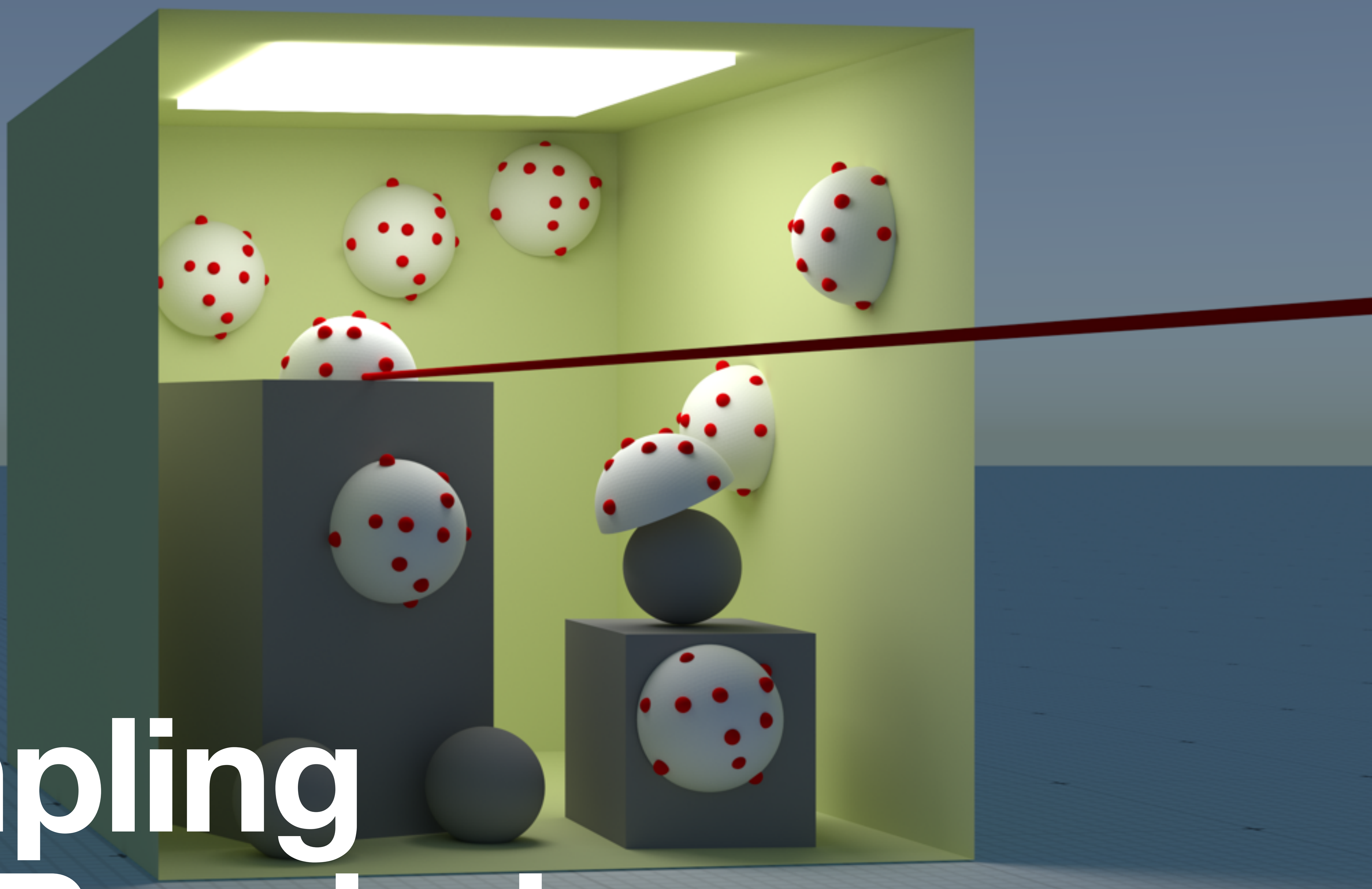
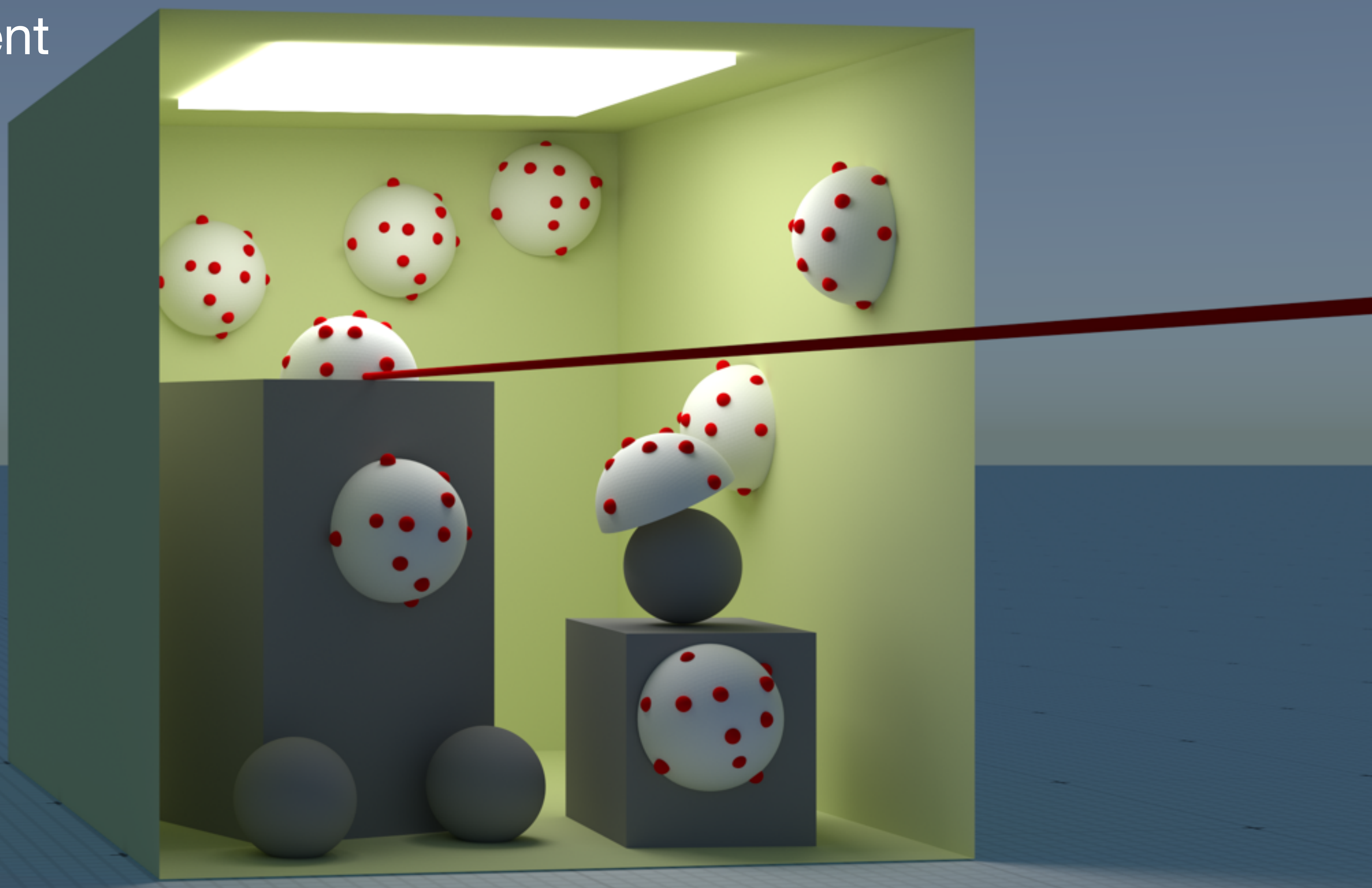


Correlated Sampling in Monte Carlo Rendering

David Coeurjolly

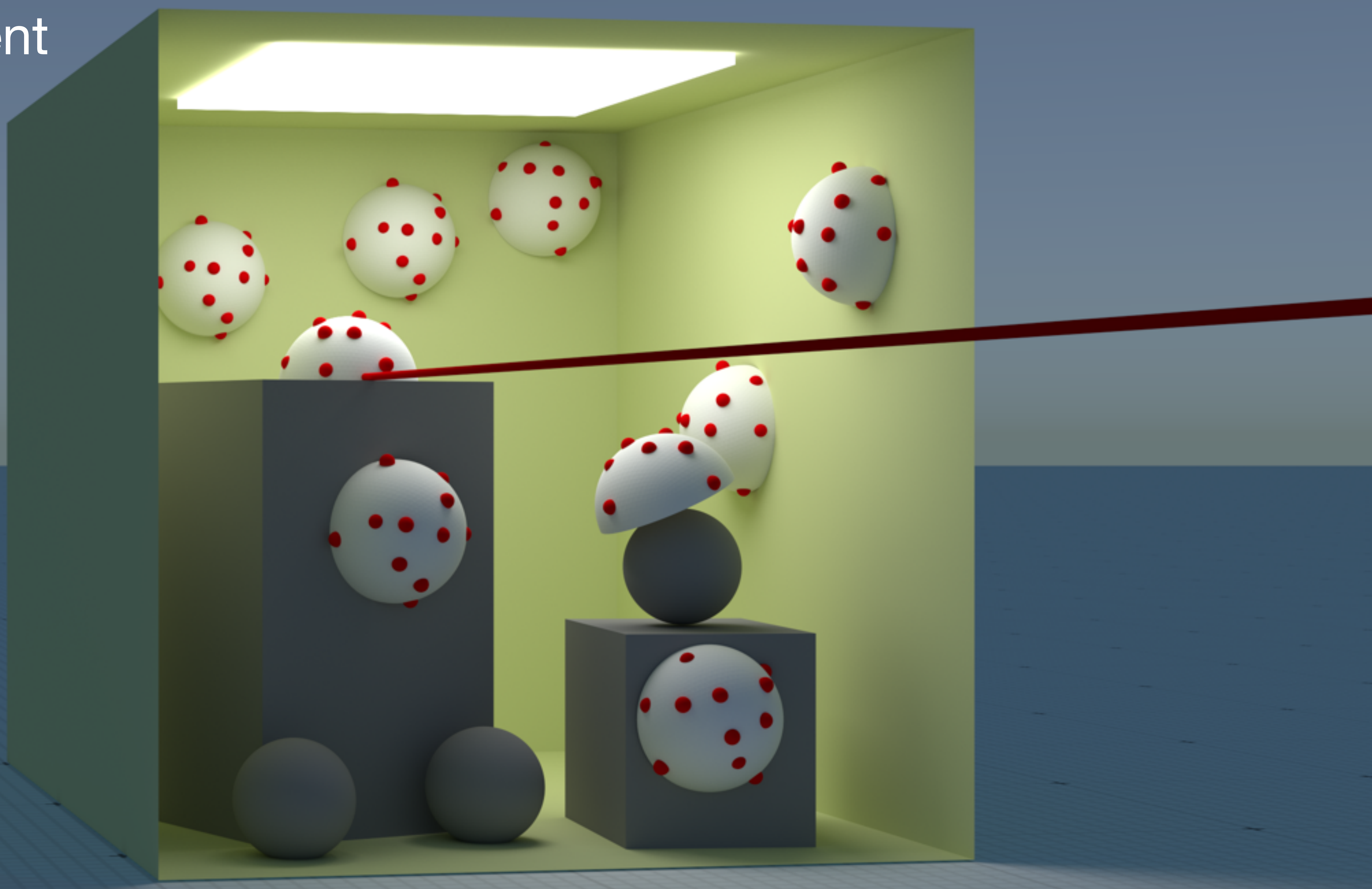
Victor Ostromoukhov, Nicolas Bonneel, Jean-Claude Iehl, Alex Keller, Mathieu Desbrun, Fernando de Goes, Pat Hanrahan, Kartic Subr, Eric Heitz, Laurent Belcour, Wojciech Jarosz...

Loïs Paulin, Hélène Perrier, Gurprit Singh, Adrien Pilleboue, Florent Wachtel, Feng Xie, Antoine Webanck...

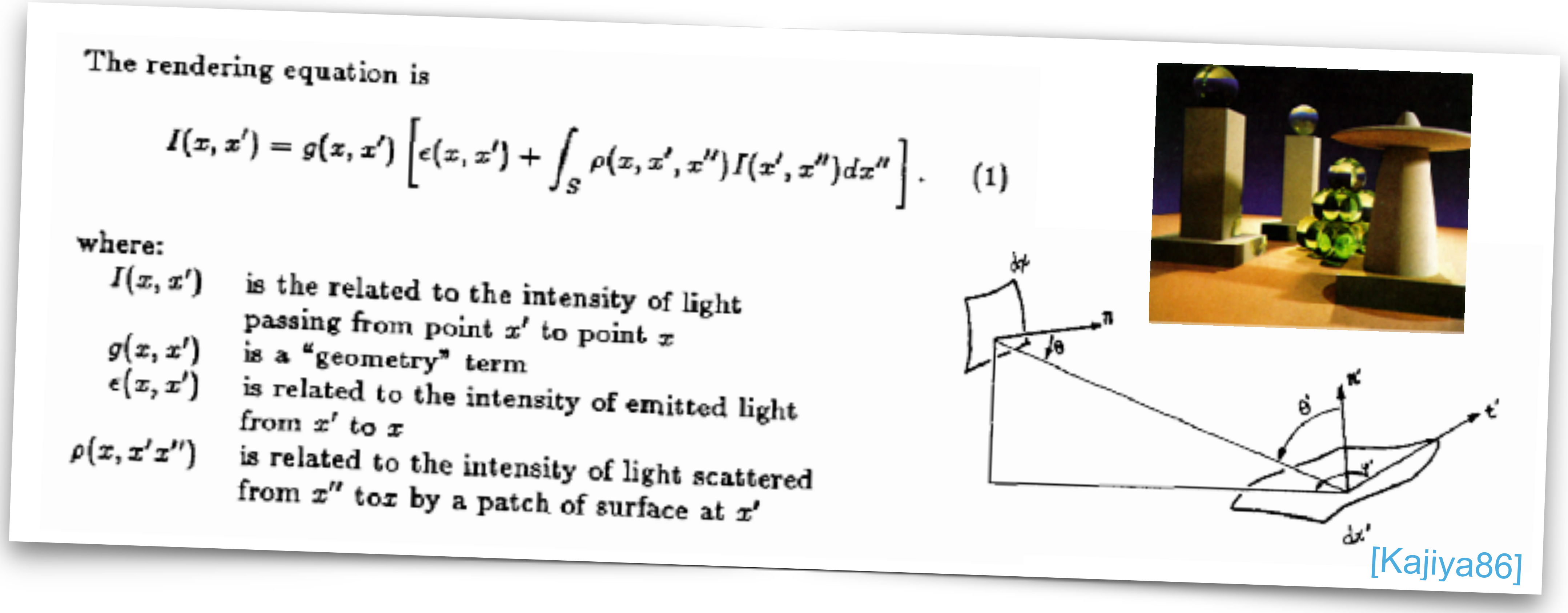


Victor Ostromoukhov, Nicolas Bonneel, Jean-Claude Iehl, Alex Keller, Mathieu Desbrun, Fernando de Goes, Pat Hanrahan, Kartic Subr, Eric Heitz, Laurent Belcour, Wojciech Jarosz...

Loïs Paulin, Hélène Perrier, Gurprit Singh, Adrien Pilleboue, Florent Wachtel, Feng Xie, Antoine Webanck...



Monte Carlo Integration


The rendering equation is

$$I(x,x') = g(x,x')\left[\epsilon(x,x') + \int_S \rho(x,x',x'')I(x',x'')dx''\right]. \quad (1)$$

where:

$I(x,x')$ is the related to the intensity of light passing from point x' to point x

$g(x,x')$ is a "geometry" term

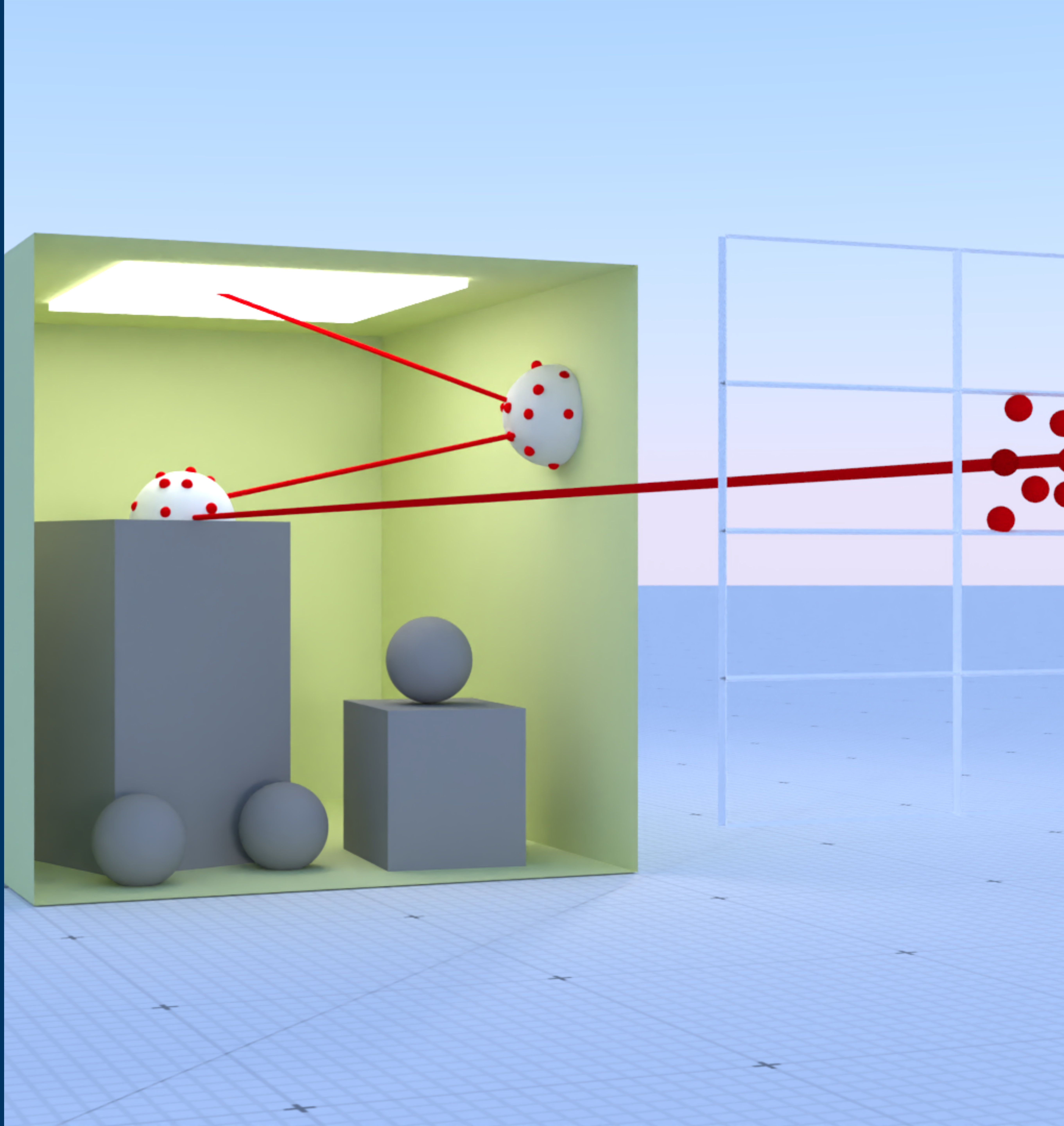
$\epsilon(x,x')$ is related to the intensity of emitted light from x' to x

$\rho(x,x'x'')$ is related to the intensity of light scattered from x'' to x by a patch of surface at x'

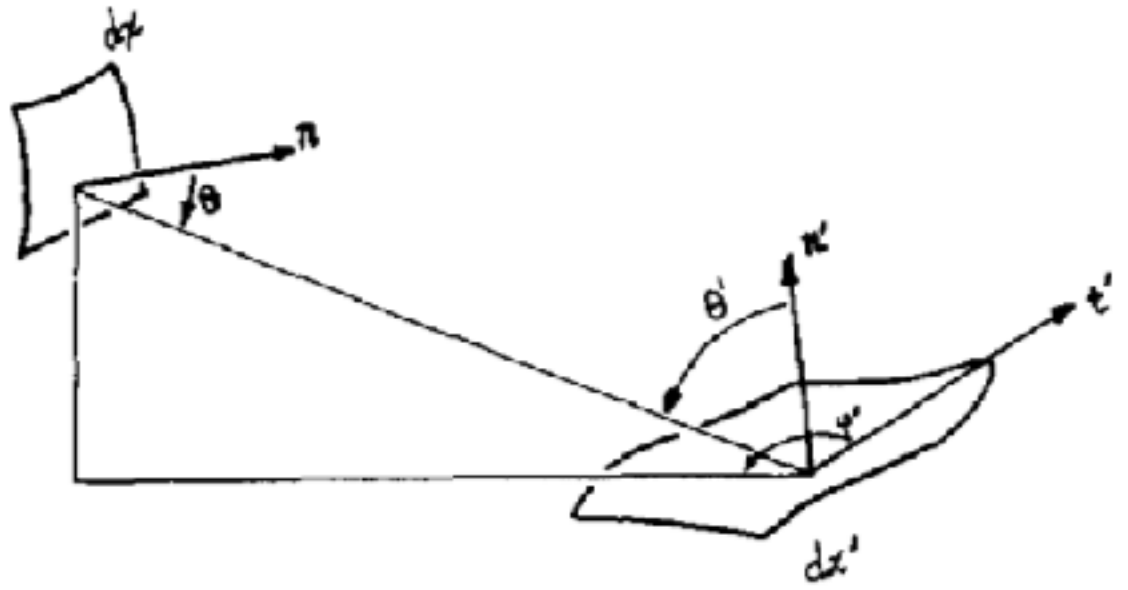
[Kajiya86]

$$\int_\Omega f(x)\,\mathrm{d}x \approx \frac{1}{|X|}\sum_{x\in X} f(x)$$

Monte Carlo Rendering



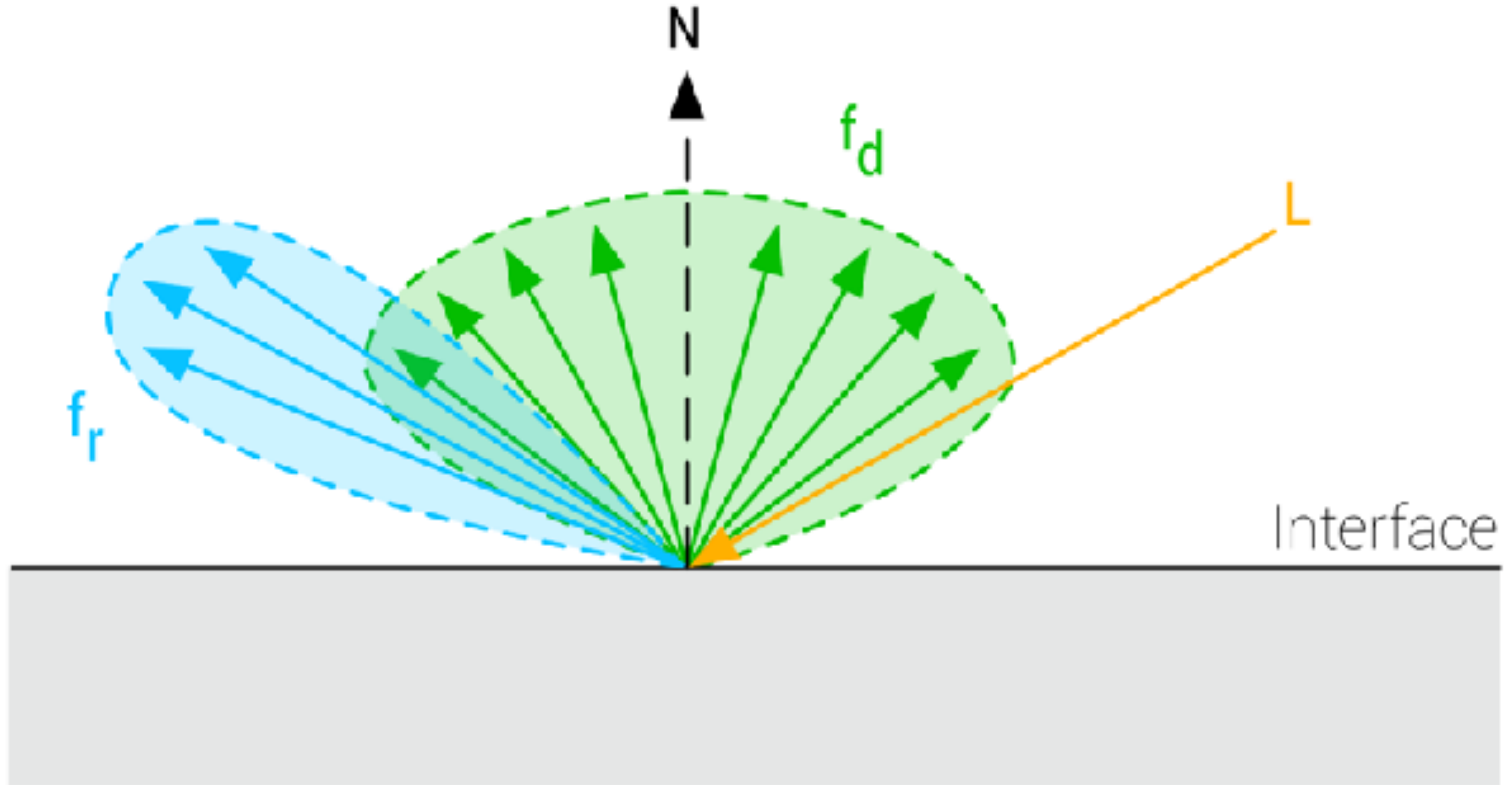
The Light Transport Equation



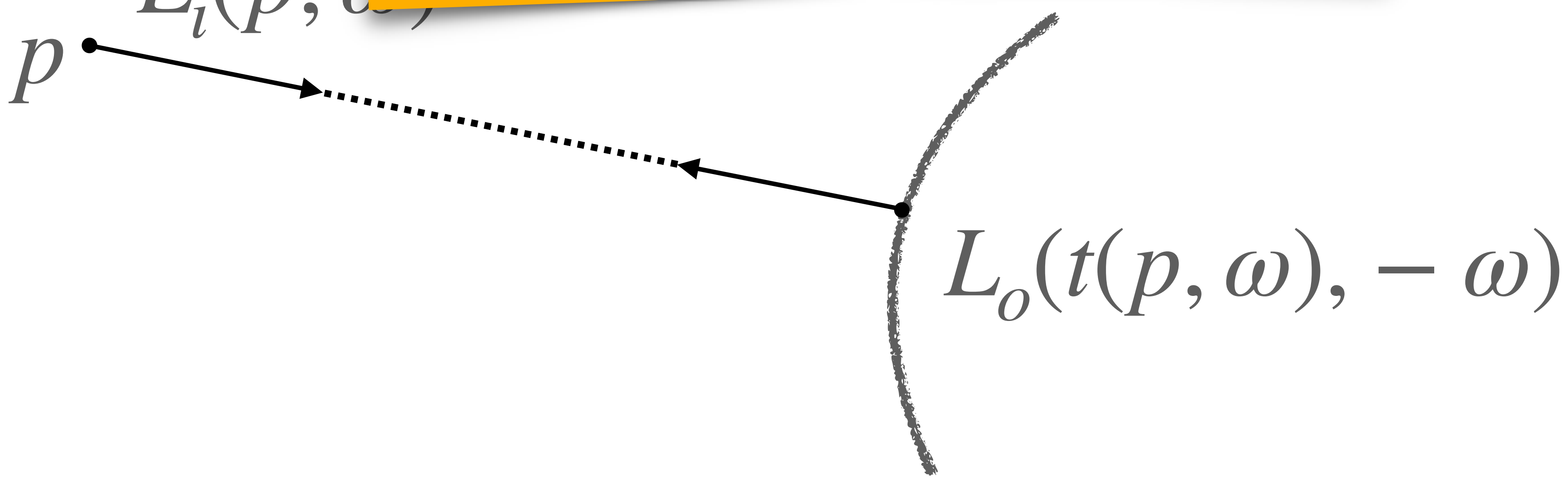
$$\boxed{L_o(p,\omega_o)} = \boxed{L_e(p,\omega_o)} + \int_{\mathbb{S}^2} \boxed{f(p,\omega_o,\omega_i)}\boxed{L_i(p,\omega_i)}\,|\,\omega_i \cdot \vec{n}\,|\,d\omega_i$$

$$L_i(p,\omega) = L_o(t(p,\omega), -\omega)$$

Recursive integral

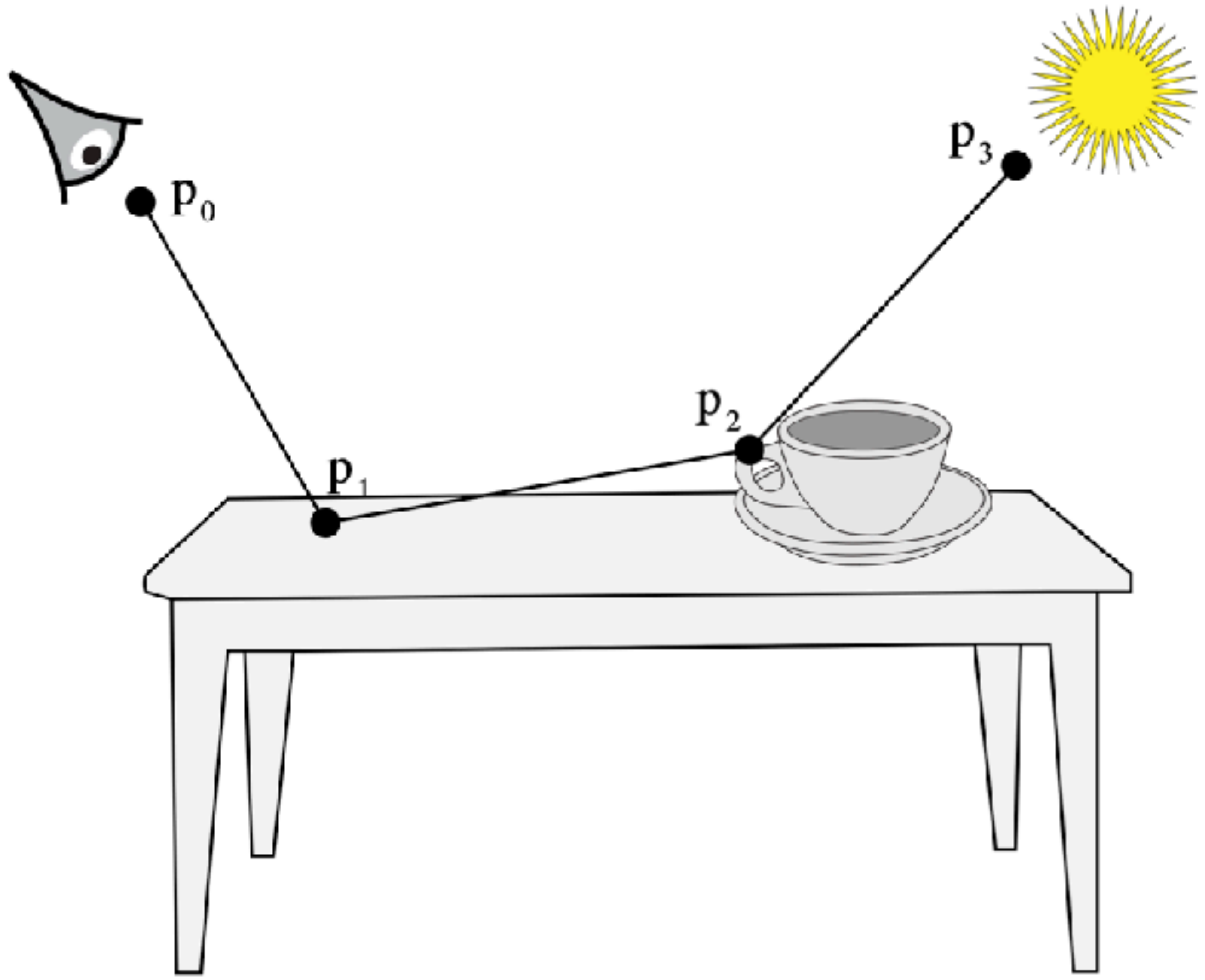


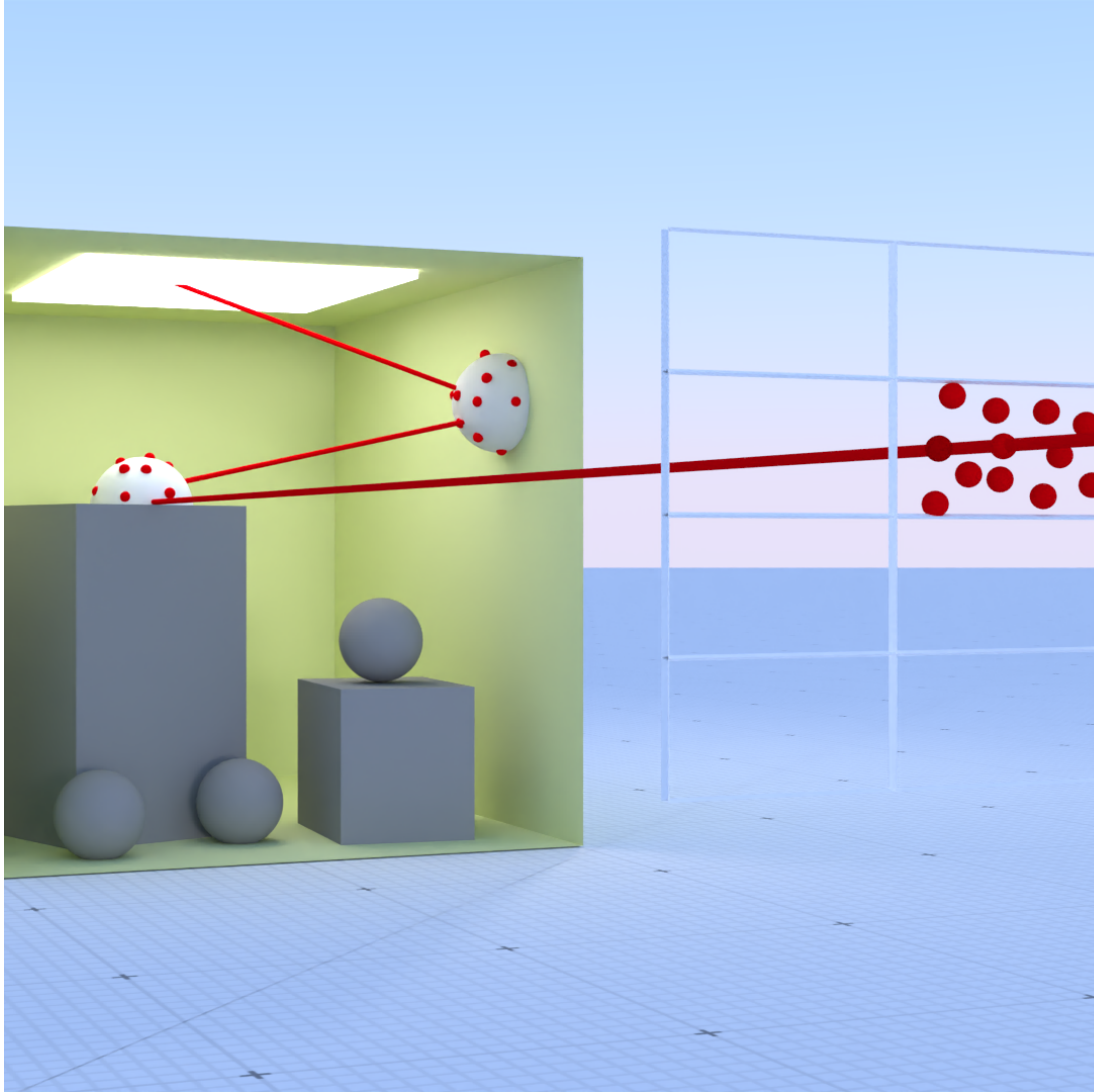
+ Visibility



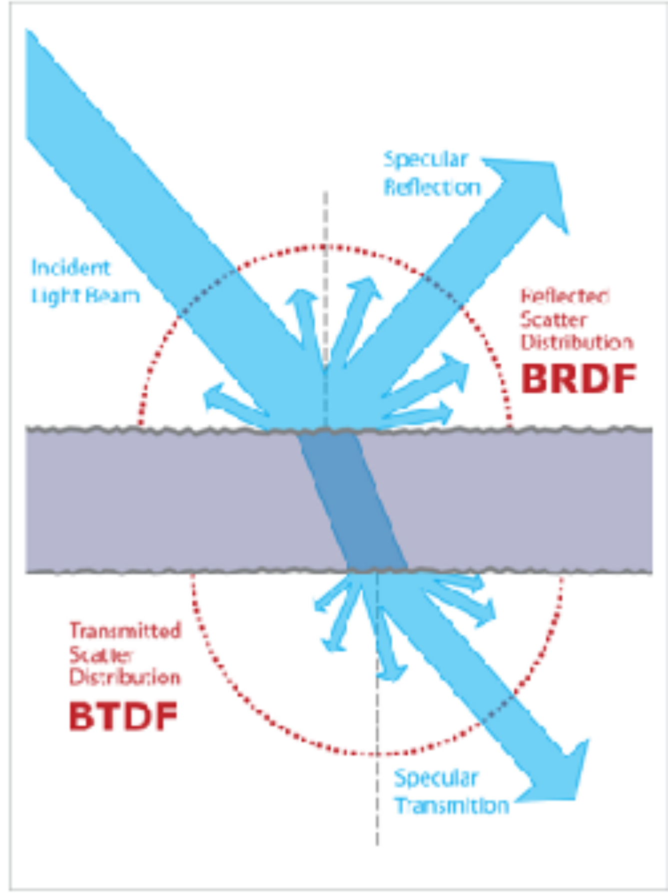
https://google.github.io/filament/Filament.html#materialsystem/standardmodel

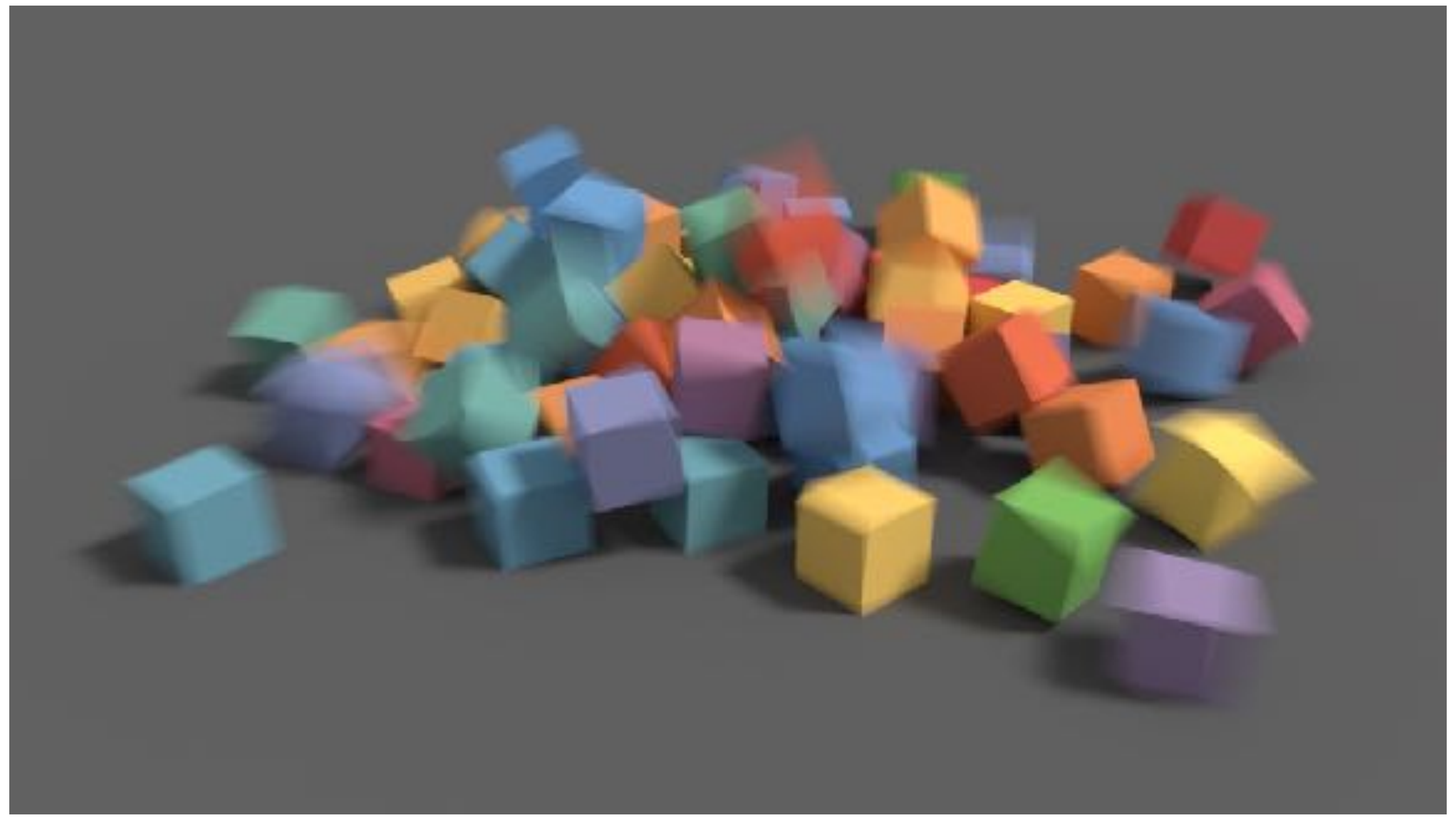
Path Sampling / Path tracing

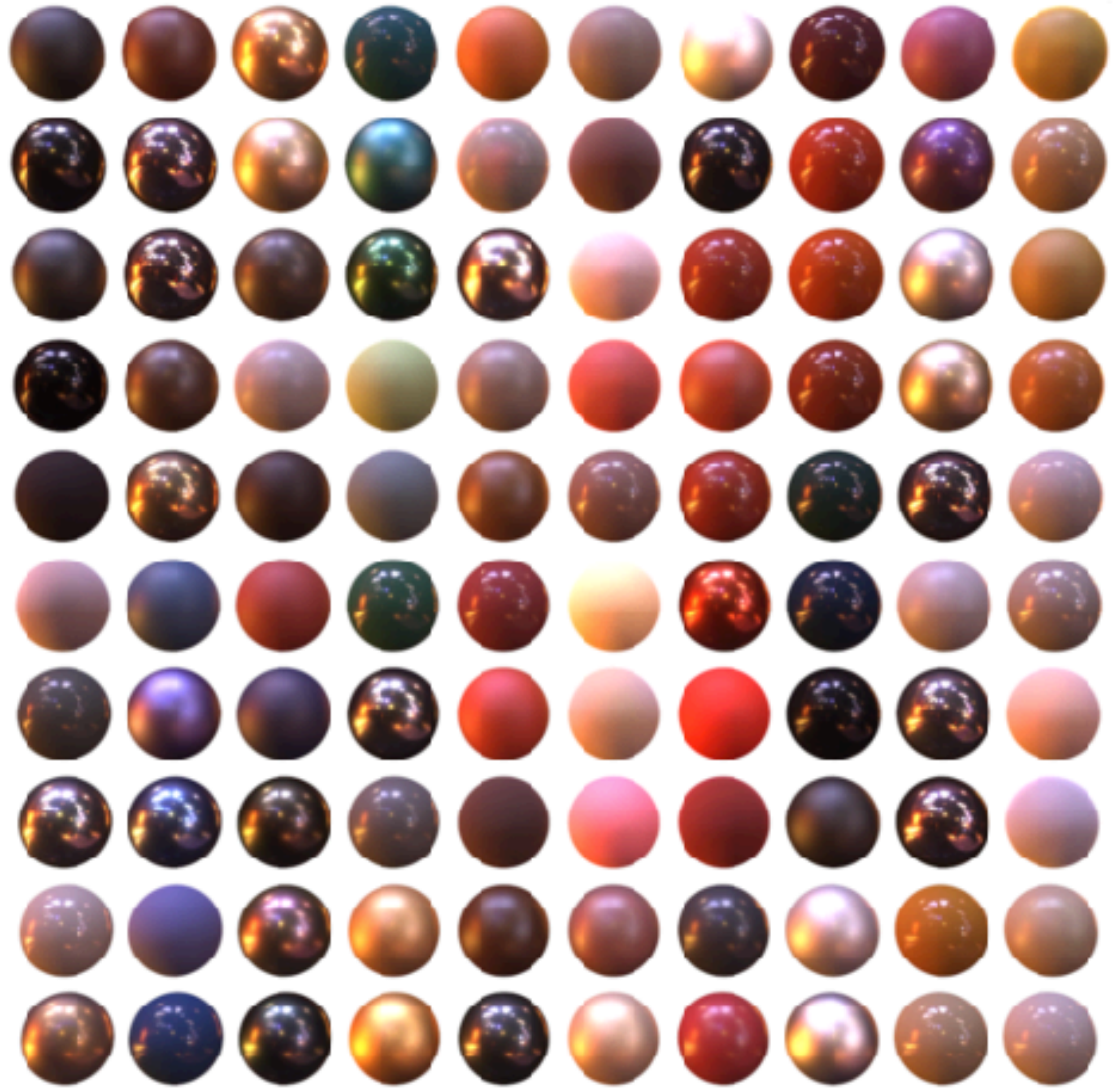


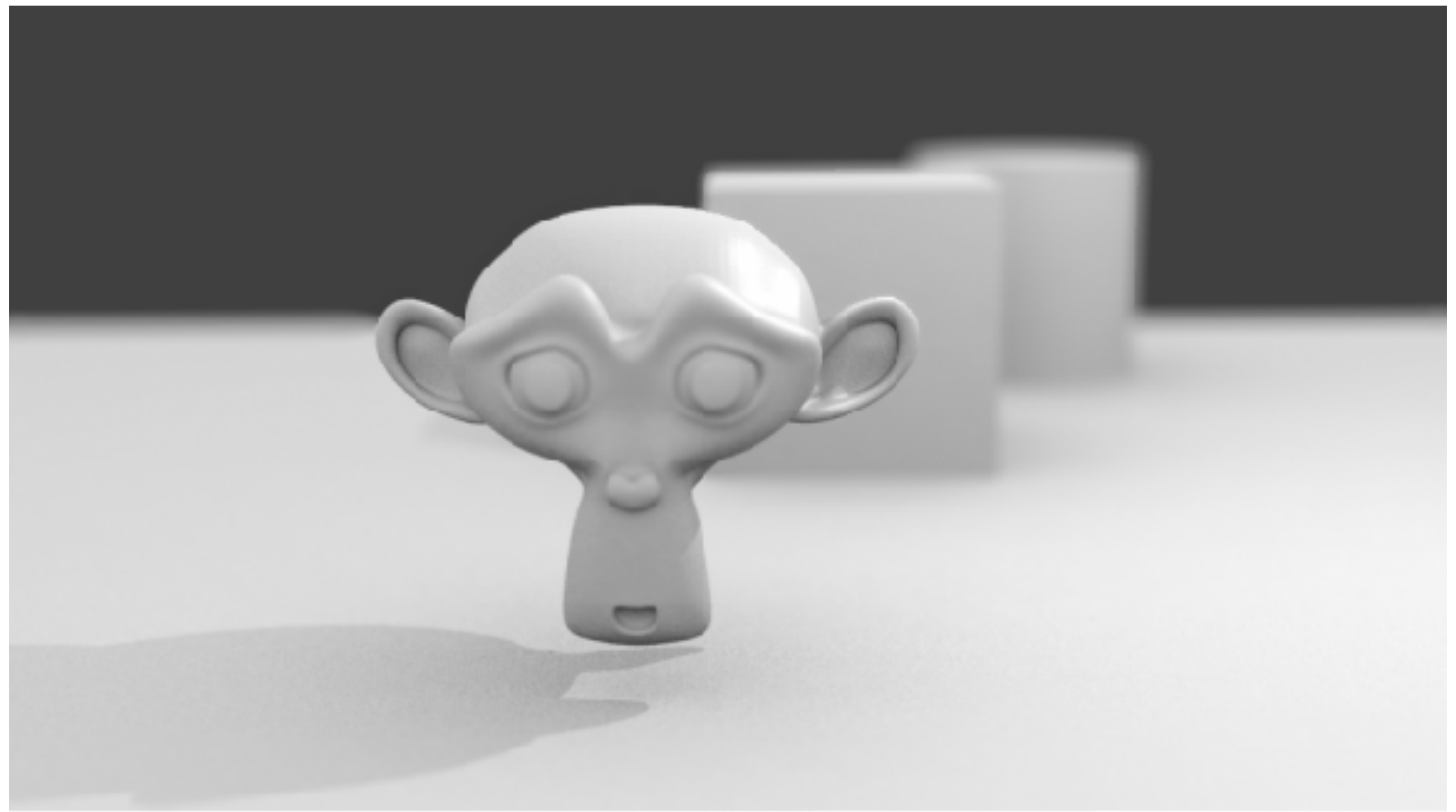


Material and physically based rendering

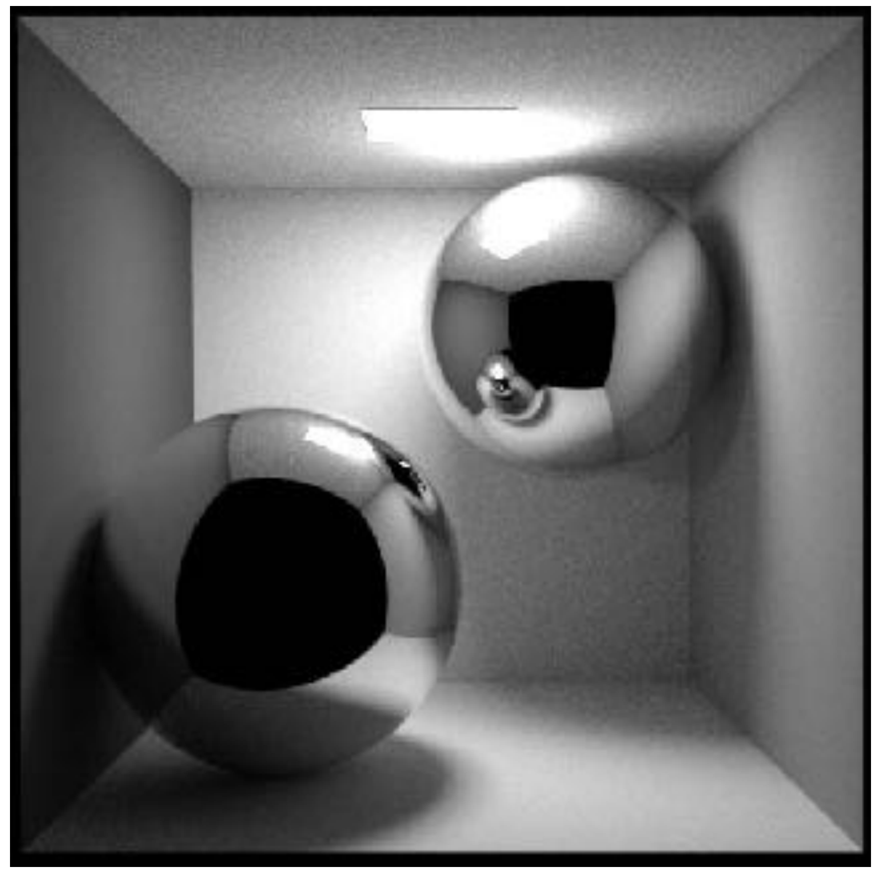


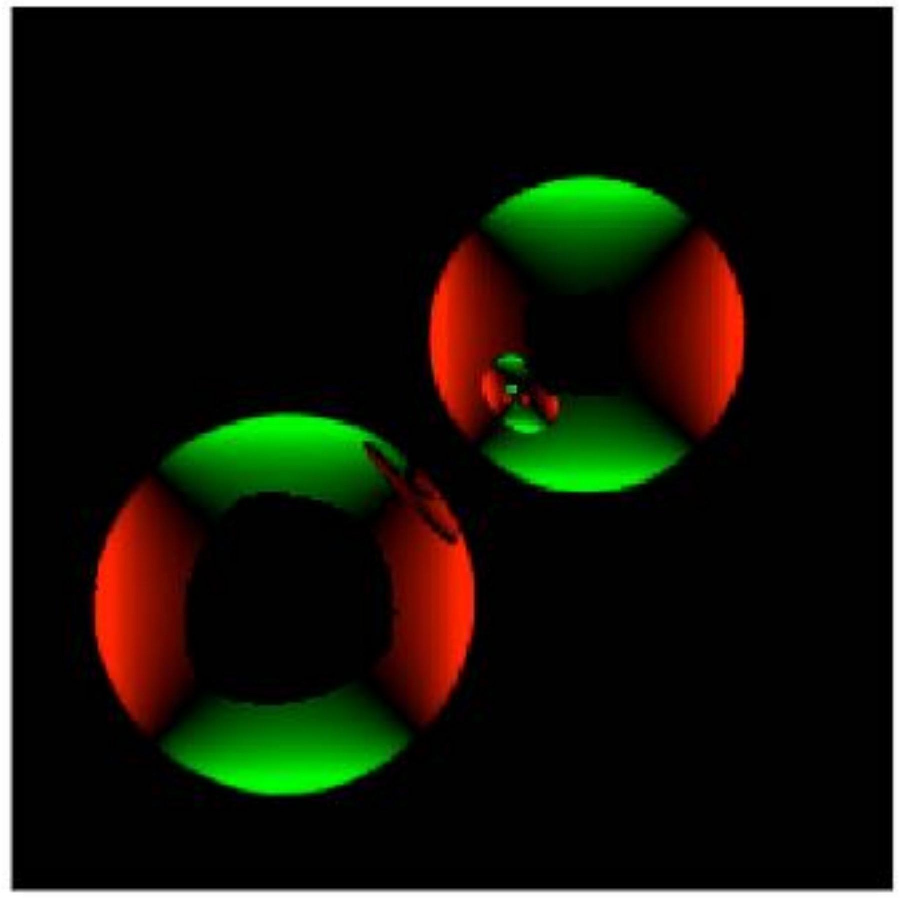


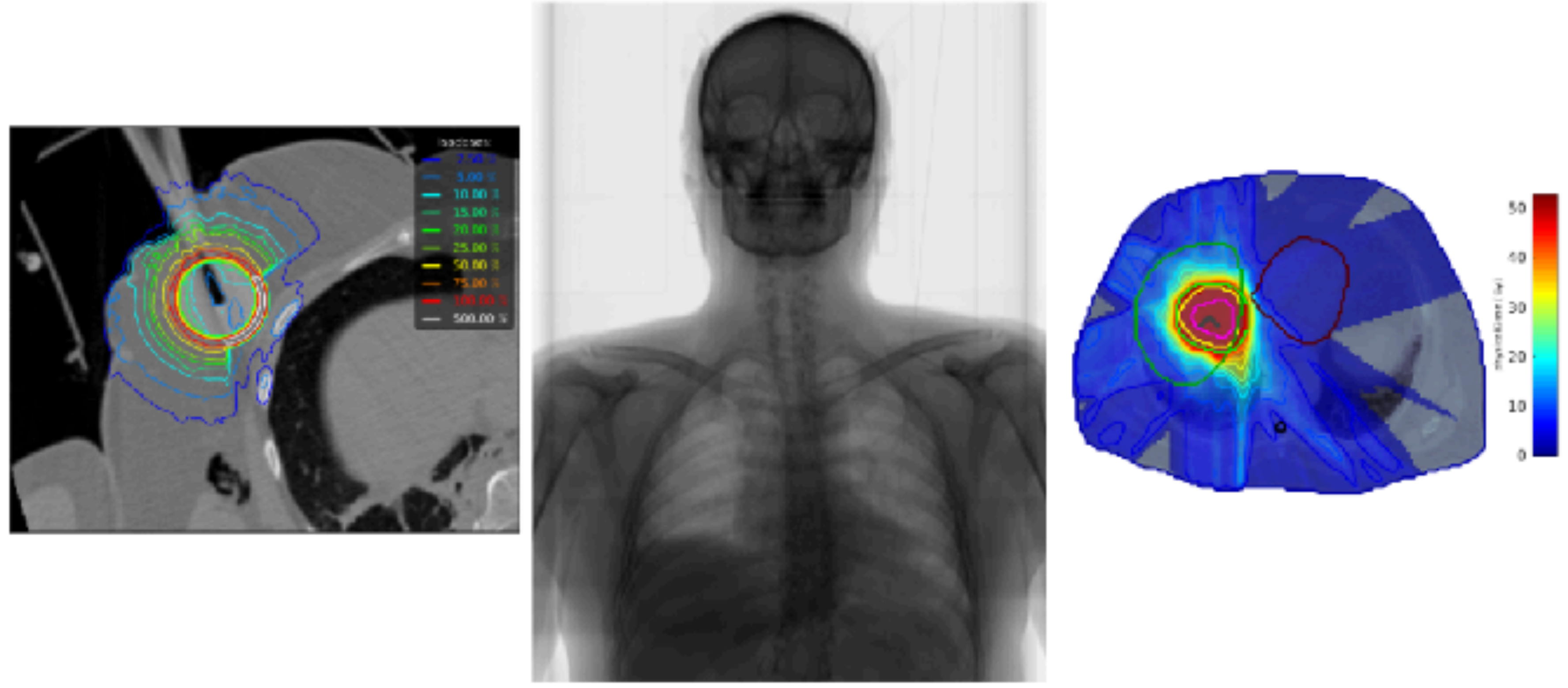




Spectral rendering, polarized rendering, high energy interactions...

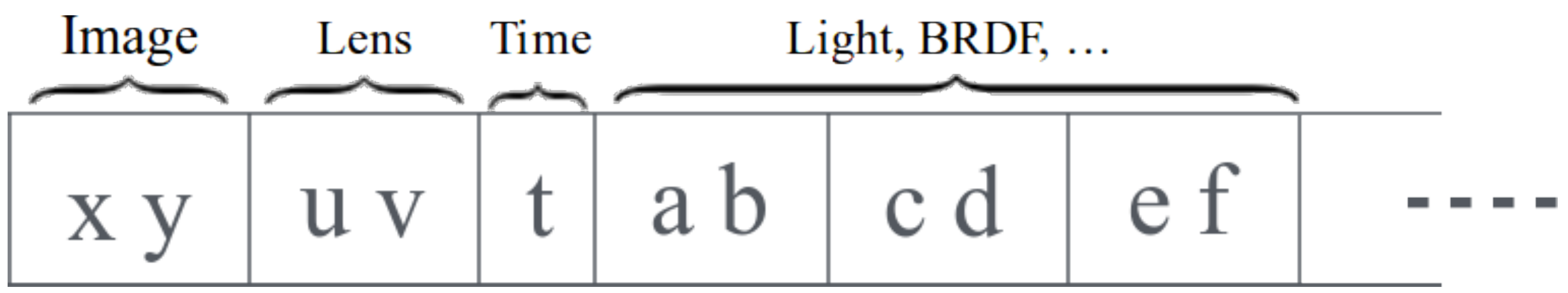






Canonical MC integration problem

- Samples in $[0,1)^s$

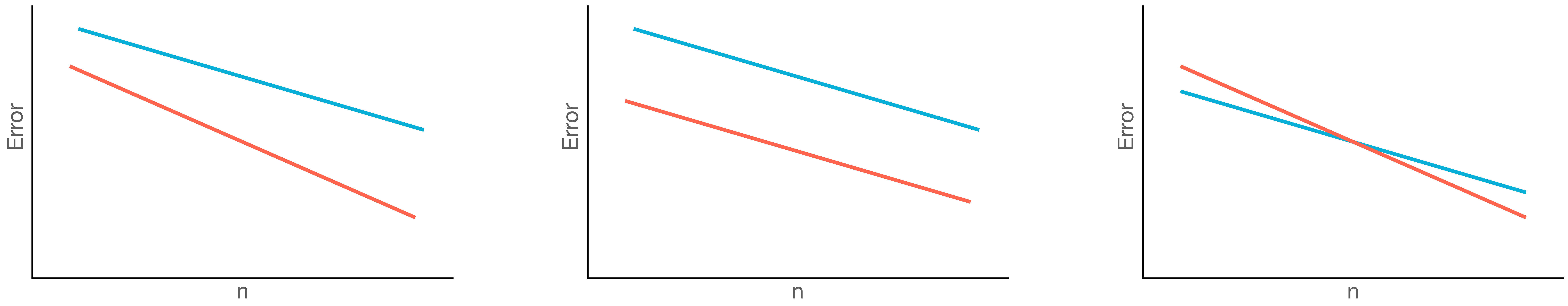


- Numerical integration

$$\mathcal{I} = \int_\Omega f(x)dx \qquad I_n = \frac{1}{n}\sum f(x_i)\,, \quad x_i \in \Omega = [0,1)^s$$

$$\Delta_n = |\,\mathcal{I} - I_n\,| \qquad \langle\Delta_n\rangle = \mathcal{I} - \langle I_n\rangle \qquad Var(I_n) = \langle I_n^2\rangle - \langle I_n\rangle^2$$

Convergence speed




1 spp


2 spp


4 spp


8 spp


16 spp


32 spp

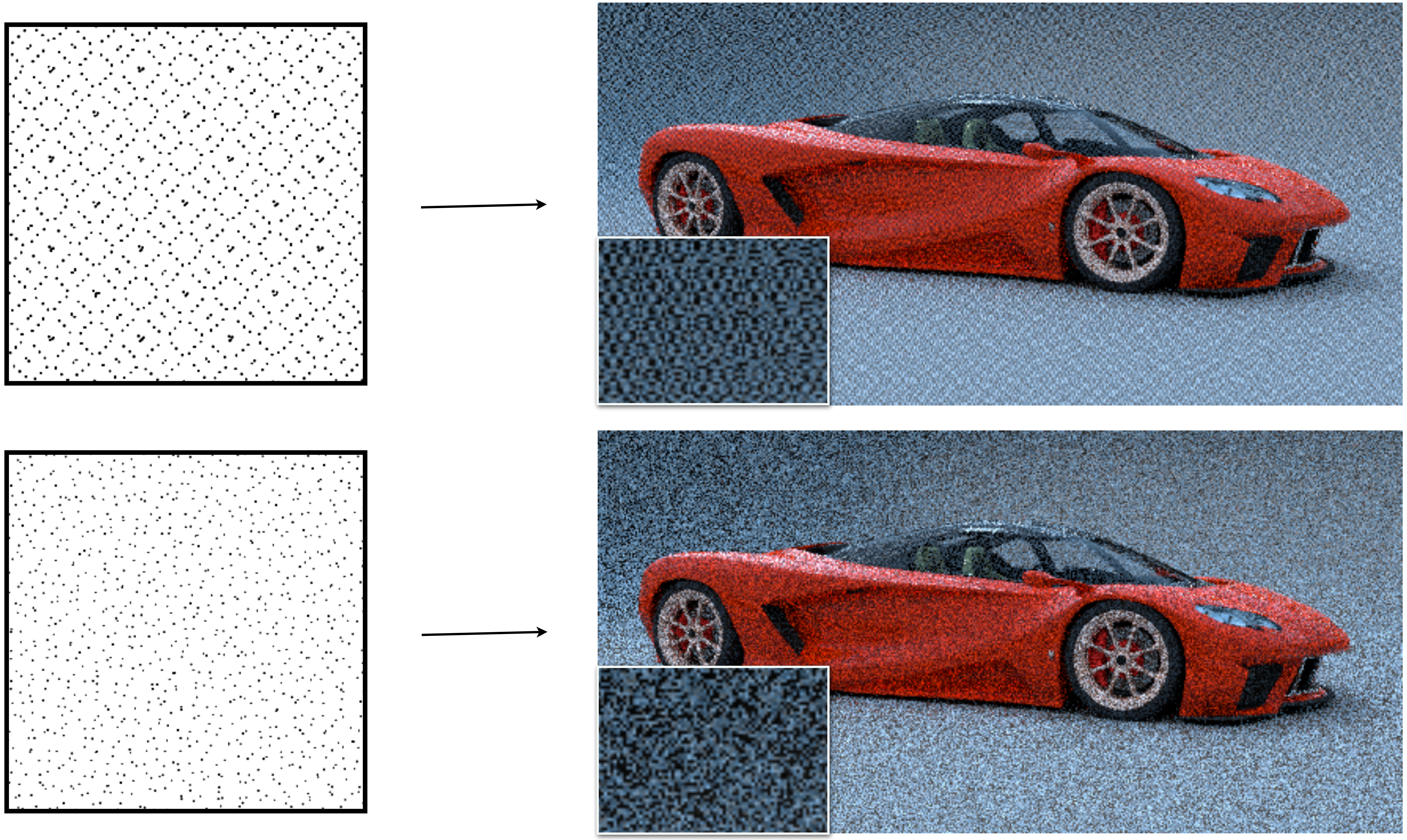

64 spp


128 spp

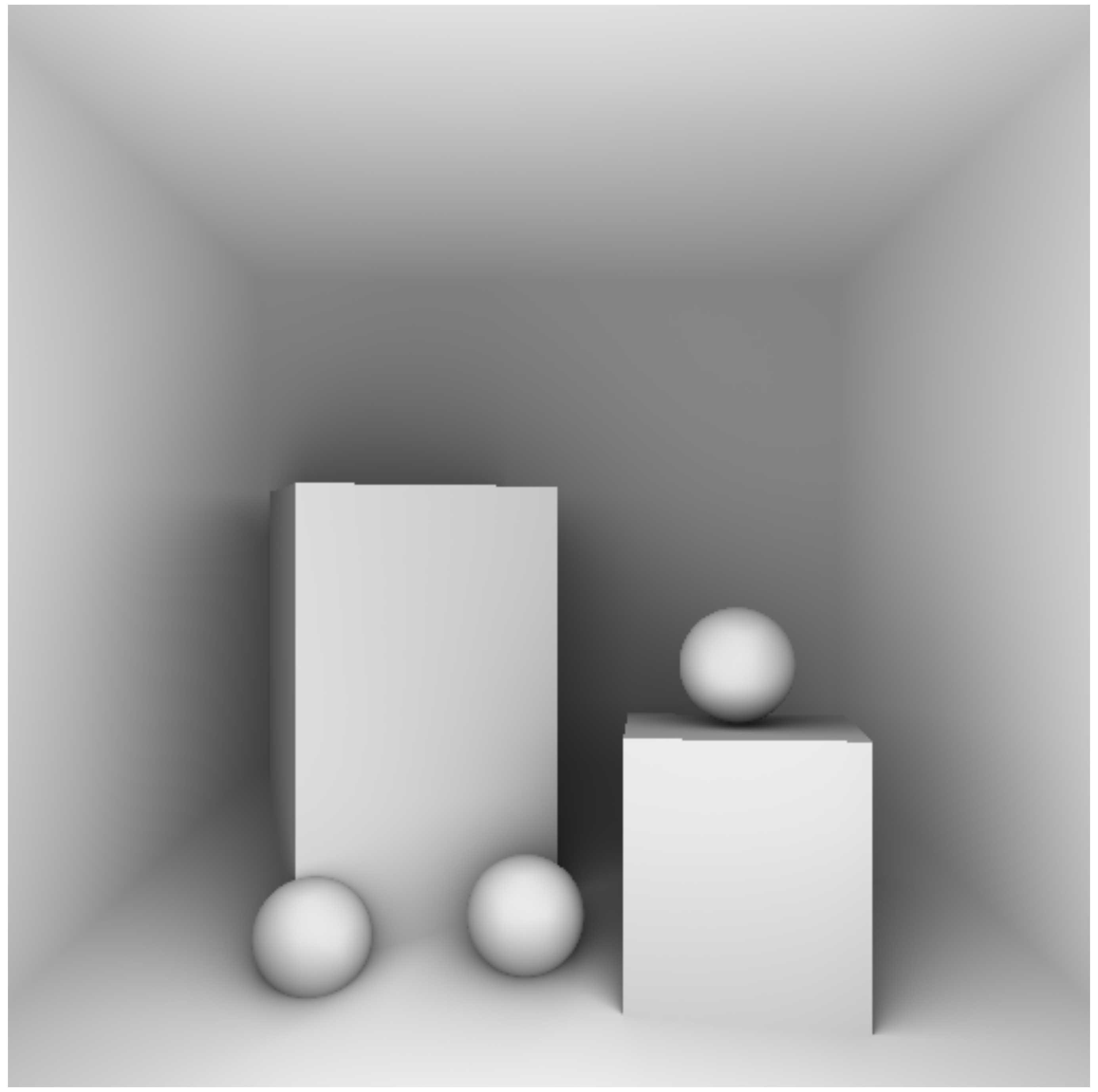

256 spp

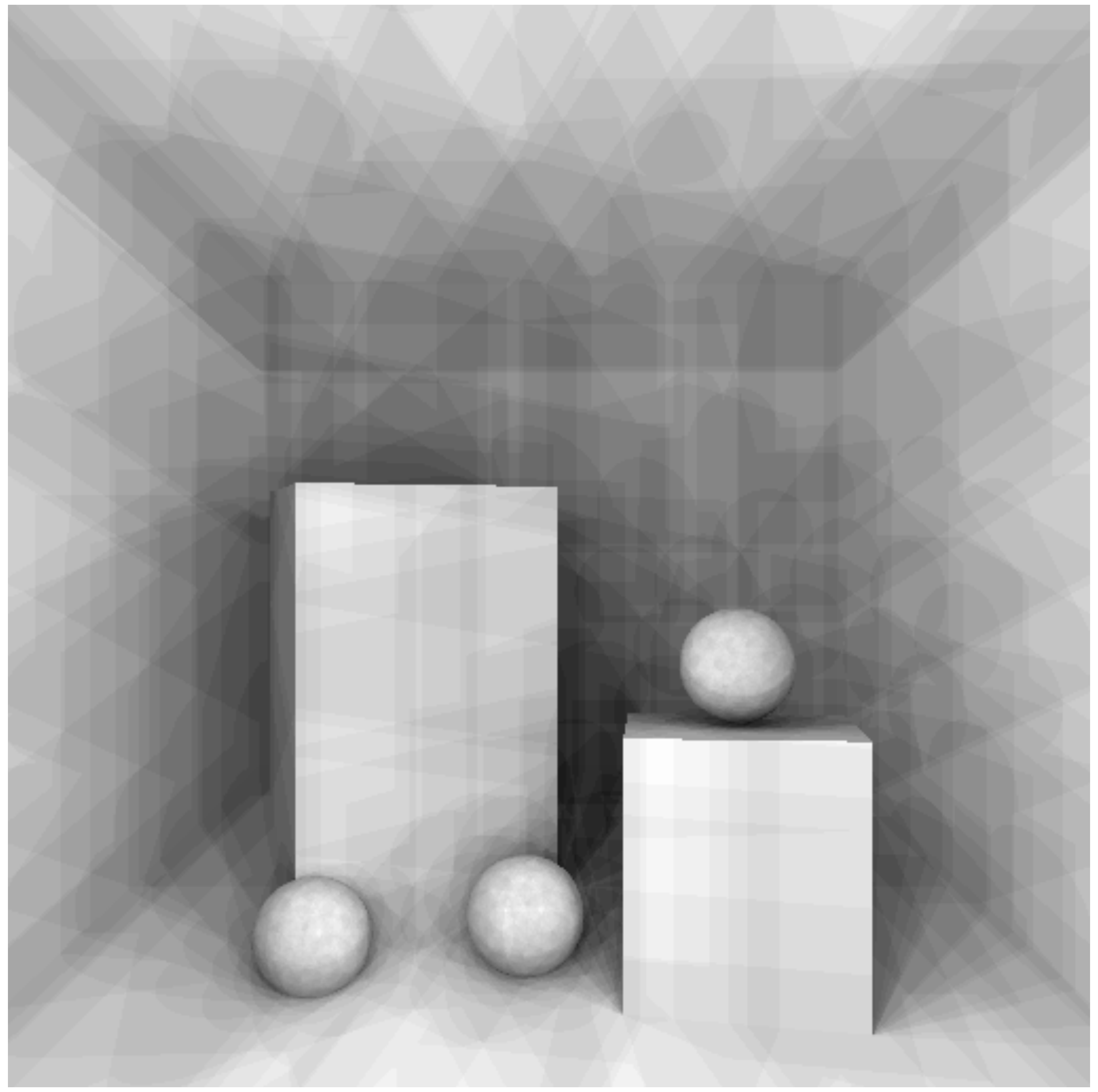

∞ spp

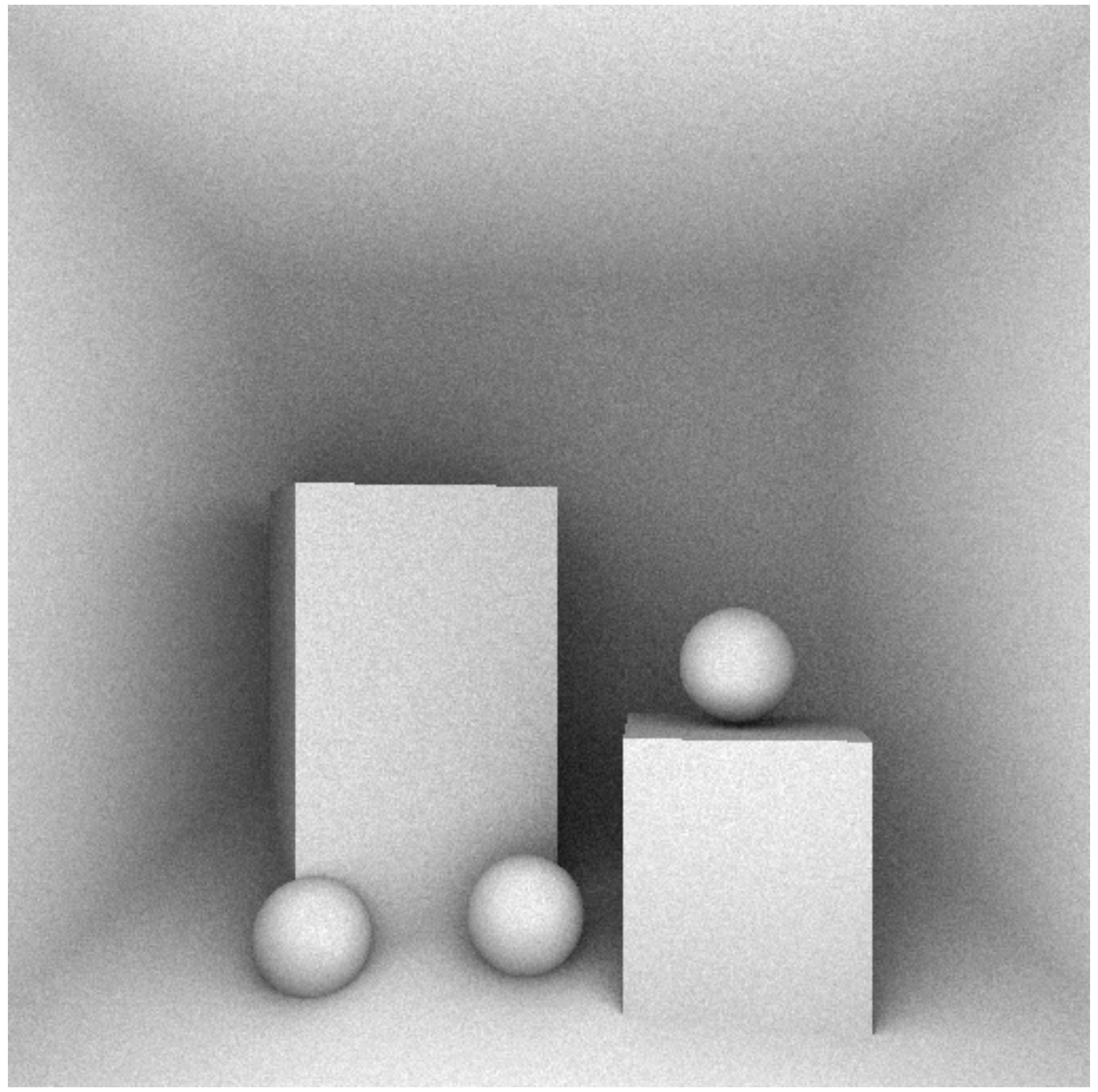
Noise vs aliasing



Noise vs aliasing





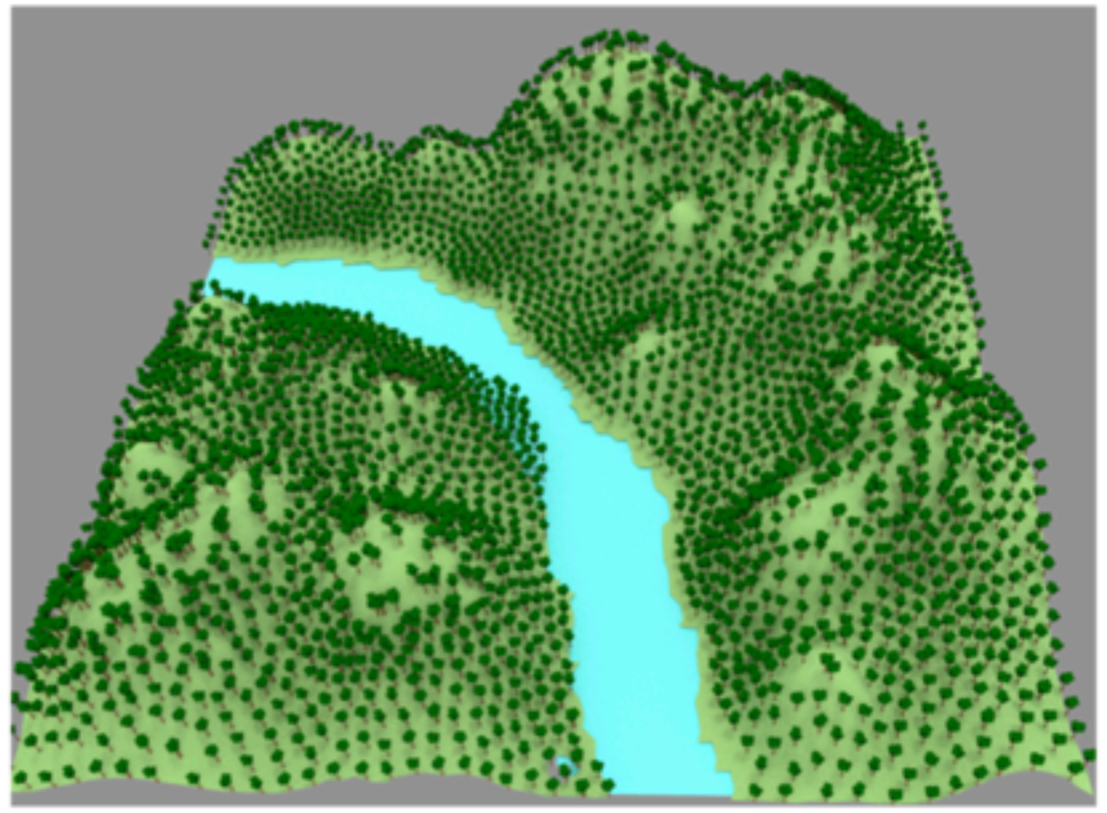


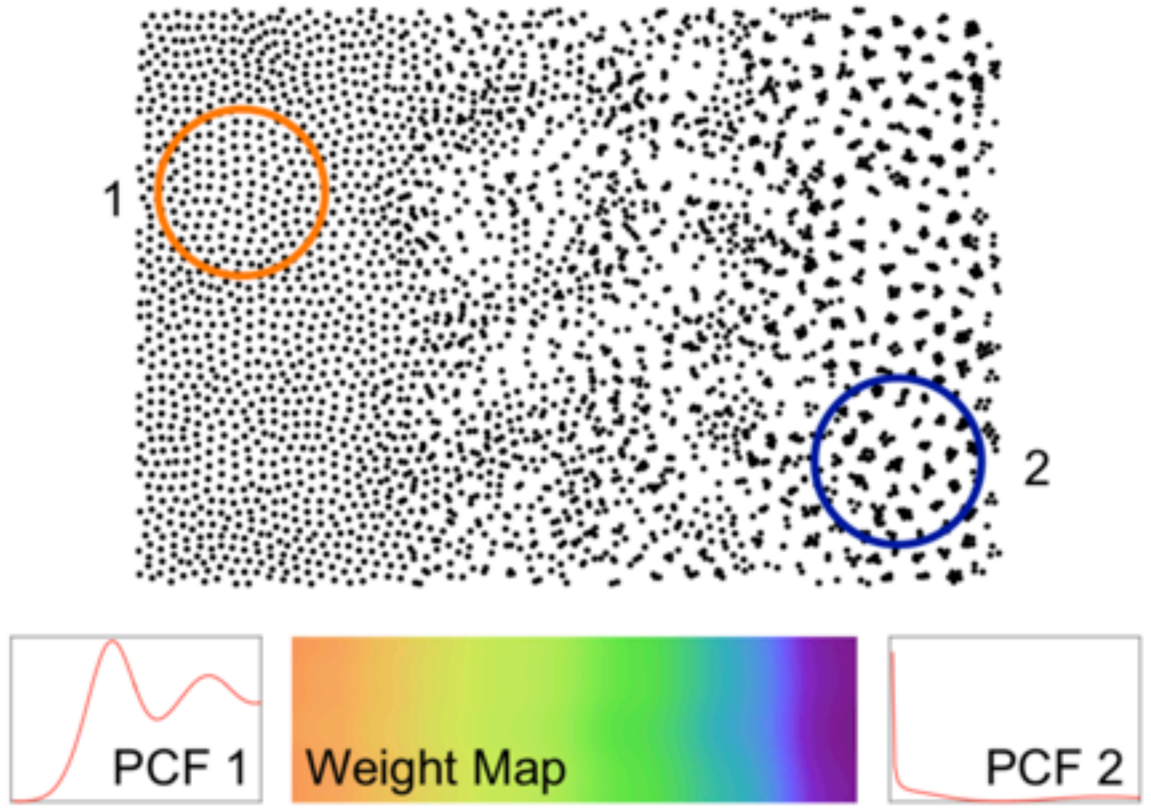
⇒ Stochastic point process or scrambling strategies

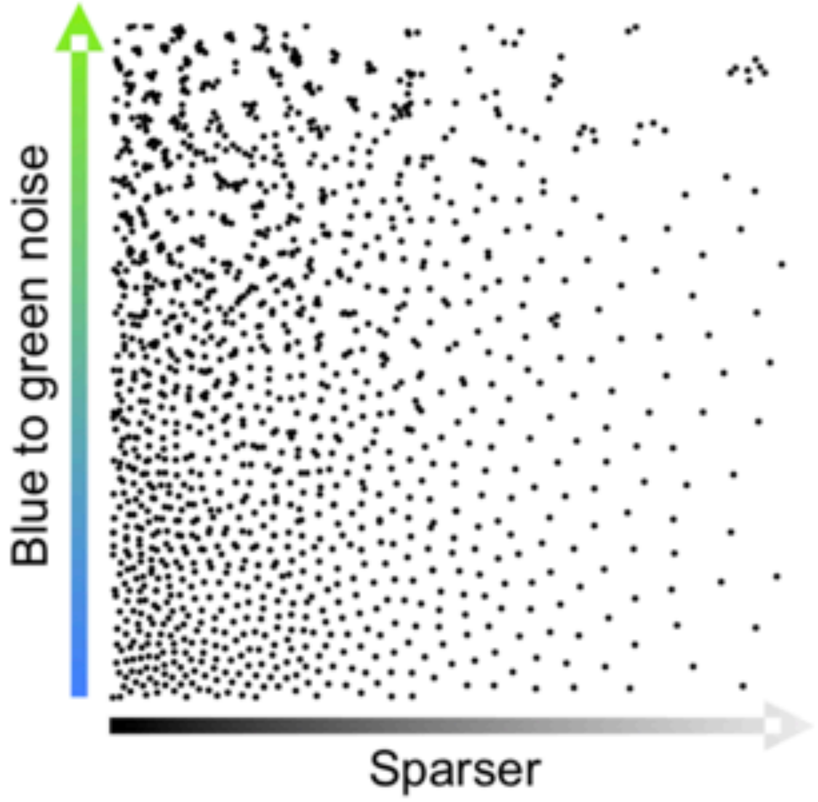
Wrap-up: MC Rendering

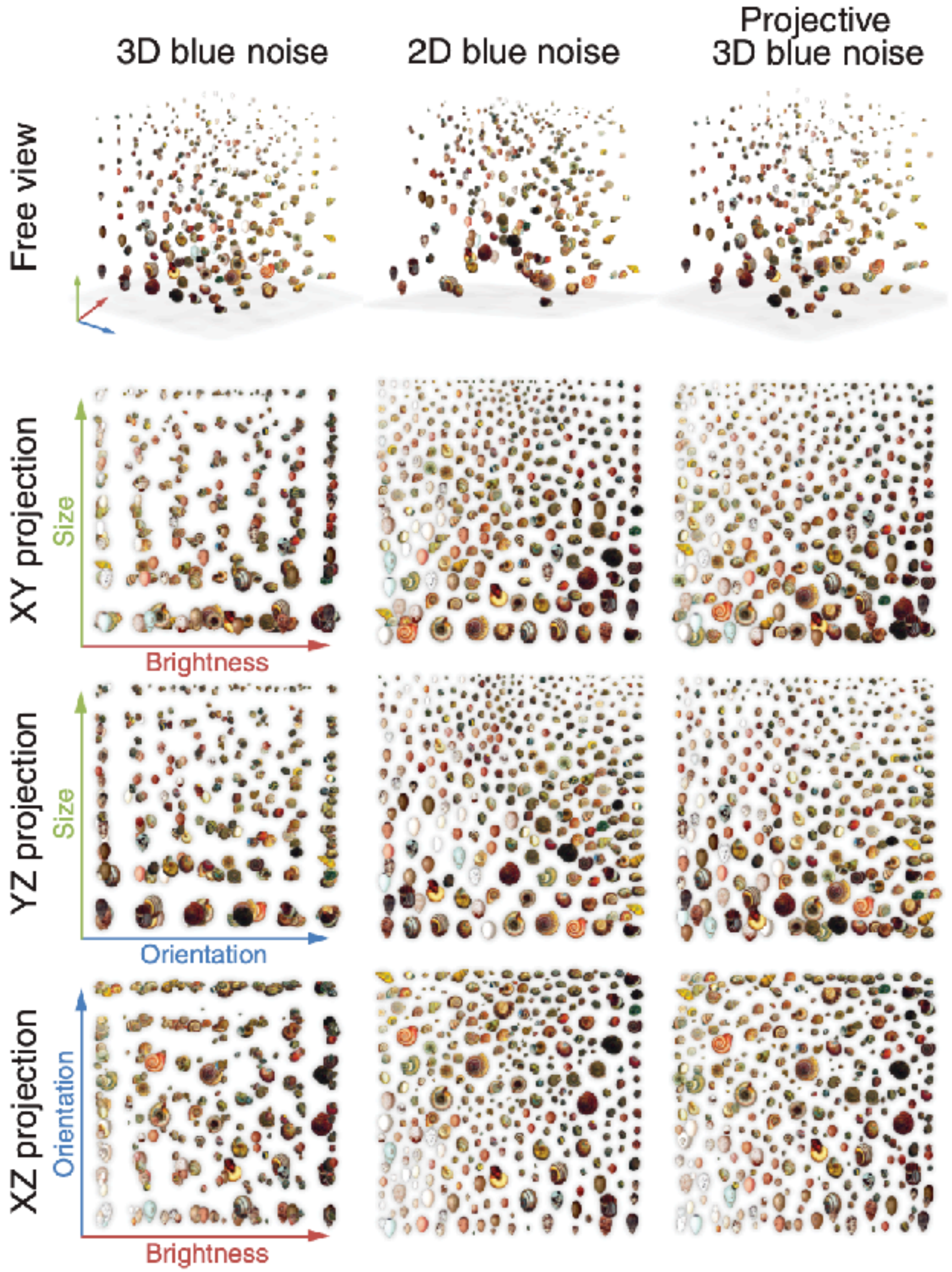
- Sampling in $[0,1)^s$ domains
- Stochastic samplers
- Best asymptotic variance reduction in MC/QMC
- Low error on low sample counts
- Rather limited number of dimensions (~40)
- Fast, adaptive and progressive sampler
- Some projective subspaces of the $[0,1)^s$ may be specific

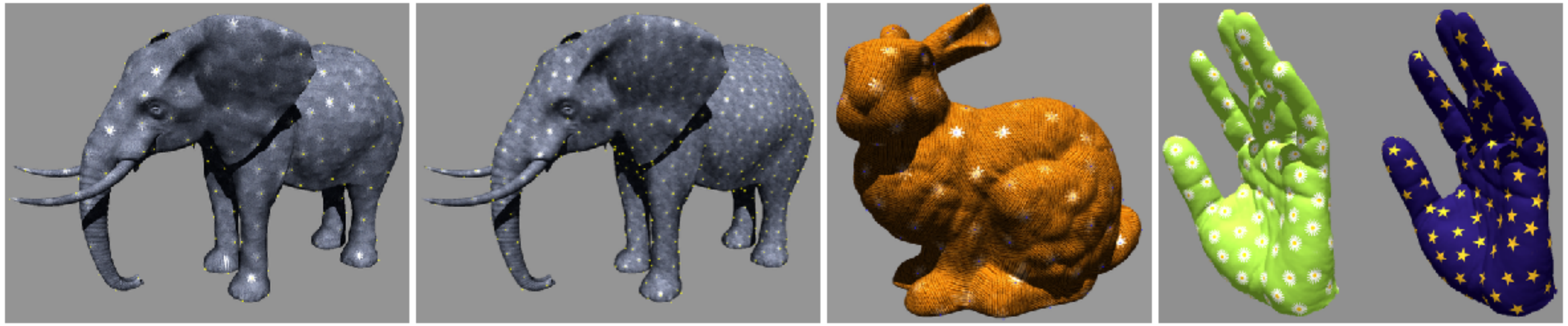
(Point processes in Computer Graphics)







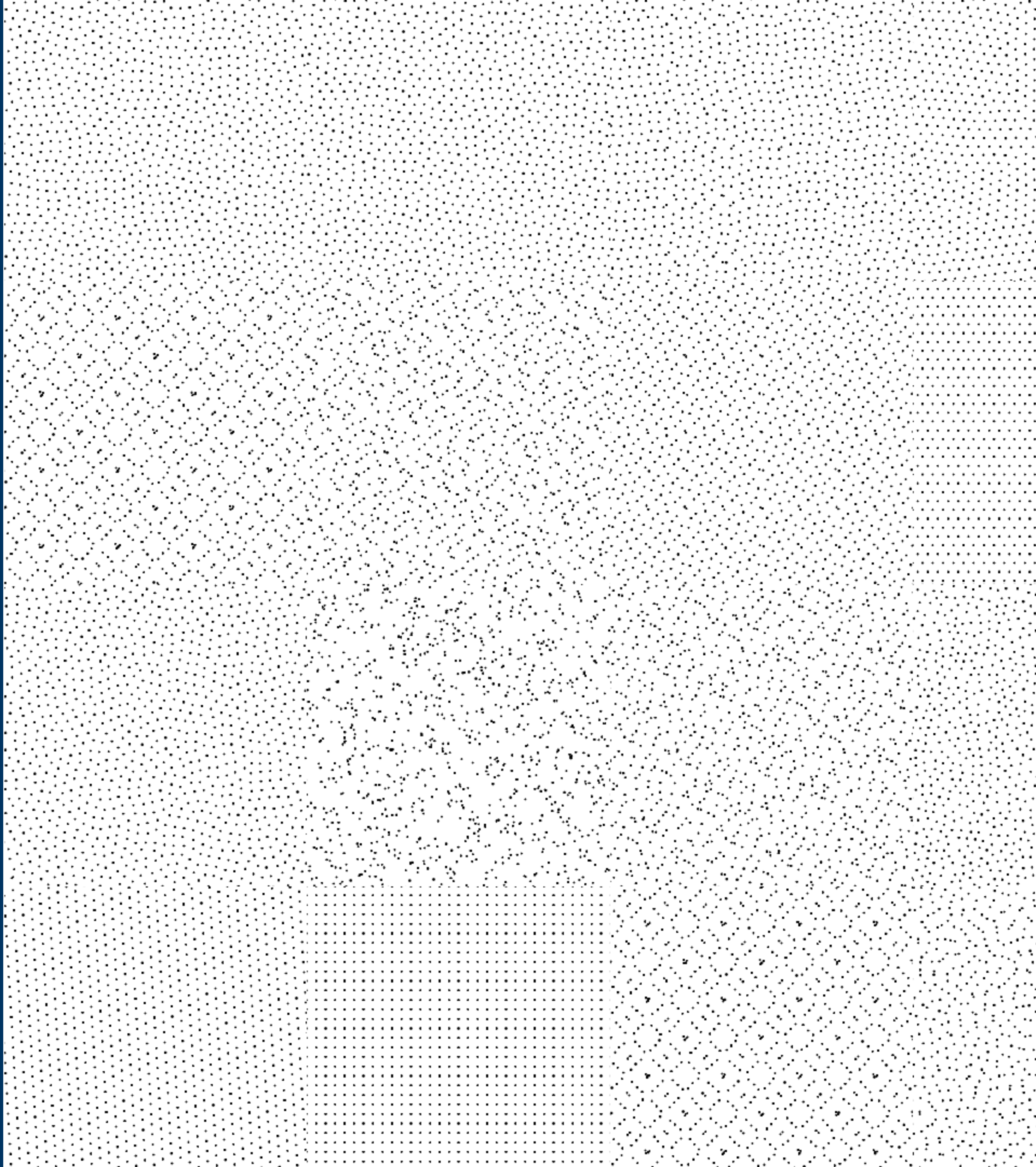




(Advanced techniques)

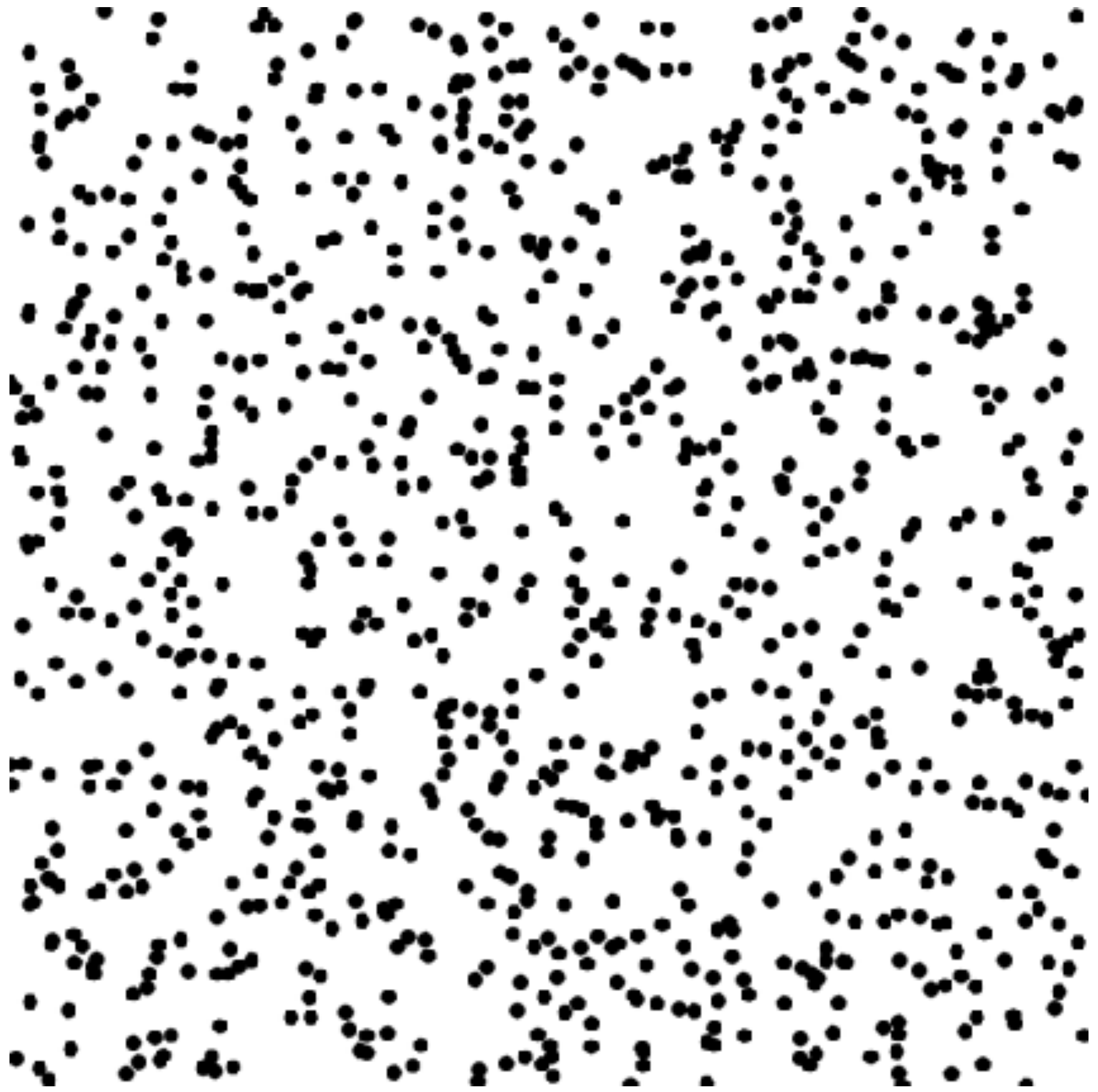
- Non-uniform densities
- Importance / Multiple importance sampling
- Control variates
- Metropolis sampling, Markov chain Monte Carlo…
- Path reuse
- Gradient domain rendering
- Denoising / Reconstruction
- Screenspace error diffusion
- ...

Equidistribution measures



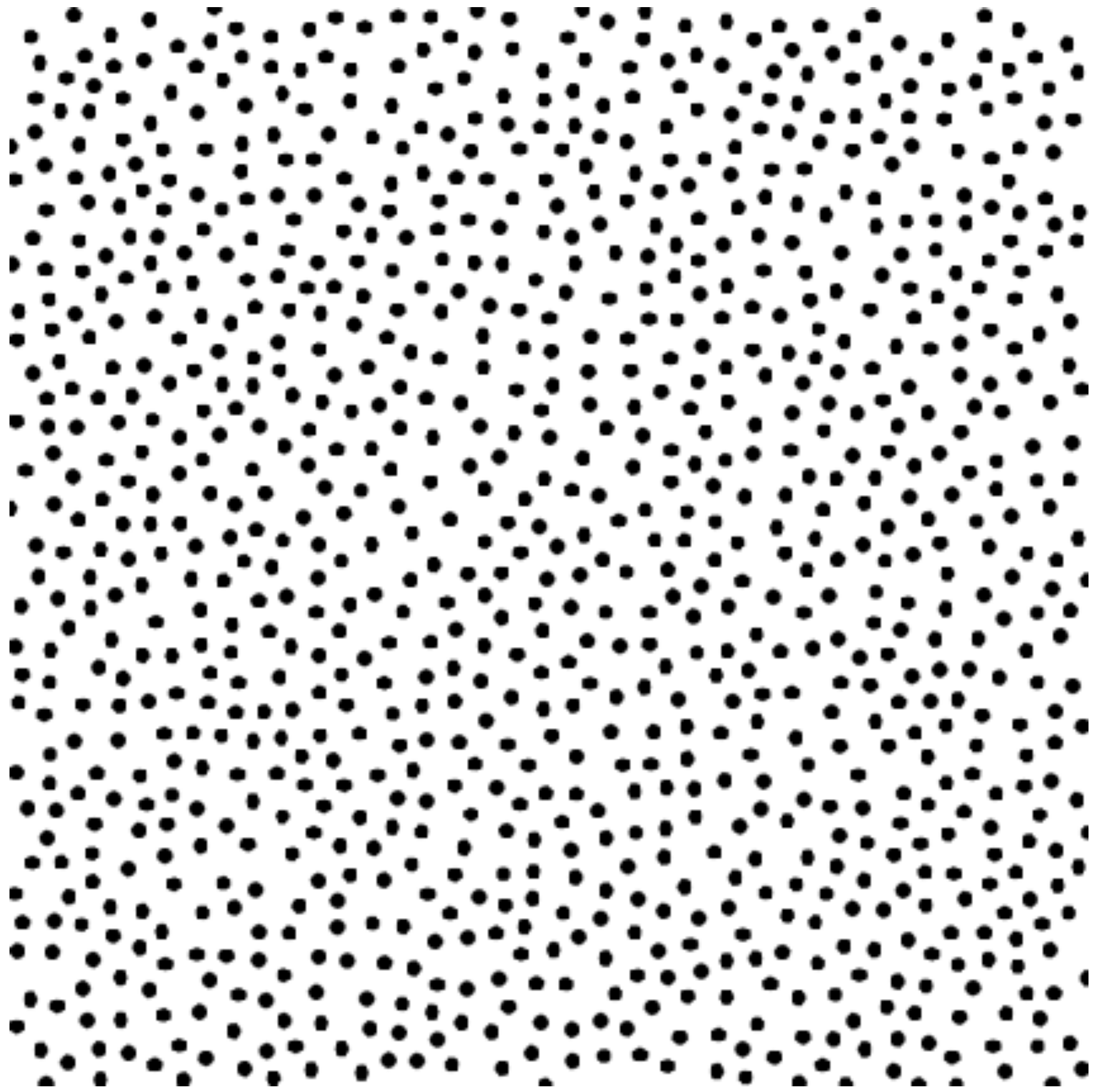
Motivation

Whitenoise



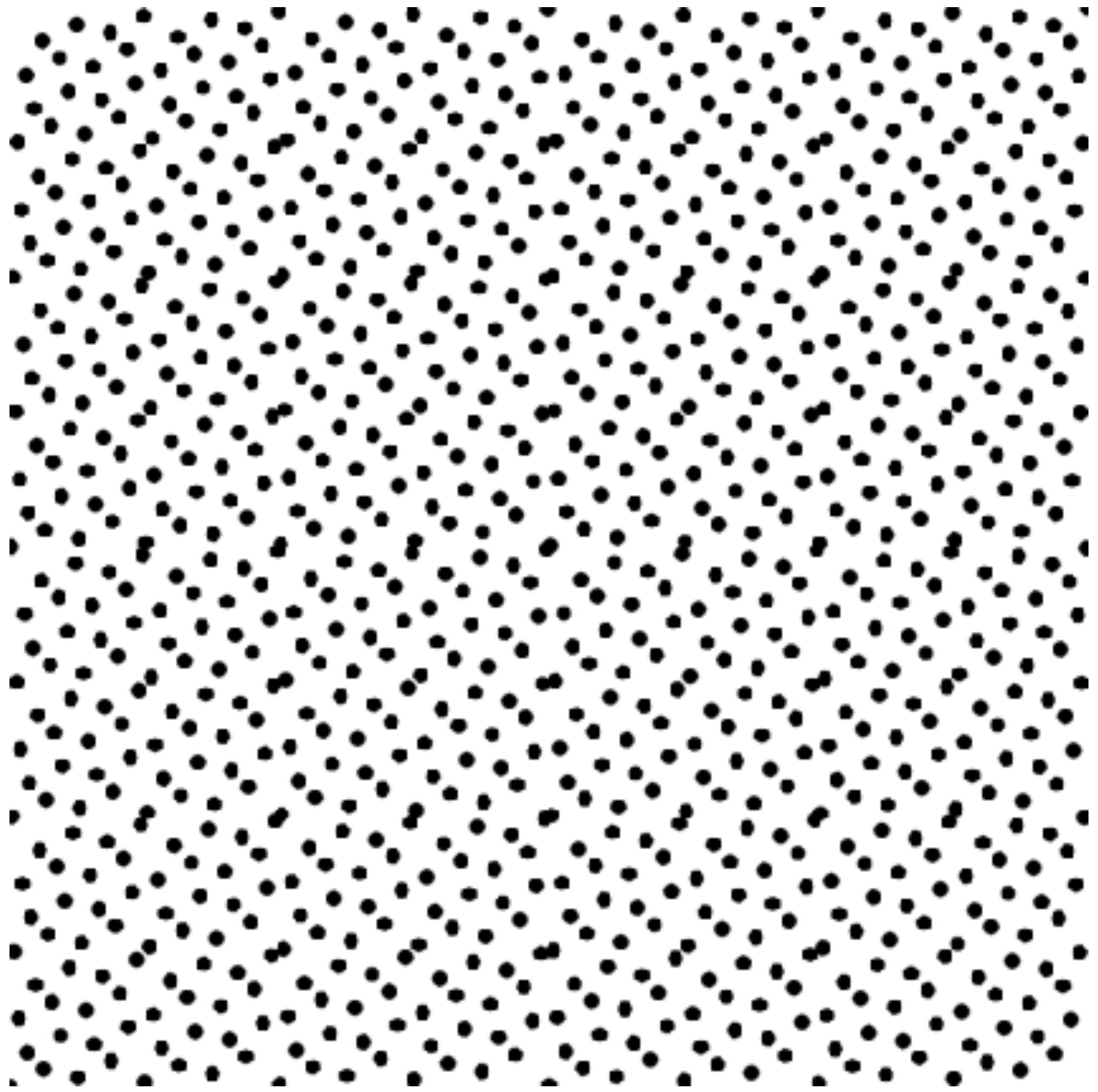
$$Var(I_n) = O\left(\frac{\sigma_f^2}{n}\right)$$

Poisson Disk



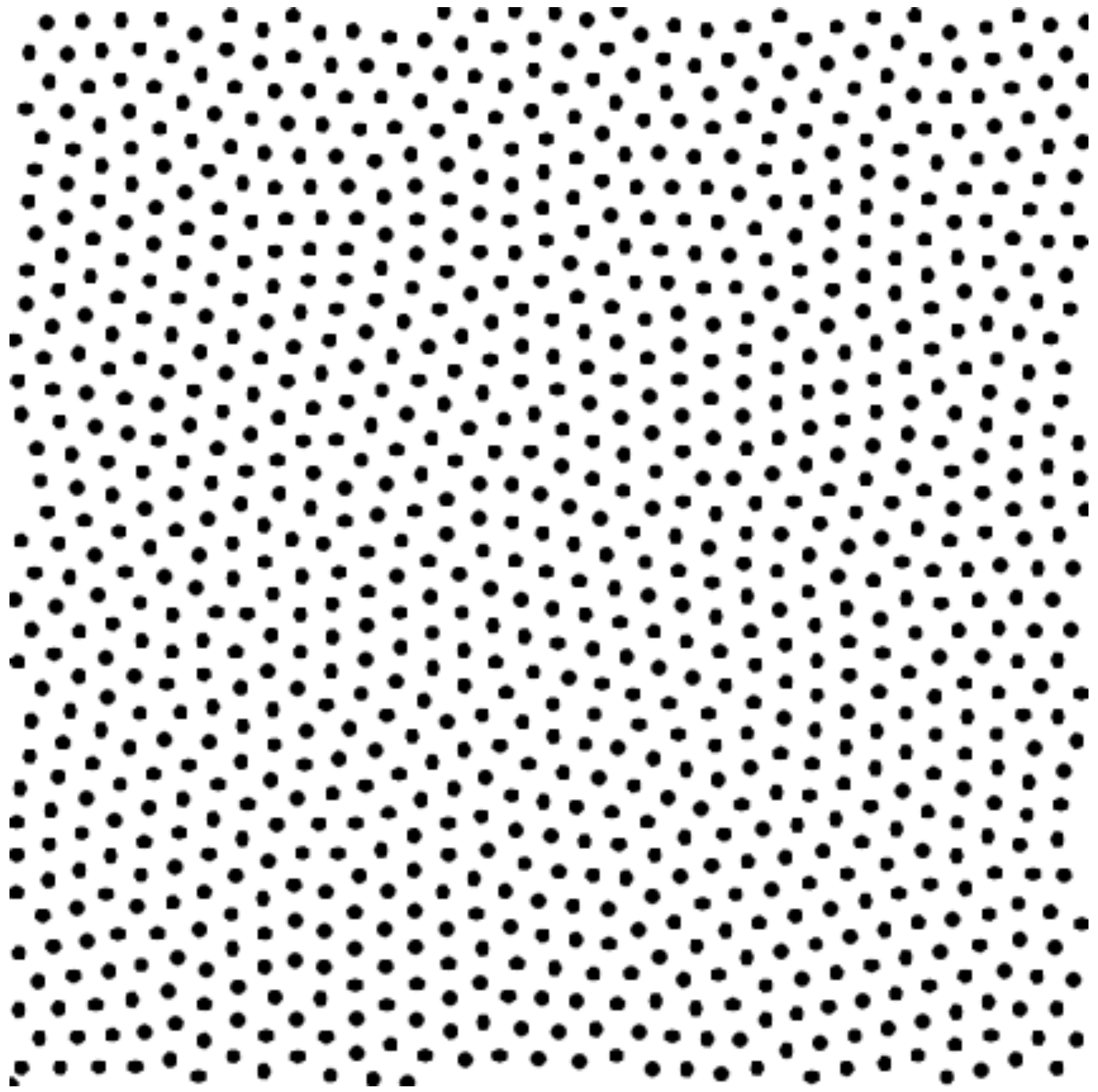
$$Var(I_n) = O\left(\frac{1}{n}\right)$$

Low discrepancy sequences



$$\Delta_n^2 = O\left(\frac{log(n)^{2(s-1)}}{n^2}\right)$$

Blue noise sampling



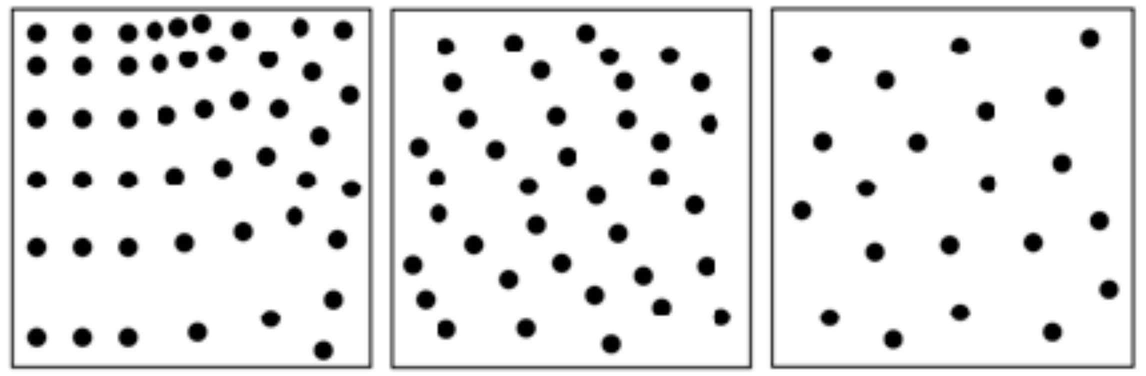
$$Var(I_n) = O\left(\frac{1}{n^{1+1/s}}\right)$$

Uniformity measure + quadrature ⇒ convergence speed

Outline

- Stochastic point process
- Discrepancy
- Kernel based approaches
- Optimal Transport
- Spectral analysis

Spatial measures: Point Process Statistics



- Product densities $\rho^{(k)}$

$$\rho^{(1)}(x)dx = \lambda(x) = \mathbb{P}(x)$$

$$\rho^{(2)}(x,y)dxdy = \mathbb{P}(x,y)$$

$$\mathbb{E}_P(N(\mathscr{B})) = \int_{\mathscr{B}} \rho^{(1)}(x)dx$$

- Stationary point process $\lambda(x) := \lambda$ $\quad \rho^{(2)}(x,y) := \rho(x-y)$ $\quad g(x,y) := \dfrac{\rho(x-y)}{\lambda^2}$

- Isotropic point process $\lambda(x) := \lambda$ $\quad \rho^{(2)}(x,y) := \rho(\|x-y\|)$ $\quad g(r) := \dfrac{\rho(r)}{\lambda^2}$

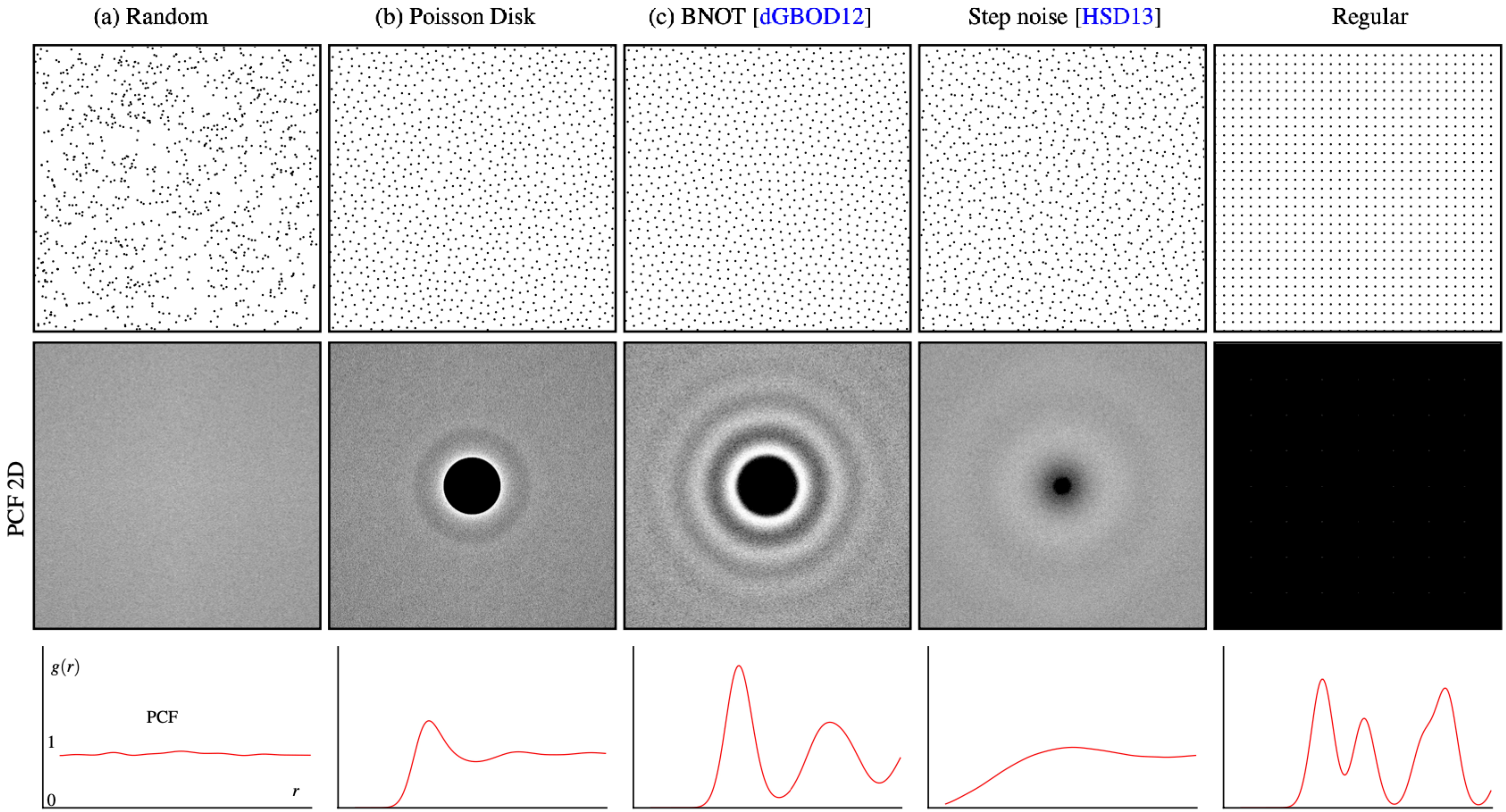
Unbiased kernel based estimator for isotropic PP

$$\hat{g}(r) := \frac{1}{\lambda^2 r^{s-1} |S|^d} \sum_{i \neq j} \kappa(r - \|x_i - x_j\|)$$

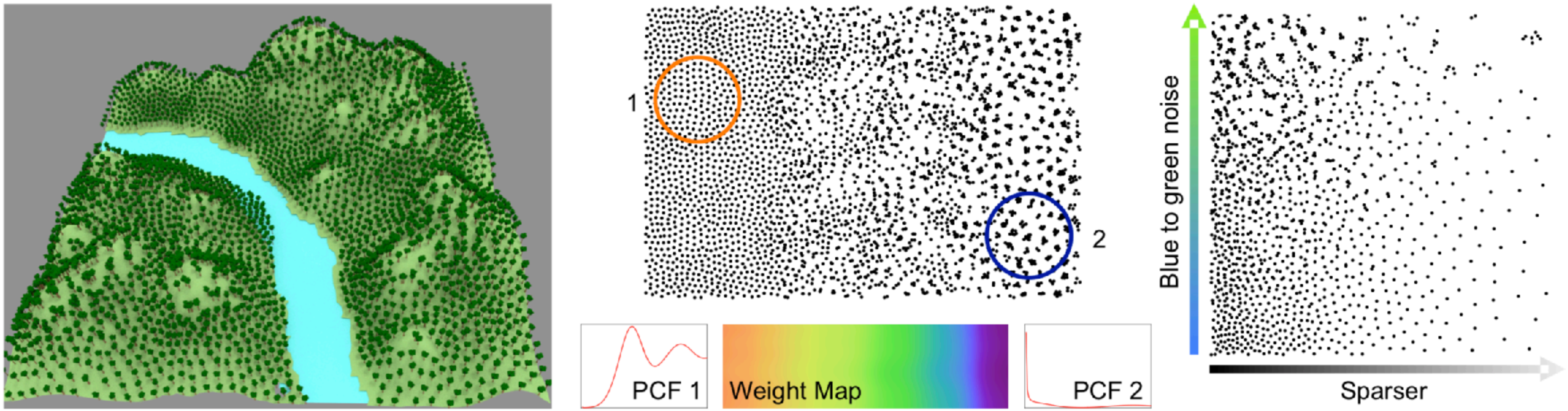
[Öztireli and Gross, 12]

$$Var(I_n) = \frac{1}{\lambda} \int_V f^2(x)dx - \left(\int_V f(x)dx \right)^2 + \int_{\mathbb{R}^s} a_f(r) g(r) dr$$

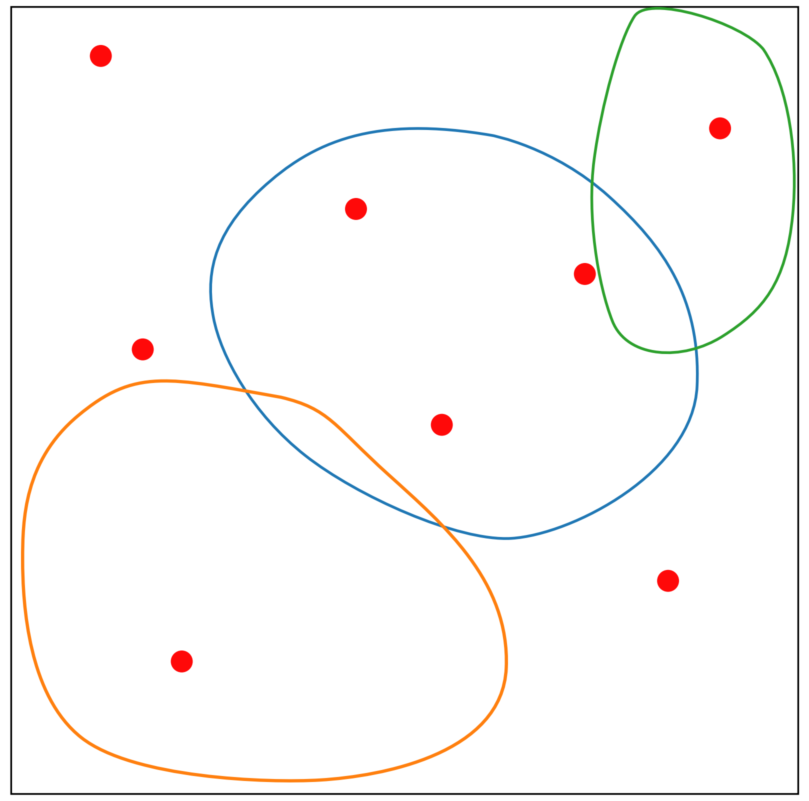
[Öztireli 16]



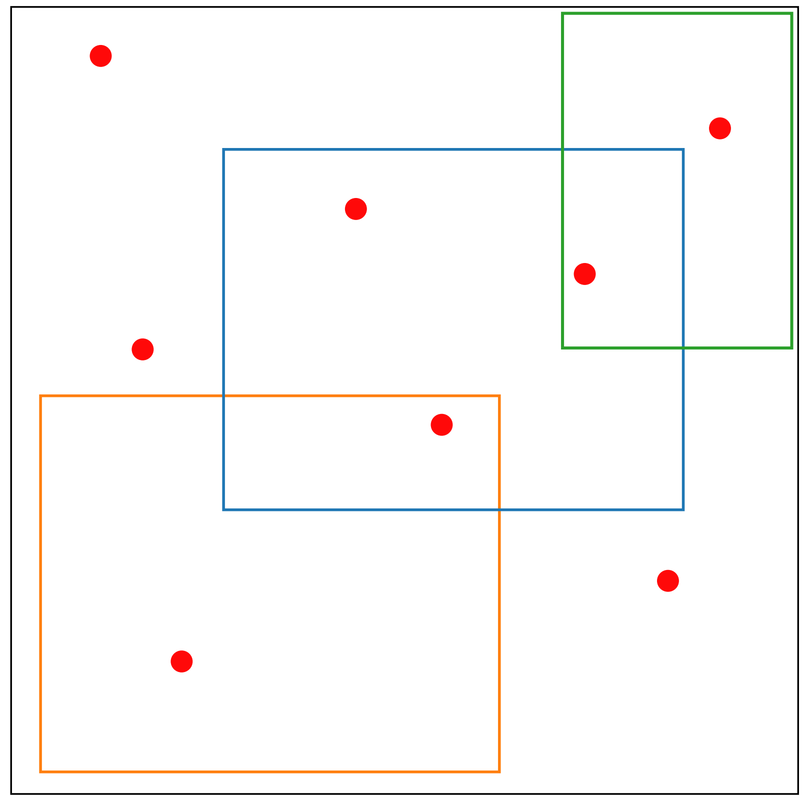
(a) Random
(b) Poisson Disk
(c) BNOT [dGBOD12]
Step noise [HSD13]
Regular
PCF 2D
g(r)
PCF
1
0
r


1
2
PCF 1
Weight Map
PCF 2
Blue to green noise
Sparser

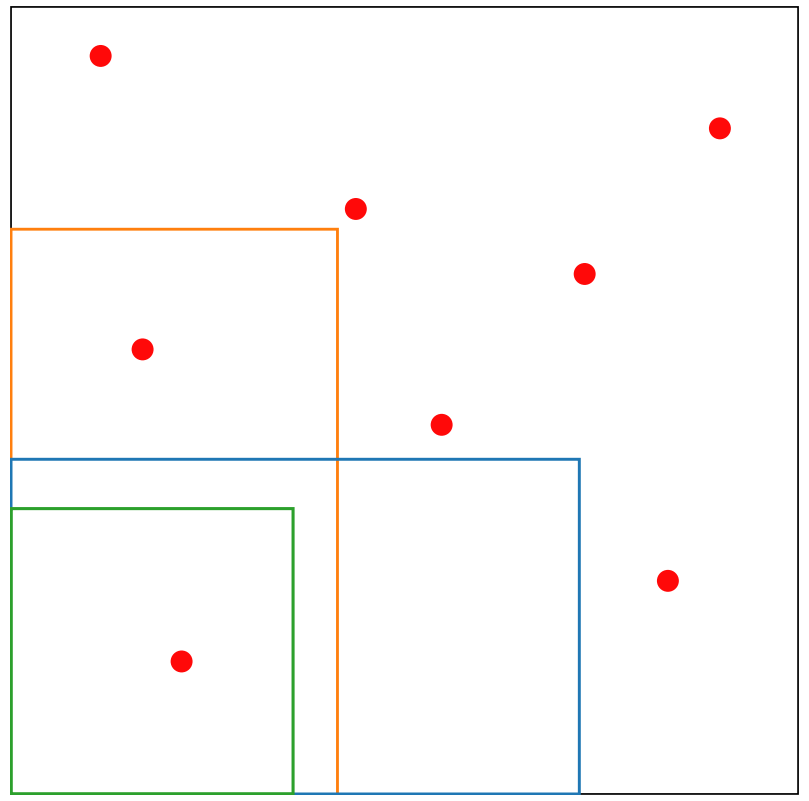
Spatial measures: Discrepancy



Discrepancy



Extreme Discrepancy



Star Discrepancy

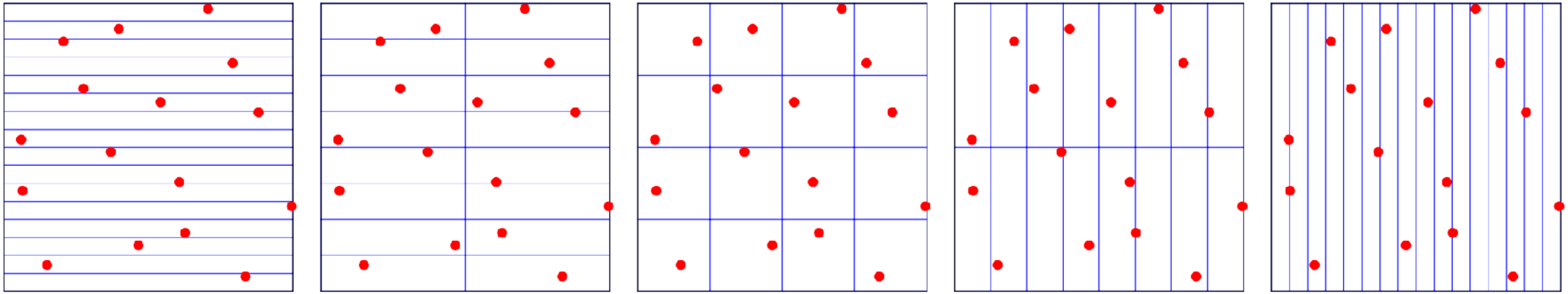
- Extensions to 1d Kolmogorov-Smirnov test
- Equivalent measures
- Many variants, numerical approximations...
- [Koksma-Hlawka] inequality:

$$\Delta_n \leq V(f) \cdot D(X_n)$$

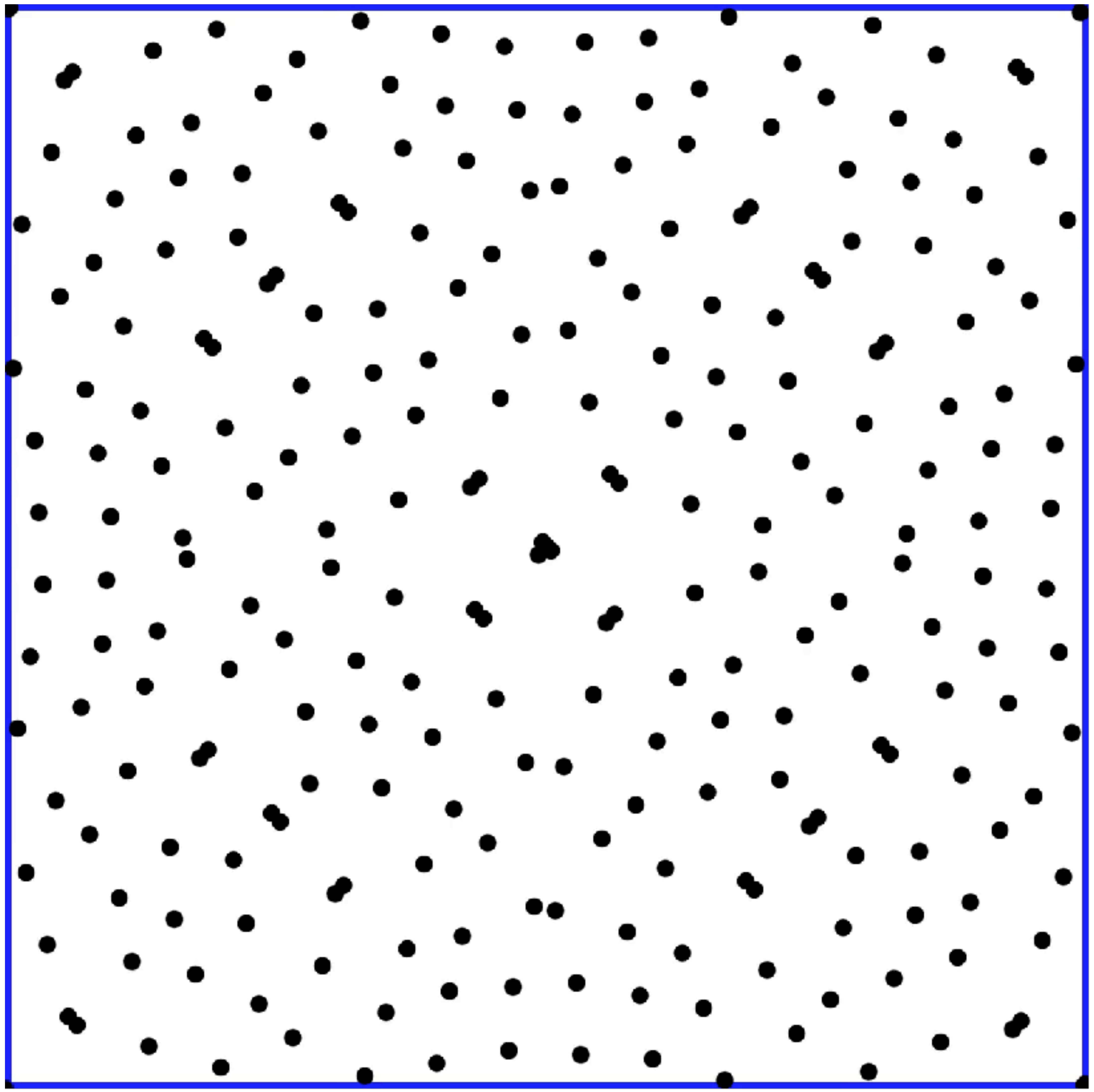
[Koksma 42, Hlawska 61]

$$D(X_n, \Omega) := \left| \lambda_2(\Omega) - \frac{A(\Omega, X_n)}{n} \right| \qquad D(X_n) := \max_{\Omega \subset [0,1)^2} D(X_n, \Omega)$$

$$\exists X_n \text{ s.t. } D(X_n) = O\left(\frac{log^s(n)}{n}\right)$$



(0,m,2)-nets in base 2: for $n = 2^m$ points, all dyadic partitions of size $1/n$ contain exactly 1 sample



Per dimension scrambling preserving the (t,m,s)-net properties [Owen, 97]

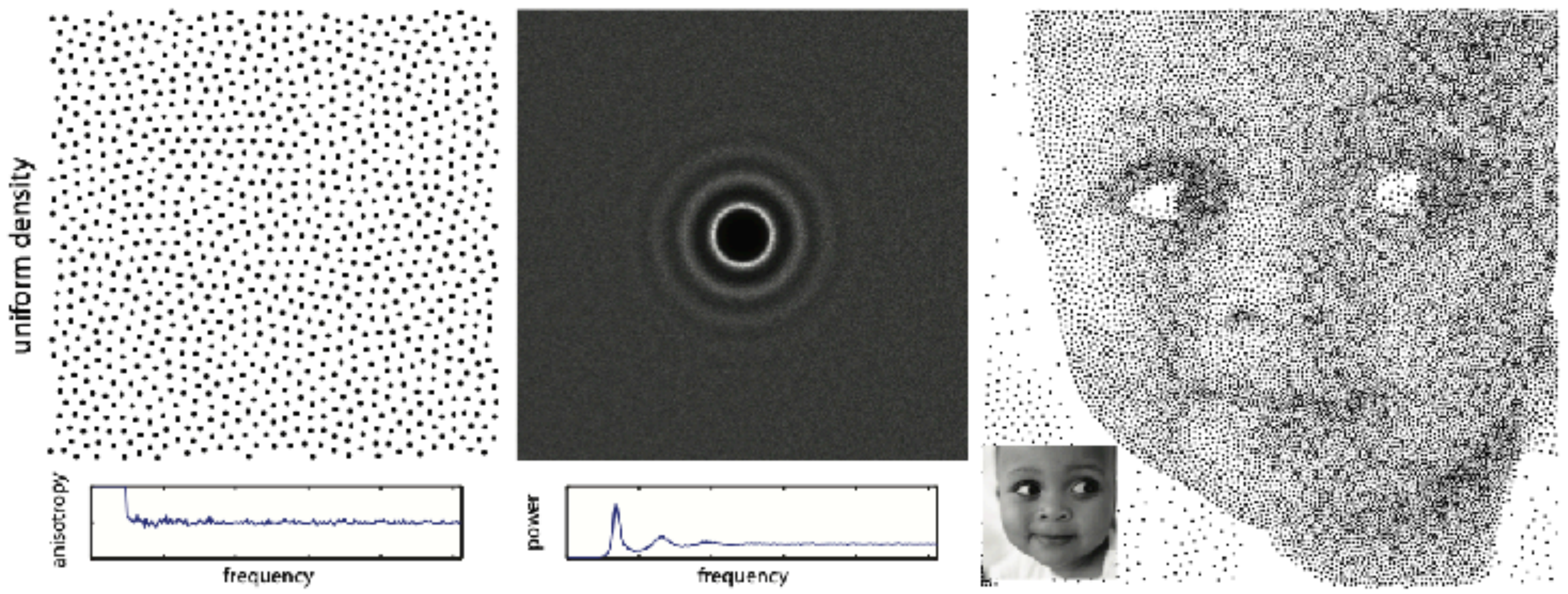
Spatial measures: Kernel based methods

- Sampling as a Gaussian Kernel based reconstruction problem of a constant function

$$A(X_n) = g(x) * \delta(X_n) = \sum_{i=1}^{n} exp\left(-\frac{\|x - x_i\|^2}{2\sigma^2}\right)$$

$$Var(A(X_n)) = \frac{\pi\sigma^2}{n} \sum_{k=1}^{n} \sum_{l=1}^{n} \exp\left(-\frac{\|x_k - x_l\|^2}{4\sigma^2}\right) - \left(2\pi\sigma^2\right)^2$$

- $\Rightarrow$ Find X_n minimizing $Var(A(X_n))$

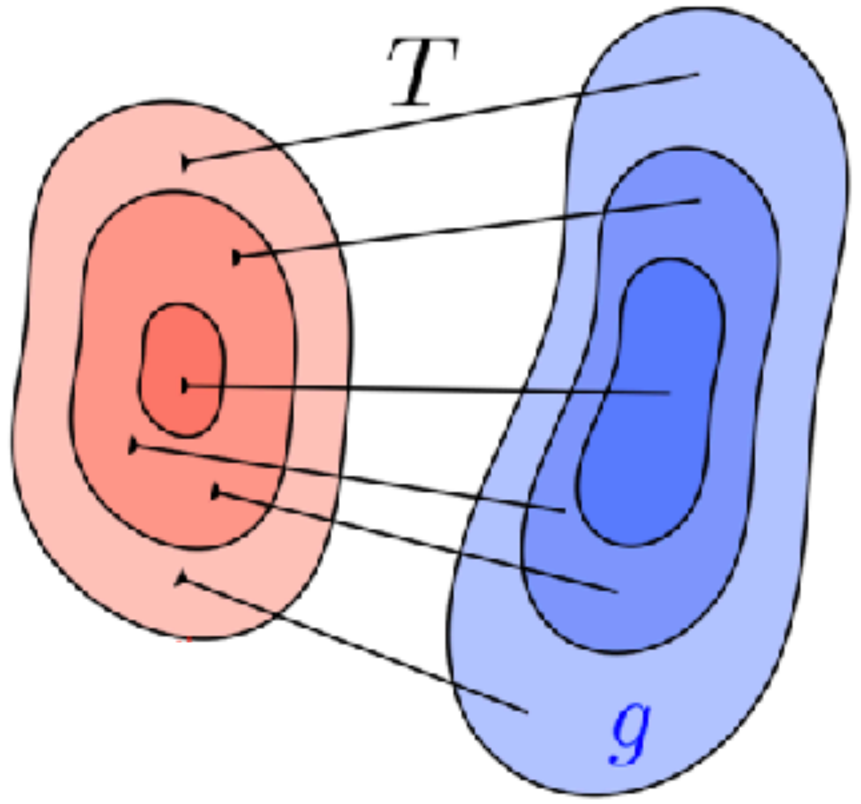


[Fattal 2011, Ahmed and Wonka 2021]

Optimal transport and Wasserstein distance

- Monge-Kantorovich formulation

$$W_p(\alpha, \beta) = \inf \left\{ \int_\Omega \|x - T(x)\|^p d\alpha(x) \right\}^{1/p}, \quad T_*(\alpha) = \beta$$

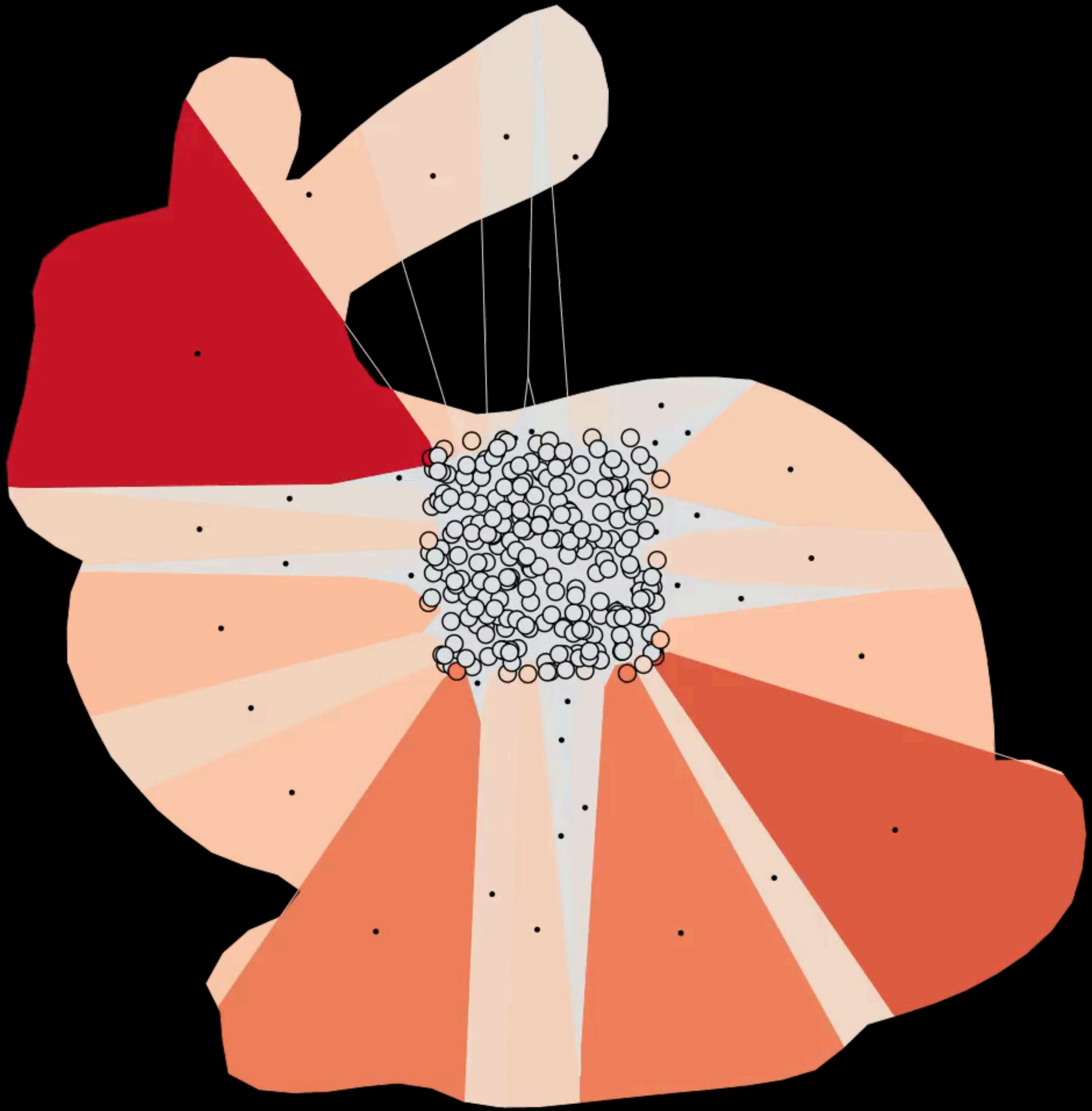


- Semi-discrete optimal transport and Wasserstein distance

$$W_p(X_n, \phi) := \left(\min_\pi \int_\Omega \|x - x^{\pi(x)}\|^p \, \phi(x) dx \right)^{1/p} \quad \pi : \Omega \to \{1 \ldots n\} \quad \int_{\pi^{-1}(j)} \phi(x) dx = \frac{1}{n} \quad \forall j \in \{1 \ldots n\}$$

[Rubinstein-Kantorovich] theorem for a uniform target distribution

$$\Delta_n \leq W_1(X_n, 1_\Omega) \, \mathrm{Lip}(f)$$



Spectral measures: Power spectrum

- Spectral formulation:

$$S(x) = \frac{1}{n}\sum_{k=1}^{n}\delta(x - x_k) \qquad \mathbf{S}_m = \frac{1}{n}\sum_{k=1}^{n}\exp^{-i2\pi(m\cdot x_k)}, \quad m \in \mathbb{Z}^s$$

- Integration + Parseval's theorem

$$I_n = \int_{[0,1)^s} f(x)S(x)dx = \int_{\mathbb{R}^s} \mathcal{F}_f^*(v)\mathcal{F}_S(v)dv = \sum_{m\in\mathbb{Z}^s} \mathbf{f}_m^*\mathbf{S}_m$$

$$\langle \Delta_n \rangle = \mathbf{f}_0^* - \sum_{m\in\mathbb{Z}^s, m\neq 0} \mathbf{f}_m^*\langle \mathbf{S}_m \rangle$$

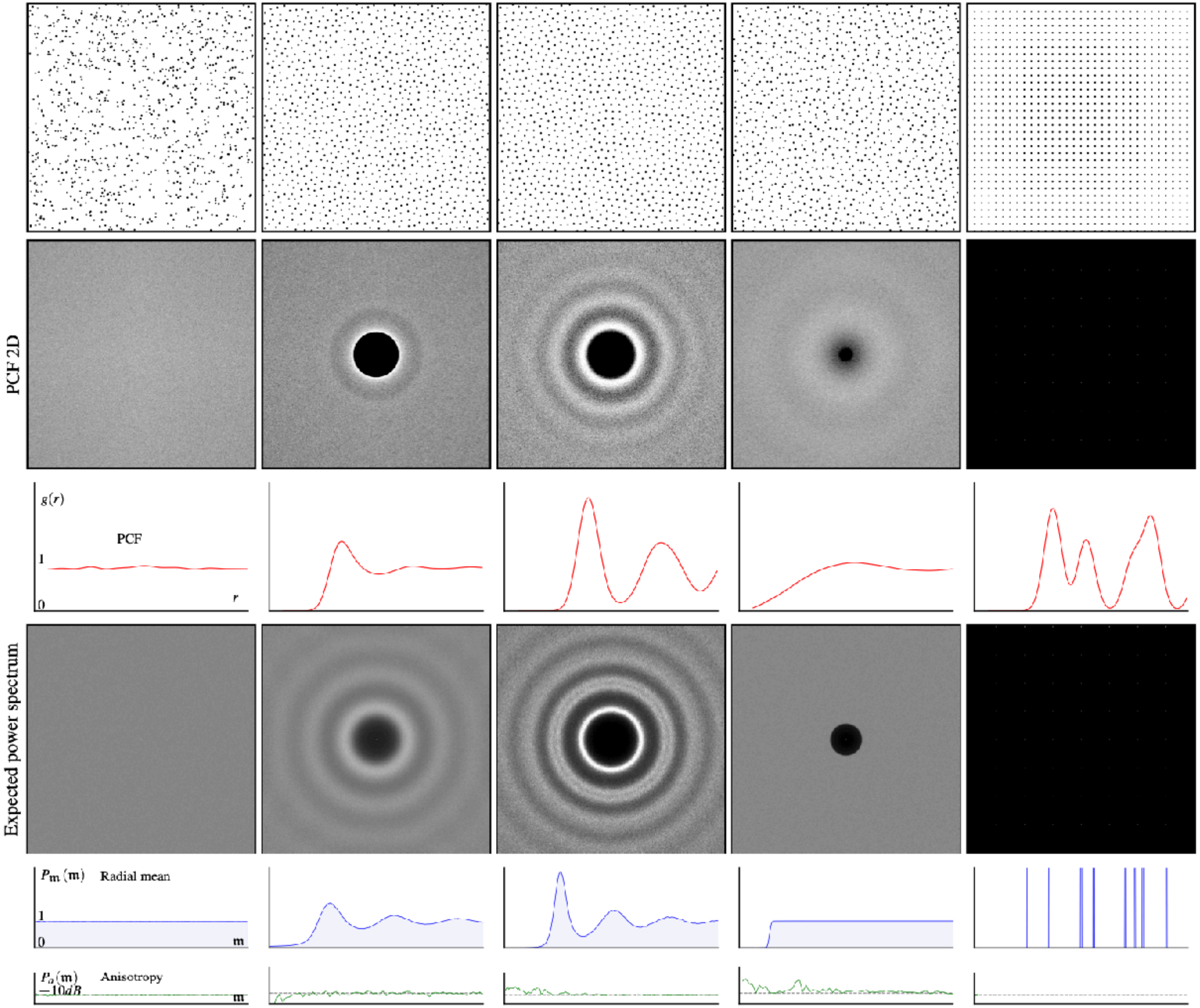
$$Var(I_n) = \mathbf{f}_0^*\mathbf{f}_0 Var(\mathbf{S}_0) + \sum_{m\in\mathbb{Z}^s, m\neq 0} \mathbf{f}_m^*\mathbf{f}_m\langle \mathbf{S}_m^*\mathbf{S}_m \rangle + \sum_{m\in\mathbb{Z}^s}\sum_{l\in\mathbb{Z}^s, l\neq m} \mathbf{f}_m^*\mathbf{f}_l\langle \mathbf{S}_m^*\mathbf{S}_l \rangle$$

For stationary processes:

$$Var(I_n) = \sum_{m\in\mathbb{Z}^s} \mathbf{f}_m^*\mathbf{f}_m\langle \mathbf{S}_m^*\mathbf{S}_m \rangle$$

⇒ Product of power spectra

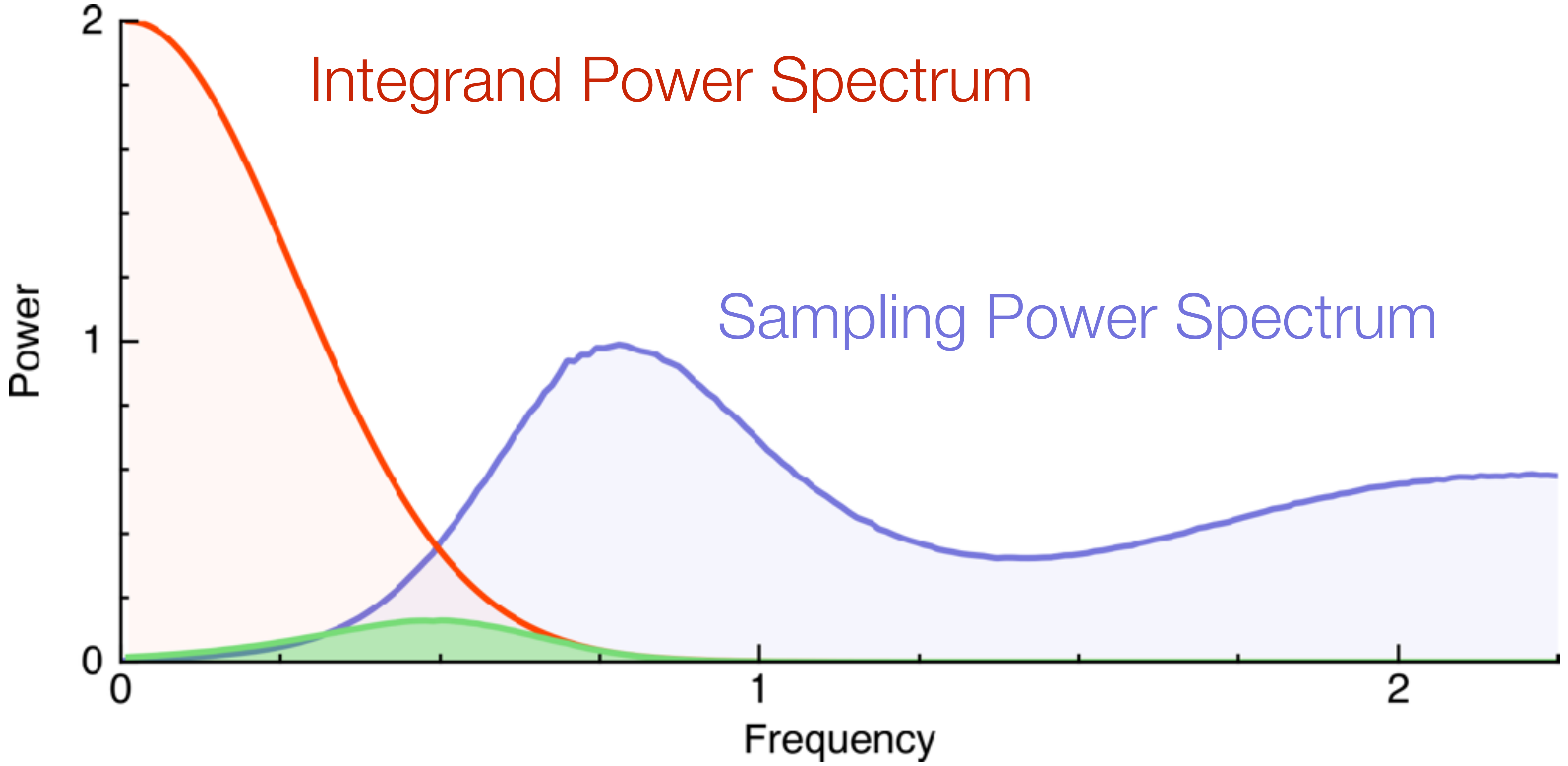
[Durand 11, Subr & Kautz 13, Pilleboue et al 15]

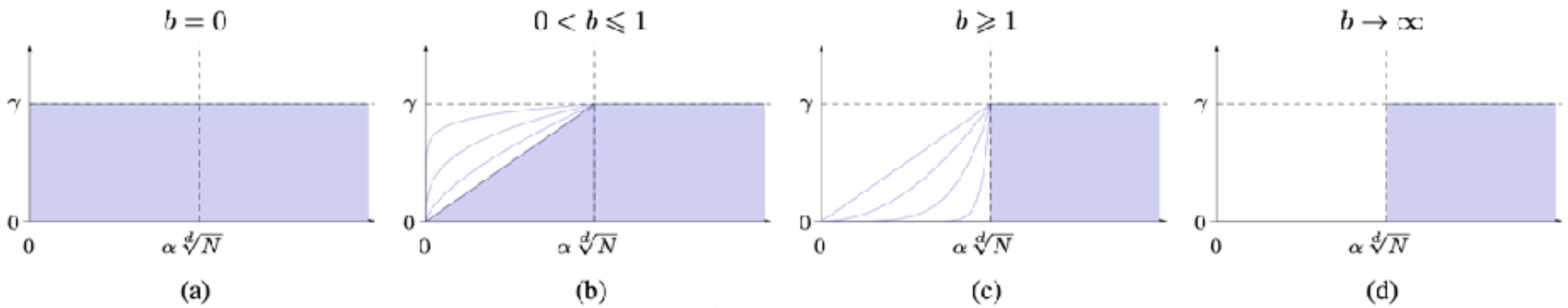


(for isotropic processes: radial mean power spectrum)

[Pilleboue et al 15]
[Singh et al 19]

MC Integration



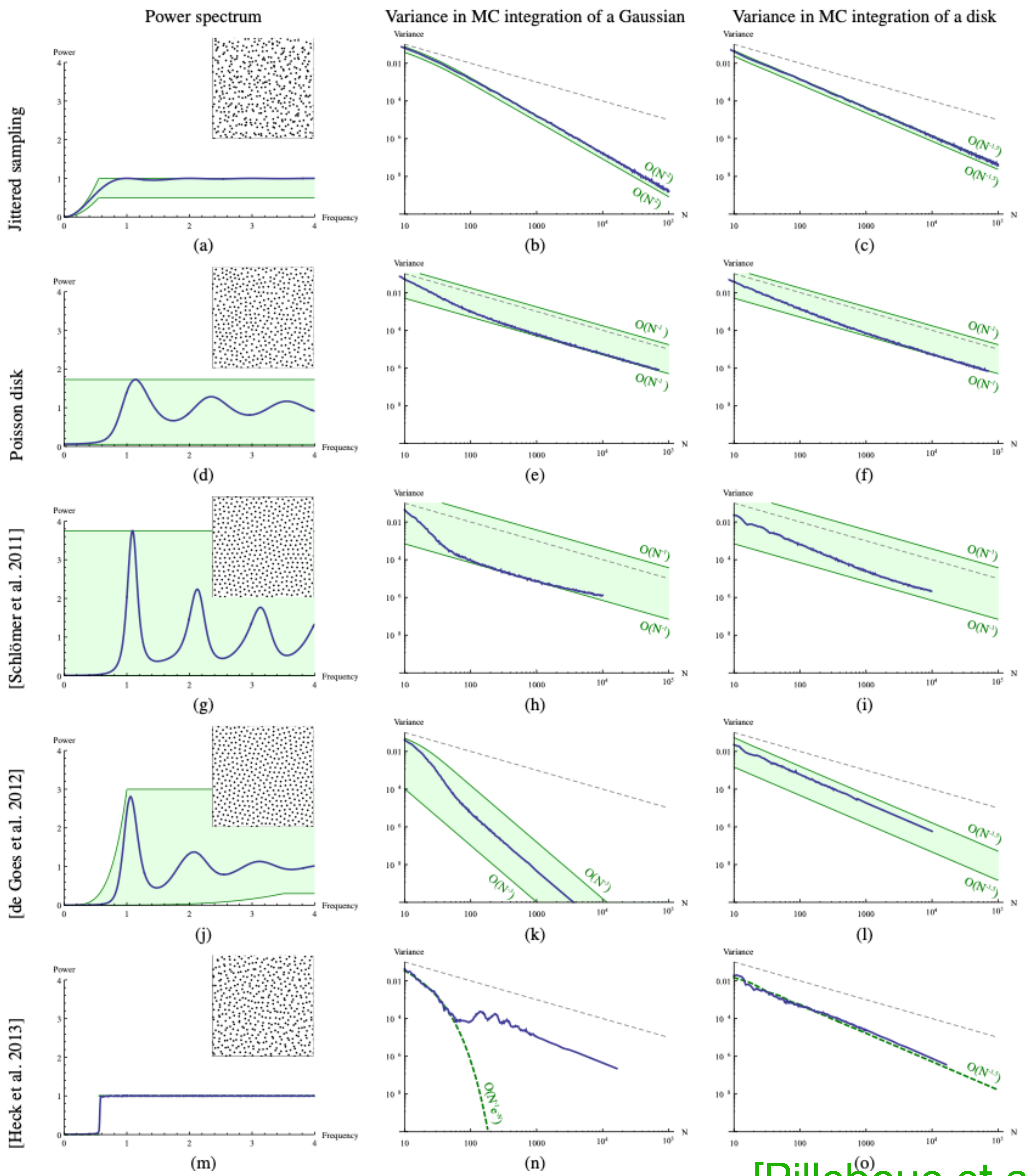


(a) Best-case: $O\left(\frac{1}{N}\right)$ Worst-case: $O\left(\frac{1}{N}\right)$

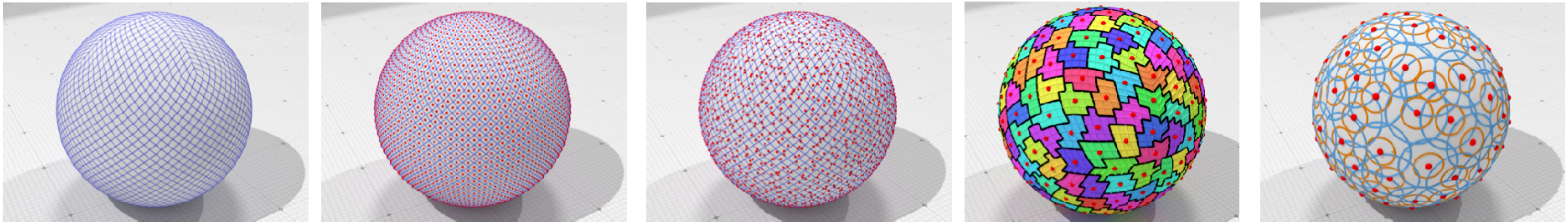
(b) Best-case: $O\left(\frac{1}{N\sqrt[d]{N}^b}\right)$ Worst-case: $O\left(\frac{1}{N\sqrt[d]{N}^b}\right)$

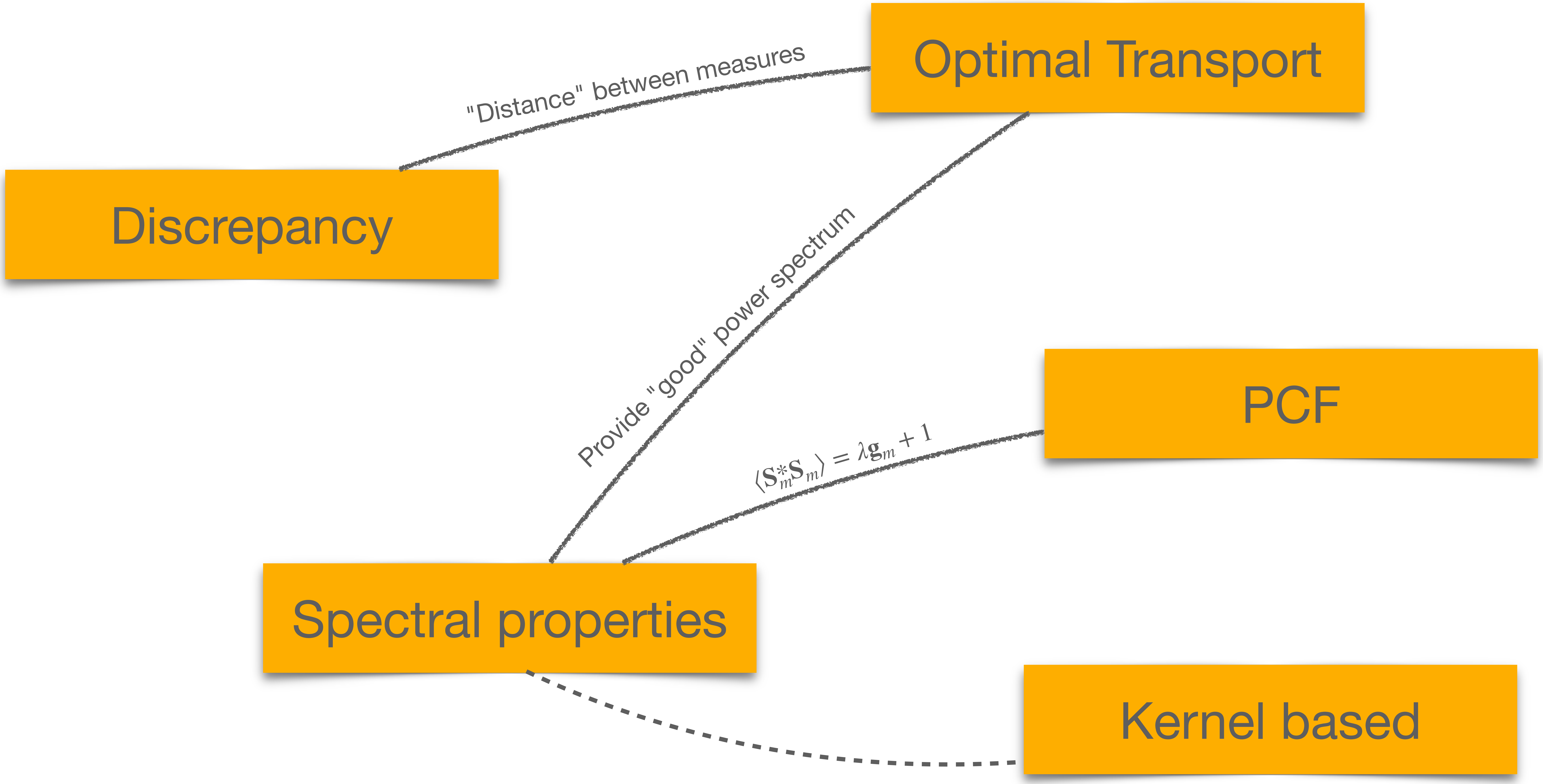
(c) Best-case: $O\left(\frac{1}{N\sqrt[d]{N}^b}\right)$ Worst-case: $O\left(\frac{1}{N\sqrt[d]{N}}\right)$

(d) Best-case: 0 Worst-case: $O\left(\frac{1}{N\sqrt[d]{N}}\right)$



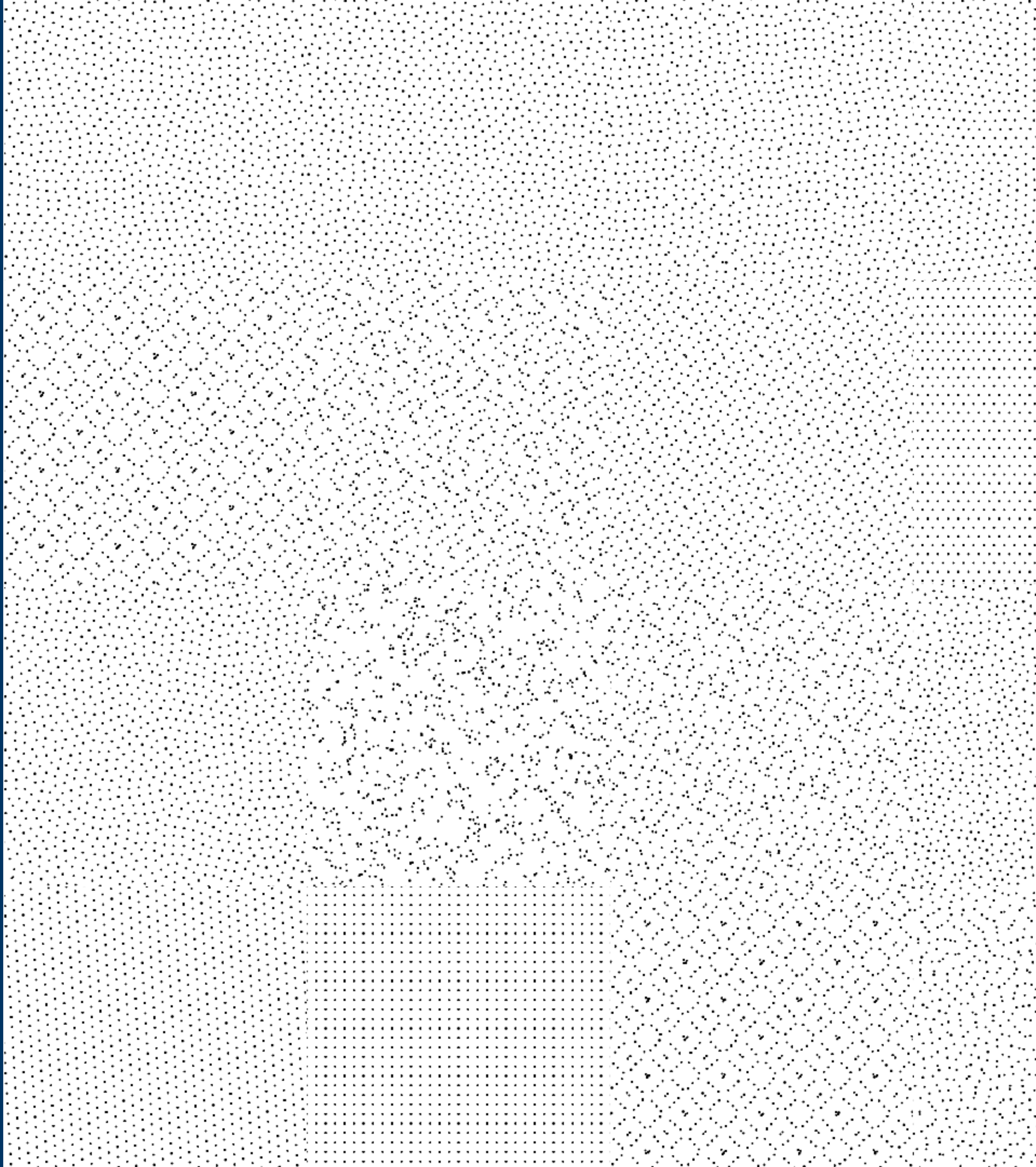
[Pilleboue et al 15]





Optimal Transport
"Distance" between measures
Discrepancy
Provide "good" power spectrum
PCF
$\langle \mathbf{S}_m^* \mathbf{S}_m \rangle = \lambda \mathbf{g}_m + 1$
Spectral properties
Kernel based

Designing samplers



Various routes

- Low Discrepancy Sequences / Pointsets [Sobol, Halton, Faure, Niederreiter, Keller...]
  - fast (matrix/vector product in base 2, or Euclidean division), nD (-ish), $O(log(n)^{s-1}/n)$ integration error, could be randomized through Owen's scrambling

  Discrepancy

- Poisson disk [Bridson 07, Wei 2008, Bowers 10, Guo et al 15, Ebeida 12...]
  - Good sample distributions for low sample counts, focus: fast algorithms (default is $O(n^2)$), genericity (anisotropic metric, on manifolds...)

  PCF | Spectral properties

- Kernel based approaches [Fatall 11, Ahmed 20]
  - Good sample distributions for low sample counts, target: fast algorithm (default is $O(n^2)$)

  Spectral properties

- Blue noise from Optimal Transport [de Goes et al 12, Qin et al 17, Paulin 20]
  - Targeting blue-noise like power spectrum, versatile tool (to sample non-uniform pdf) but numerically expensive in higher dimensions

  Optimal Transport

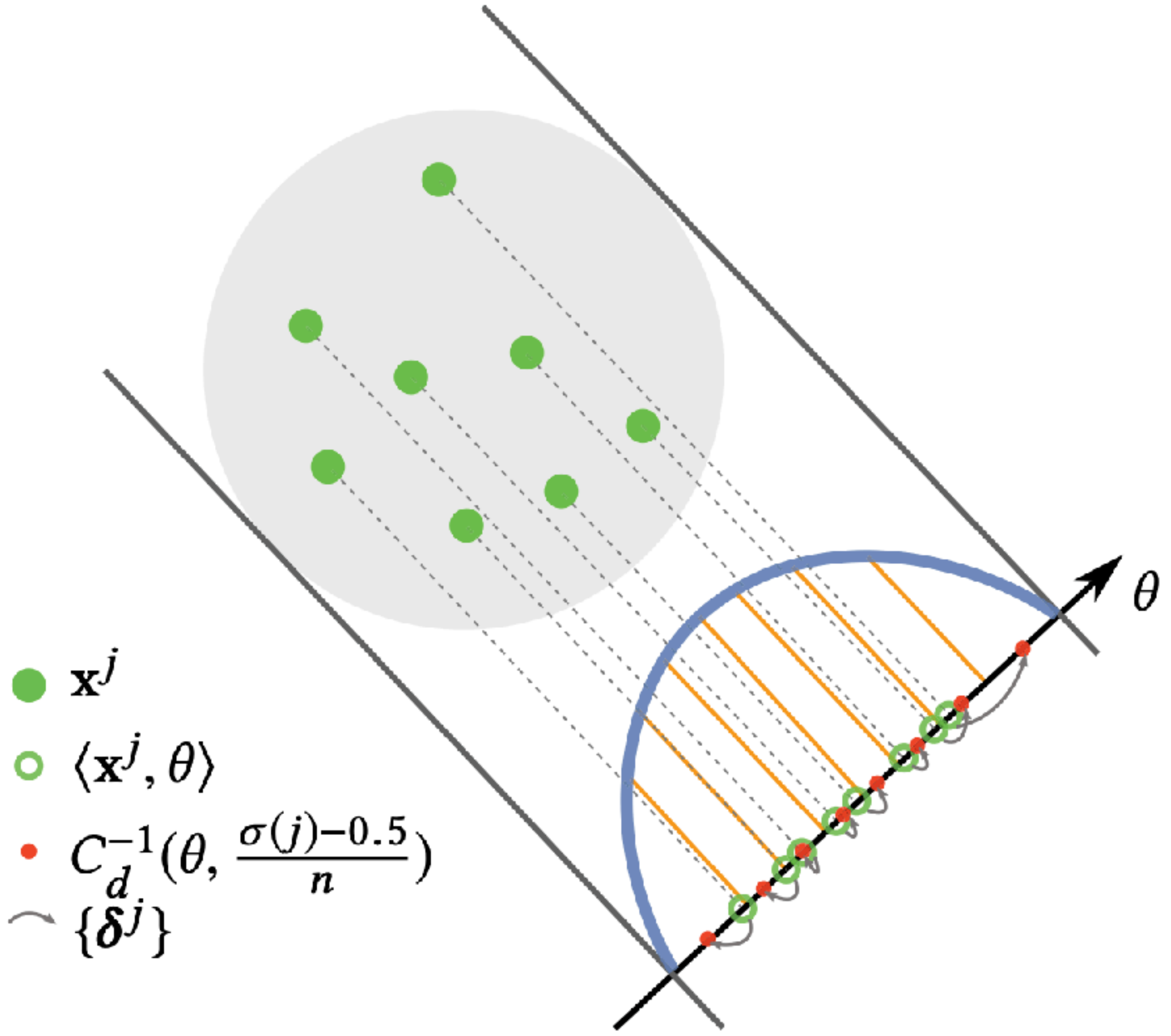
- Mixed solutions
  - Optimized LDS with Blue-Noise targets [Ahmed 16, Perrier 18, Ahmed 20]

    Discrepancy | PCF | Spectral properties
  - Tiled based Blue-noise for fast (adaptive) sampling [Ostromoukhov, , Wachtel 14]

    Aperiodic tilings + Spectral properties
  - Projective sampling (with Blue-noise or LDS targets) [Reinert 15, Paulin 21]

    Discrepancy | Spectral properties

Sliced Optimal Transport Sampling

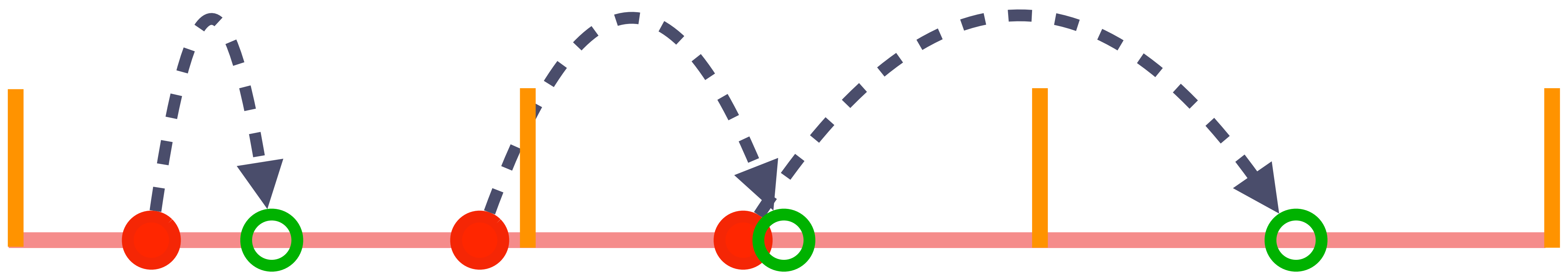


Loïs Paulin, Nicolas Bonneel, David Coeurjolly, Jean-Claude Iehl, Antoine Webanck, Mathieu Desbrun, Victor Ostromoukhov
ACM Transactions on Graphics (Proceedings of SIGGRAPH) July 2020

Motivation

$$\Delta_n \leq W_1(X_n, 1_\Omega)\ \mathrm{Lip}(f)$$

- Find X_n minimizing $W_1(X_n, 1_\Omega)$
- 1D case: sorting + advection



- Key ingredient in high dimension: (semi-discrete) Sliced Optimal Transport

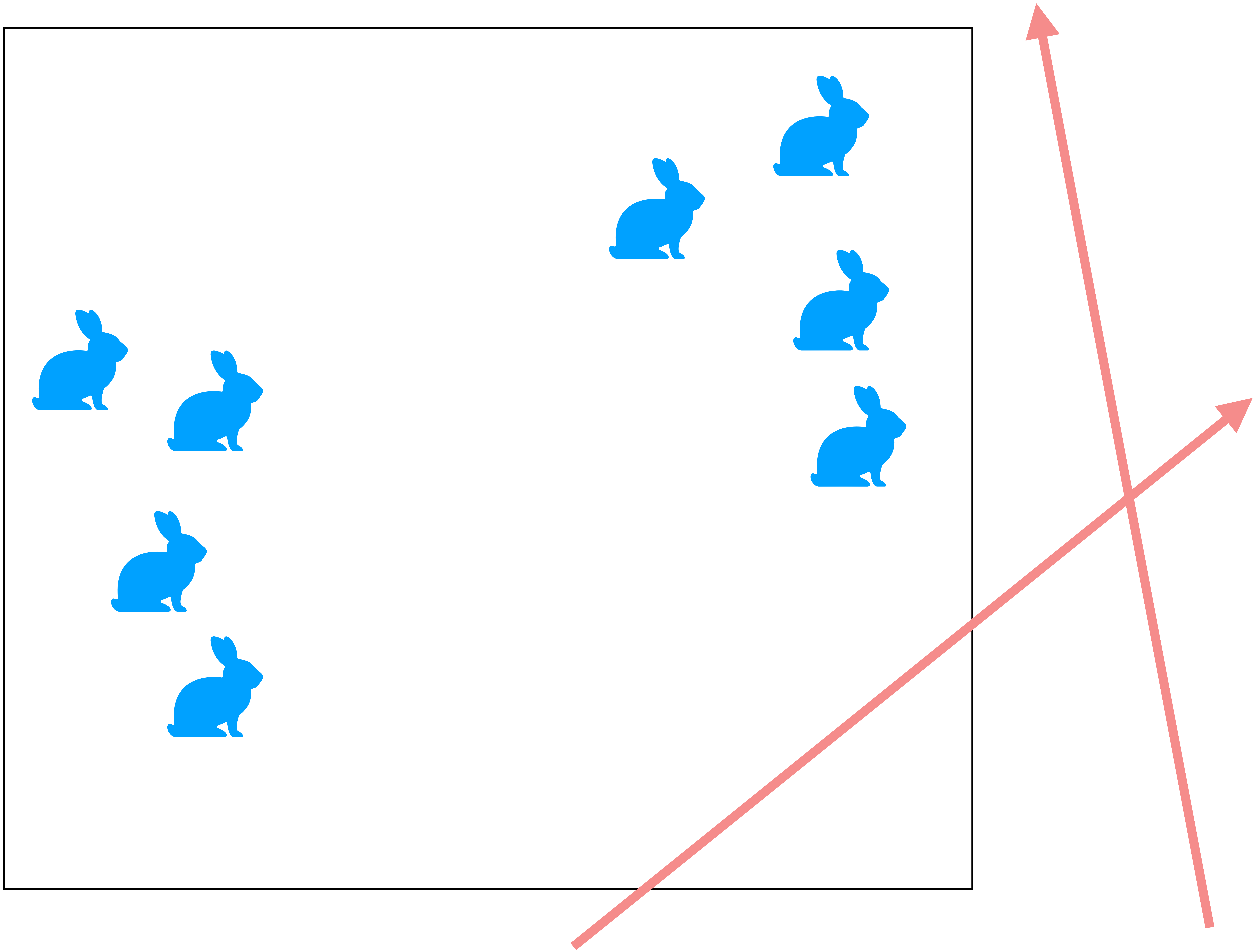
$$W_{\mathrm{SOT}}(X, \phi) = \int_{\mathbb{S}^{d-1}} \left(\min_{\pi} \int_{\mathbb{R}} \|x - x_\theta^{\pi(x)}\| \boxed{R_\theta \phi(x)} \phi_\theta) \right) \mathrm{d}\theta$$

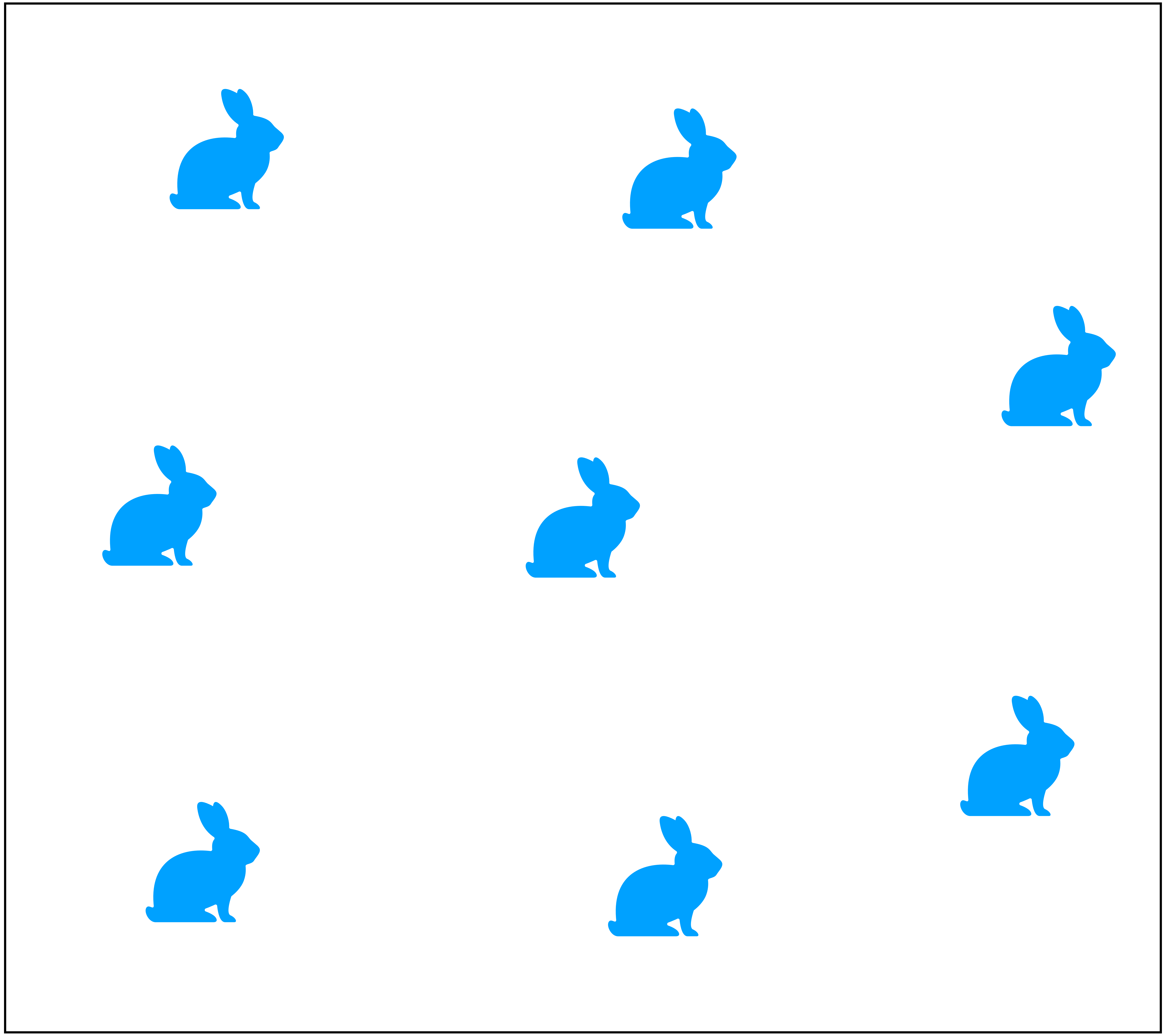
MC estimate

Radon Transform of ϕ

$$\Delta_n \leq C_s W_{SOT}(X_n, 1_\Omega)^{\frac{1}{s+1}}\ \mathrm{Lip}(f)$$

[Pitié et al 2006, Rabin et al 2011, Bonnotte 2013, Bonneel et al 2015, Paulin et al 2020]

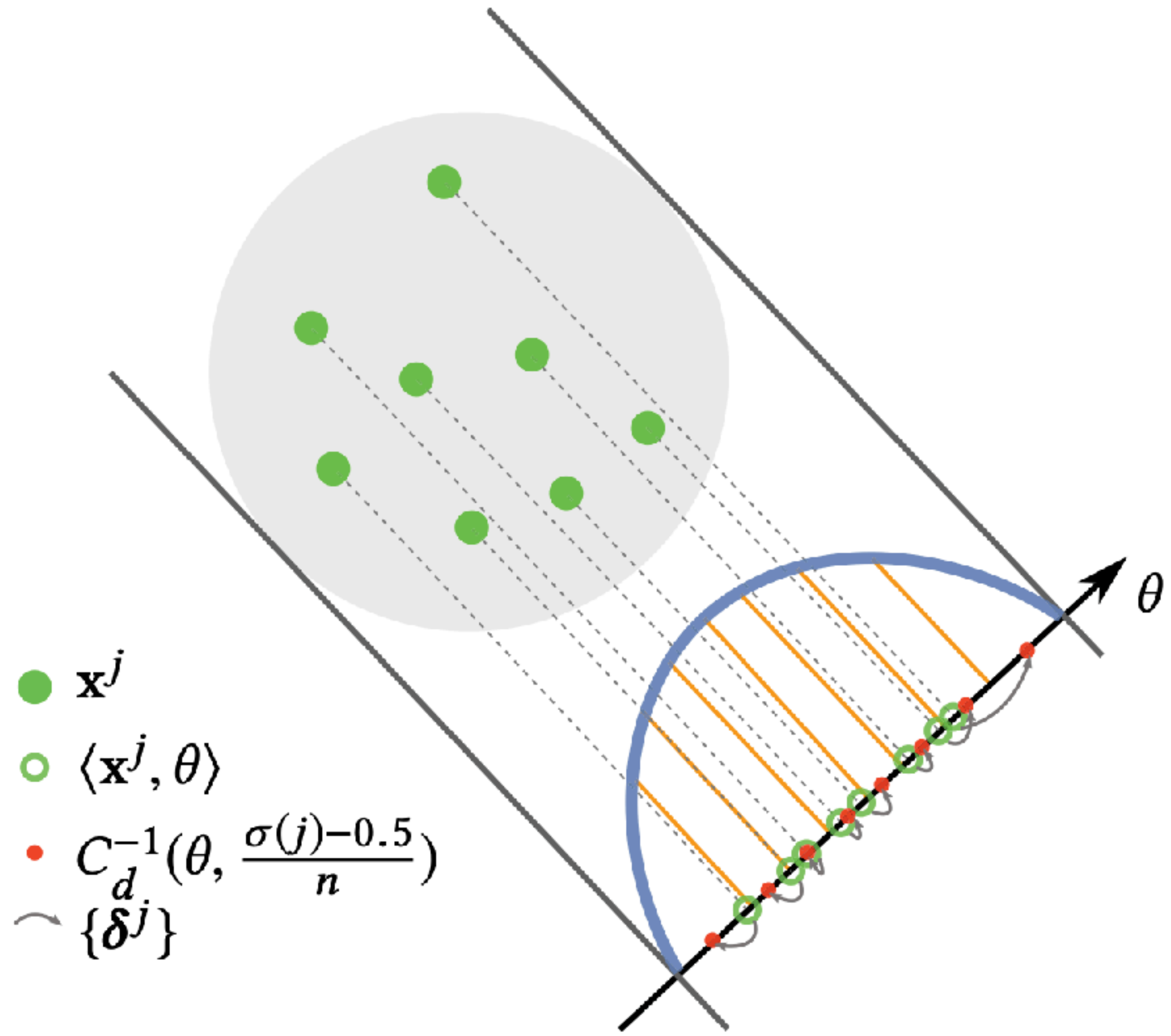


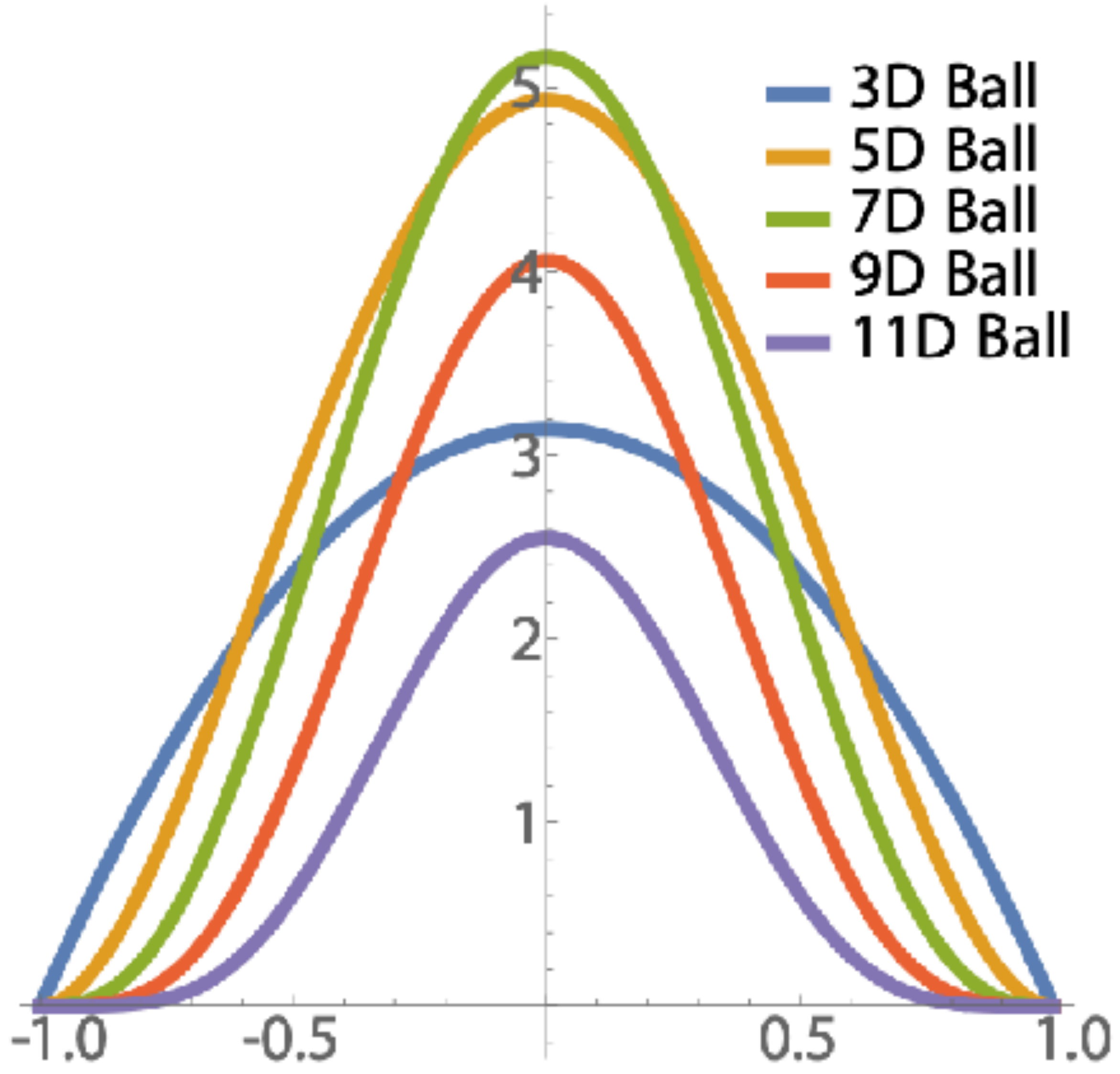


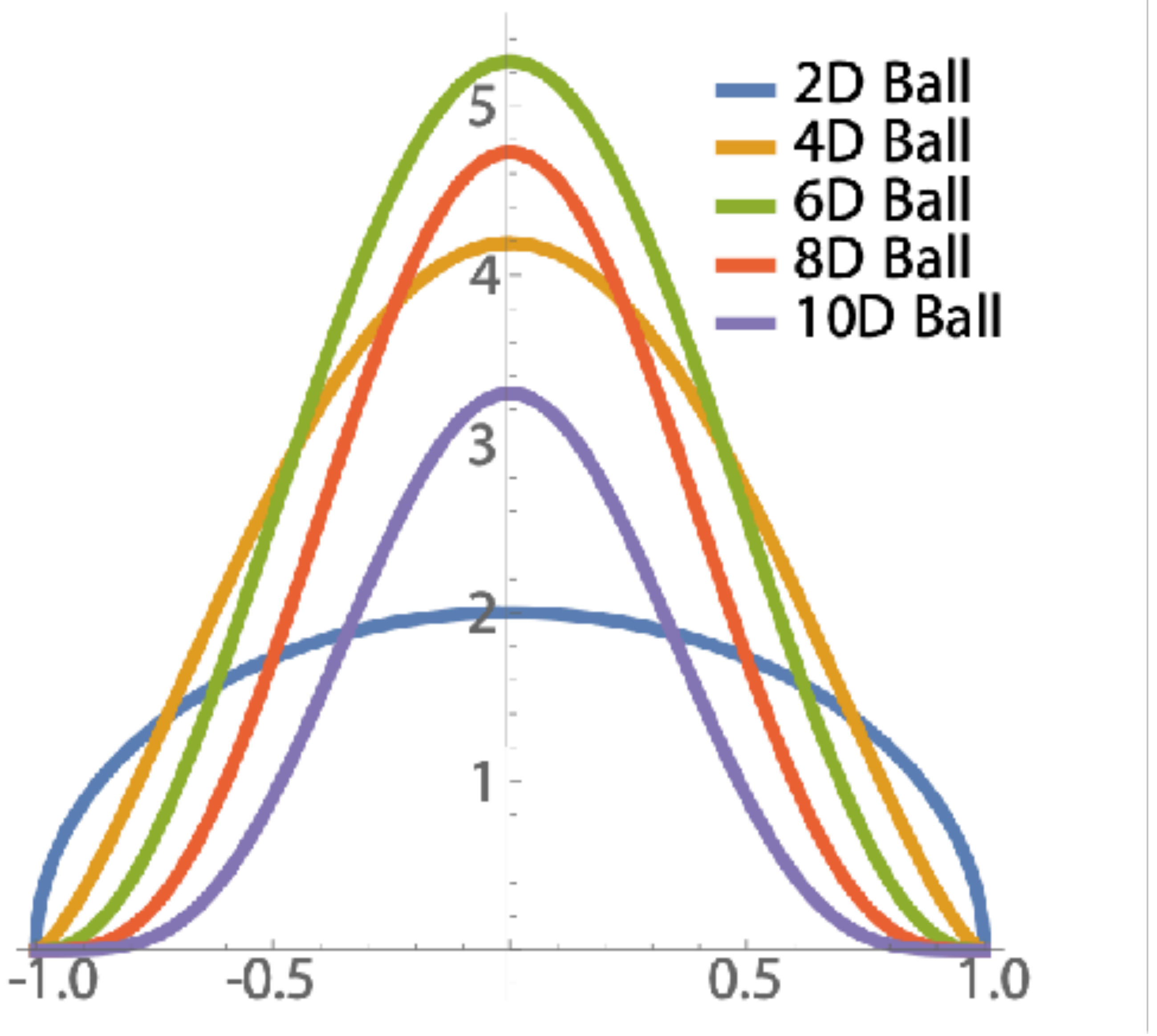
1d sliced advection ≈ stochastic gradient descent on W_{SOT}

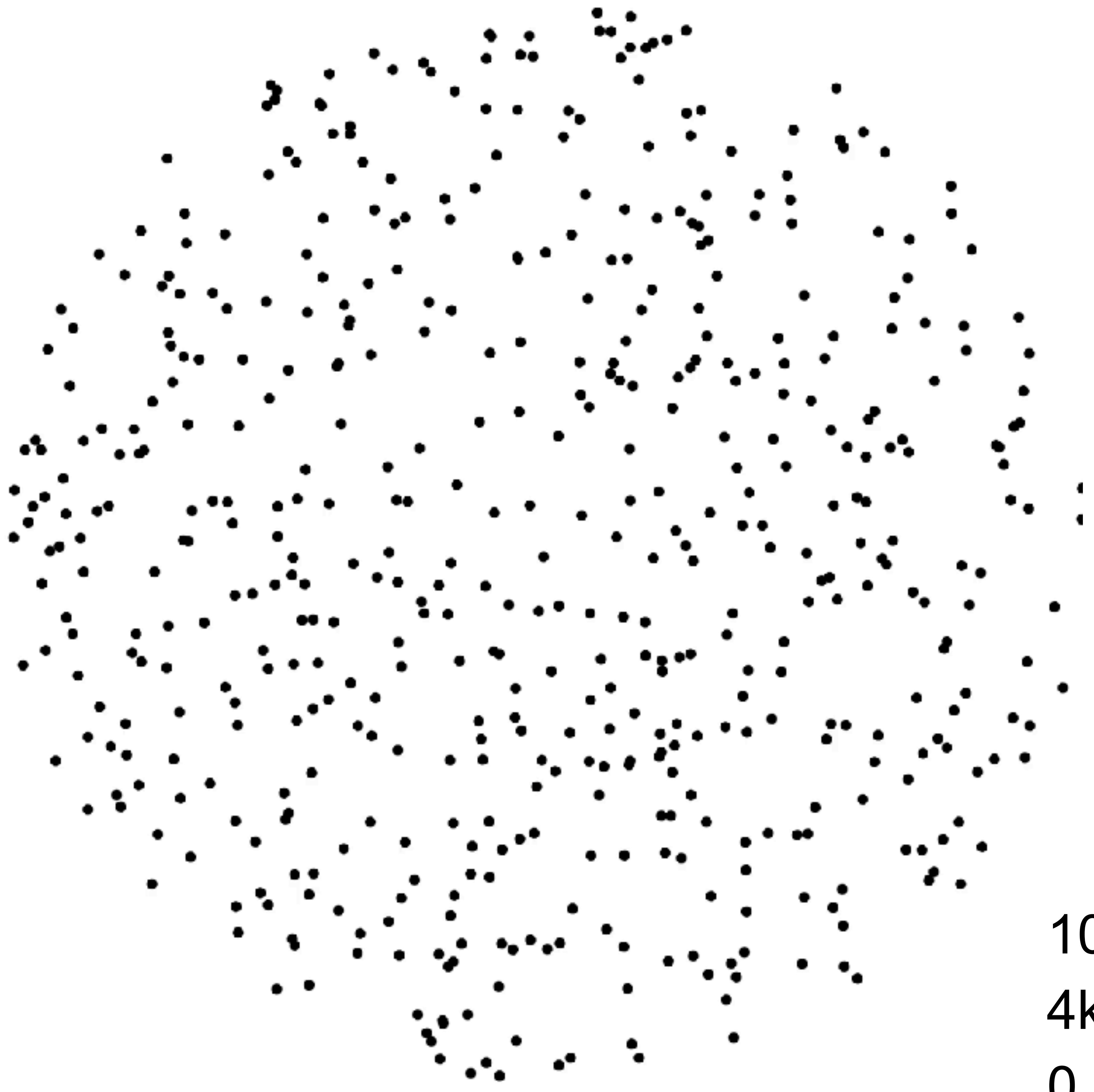
s-balls domains

- Closed formulas for $R_\theta 1_\Omega(x)$ for d-balls





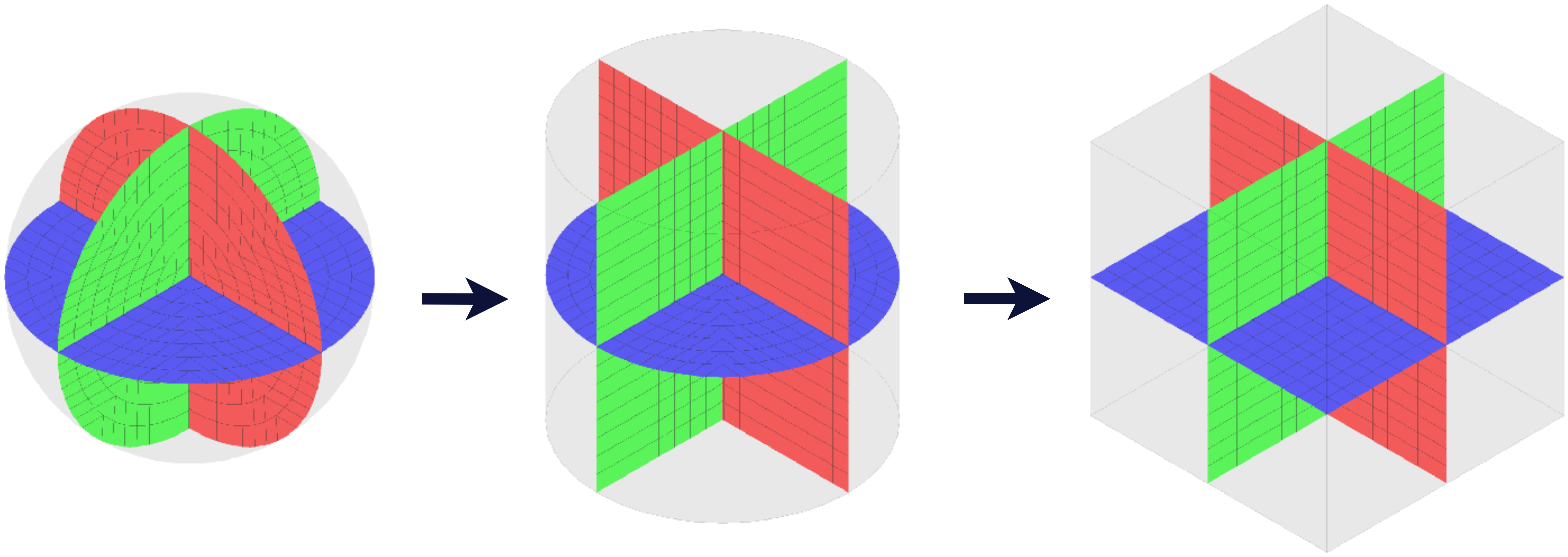




1024 samples in 2D
4k slices per sec
0.25 sec for convergence (1024 slices)

$[0,1)^s$ domains

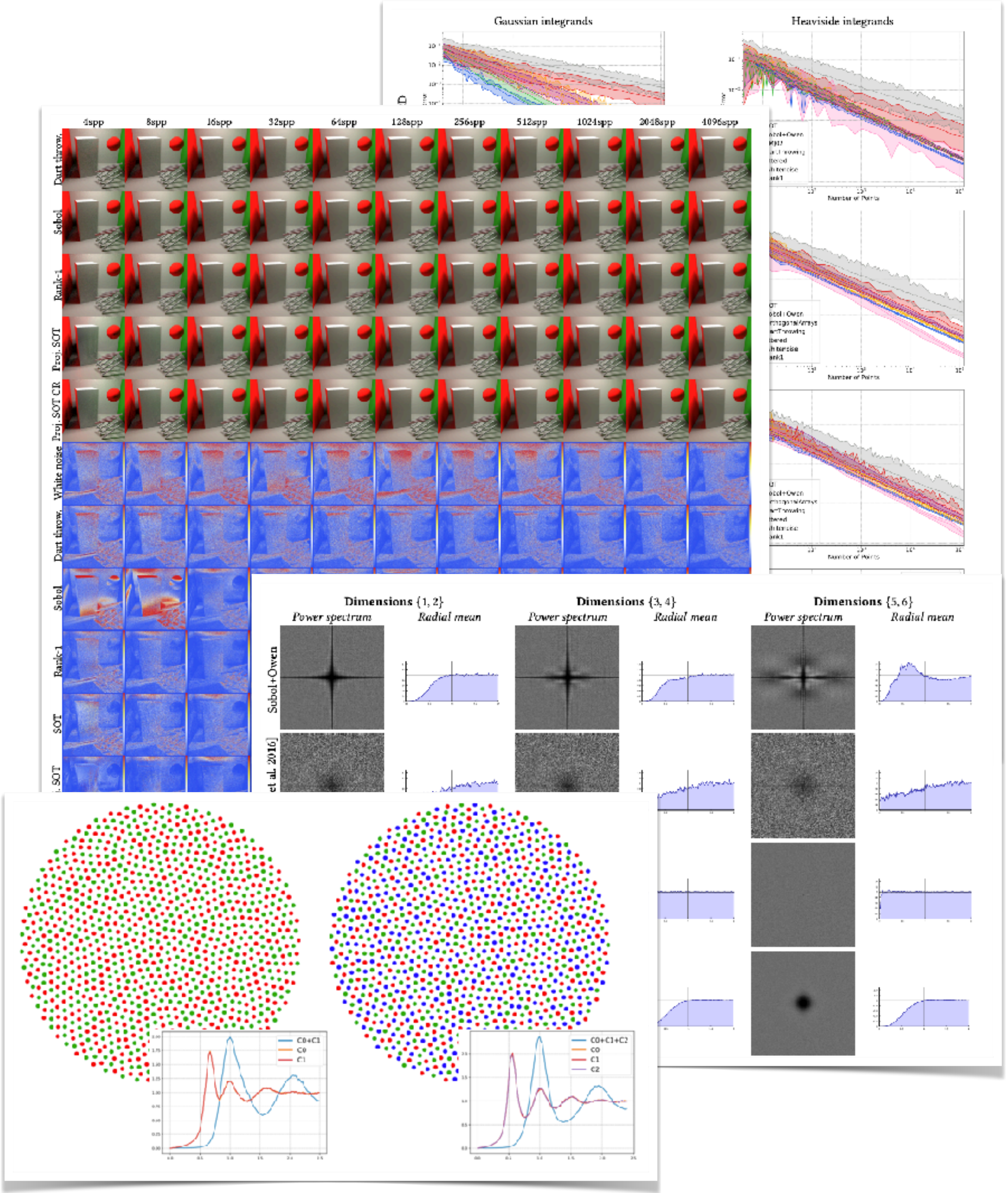
- Volume preserving mapping from s-balls to $[0,1)^s$



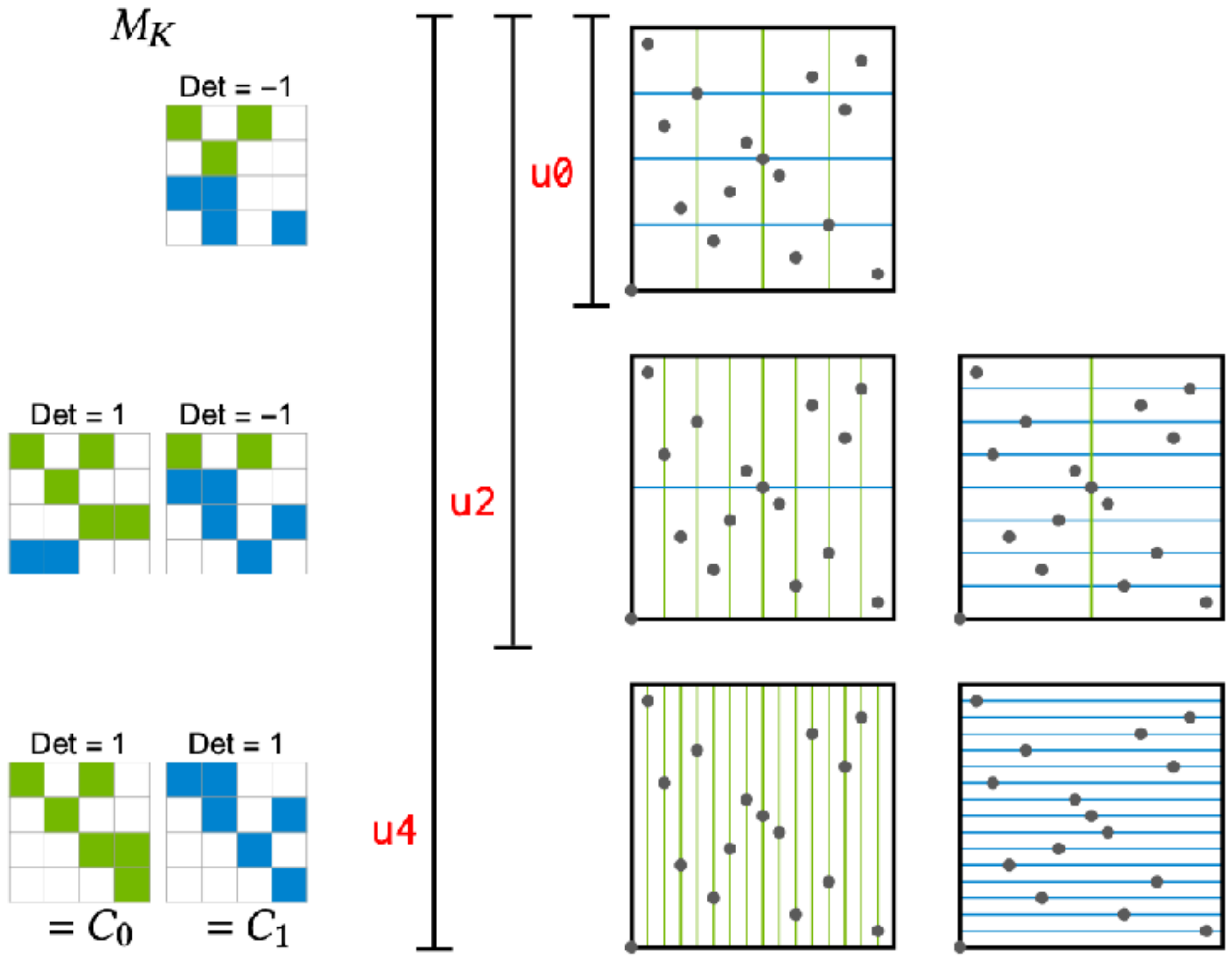
Conclusion

- Fast and effective optimal transport sampling (1Mpts = 200 sec) in high dimension
- Blue noise through Optimal Transport is relevant for MC integration
- Projective sampling for MC rendering

{1,2,3,4,5,6} + {1,2} + {3,4} + {5,6}



MatBuilder: Mastering Sampling Uniformity Over Projections

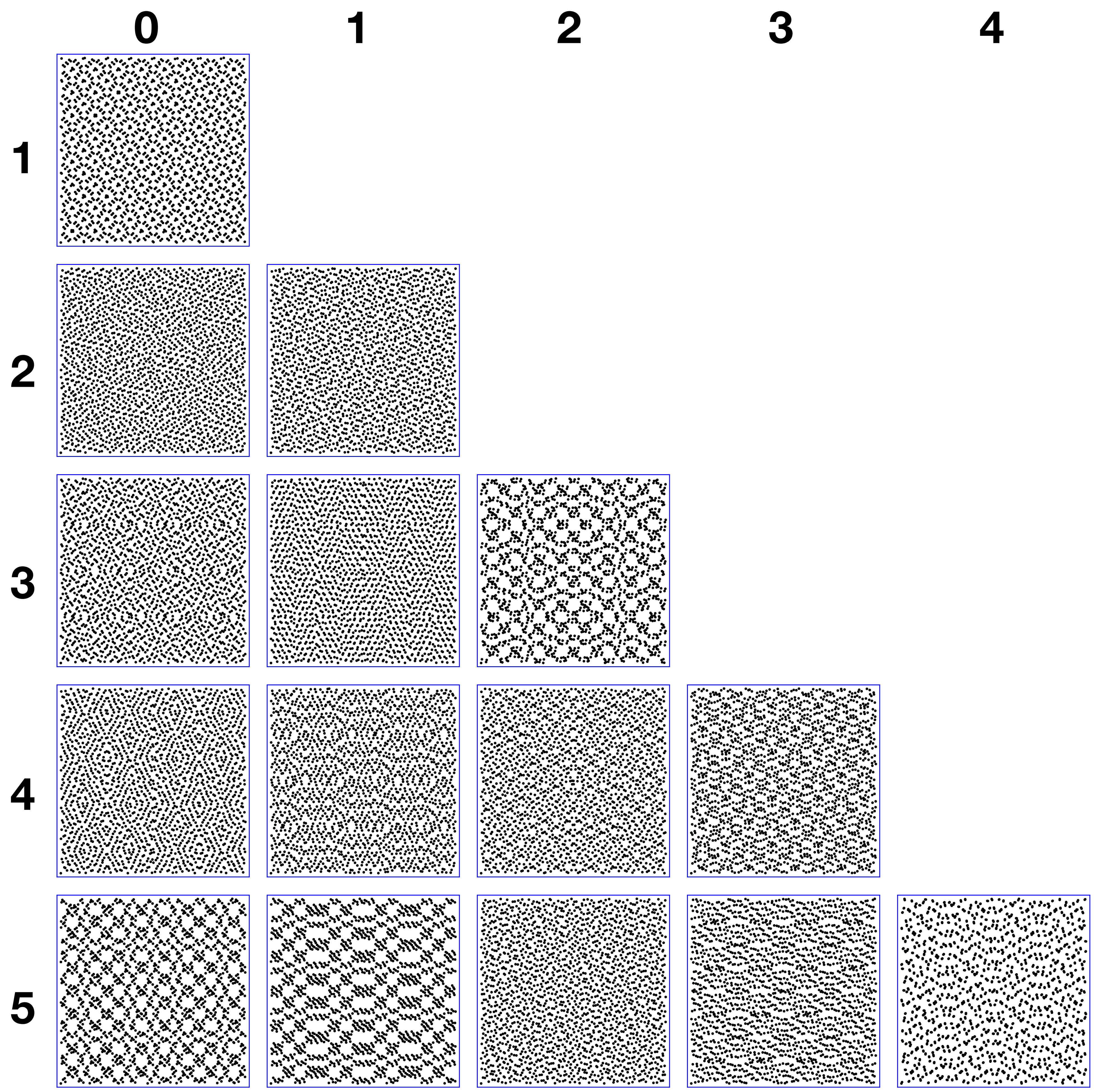



```
s=4
p=3
m=17
# particular stratifications were skipped due to
# lack of solutions
from 3 to 5 stratified 0 1 2 3
from 7 to 9 stratified 0 1 2 3
from 11 to 13 stratified 0 1 2 3
from 15 stratified 0 1 2 3
weak 1 net 0 1 2 3
```

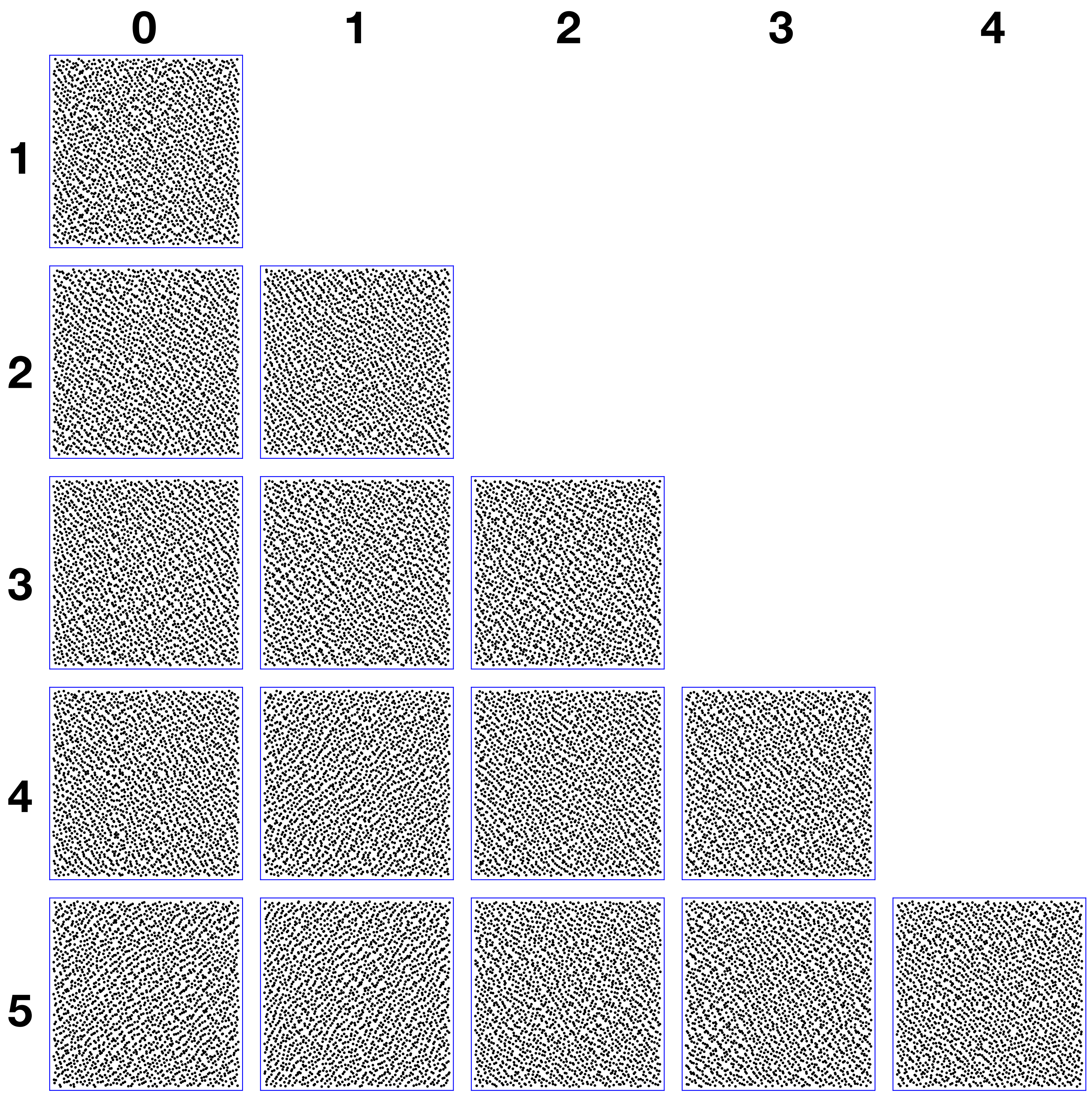
Loïs Paulin, Nicolas Bonneel, David Coeurjolly, Jean-Claude Iehl, Alexander Keller, Victor Ostromoukhov
ACM Transactions on Graphics (Proceedings of SIGGRAPH) July 2022

Low discrepancy sequences are great but..

$$\Delta_n \leq V(f) \cdot O\left(\frac{log(n)^s}{n}\right)$$

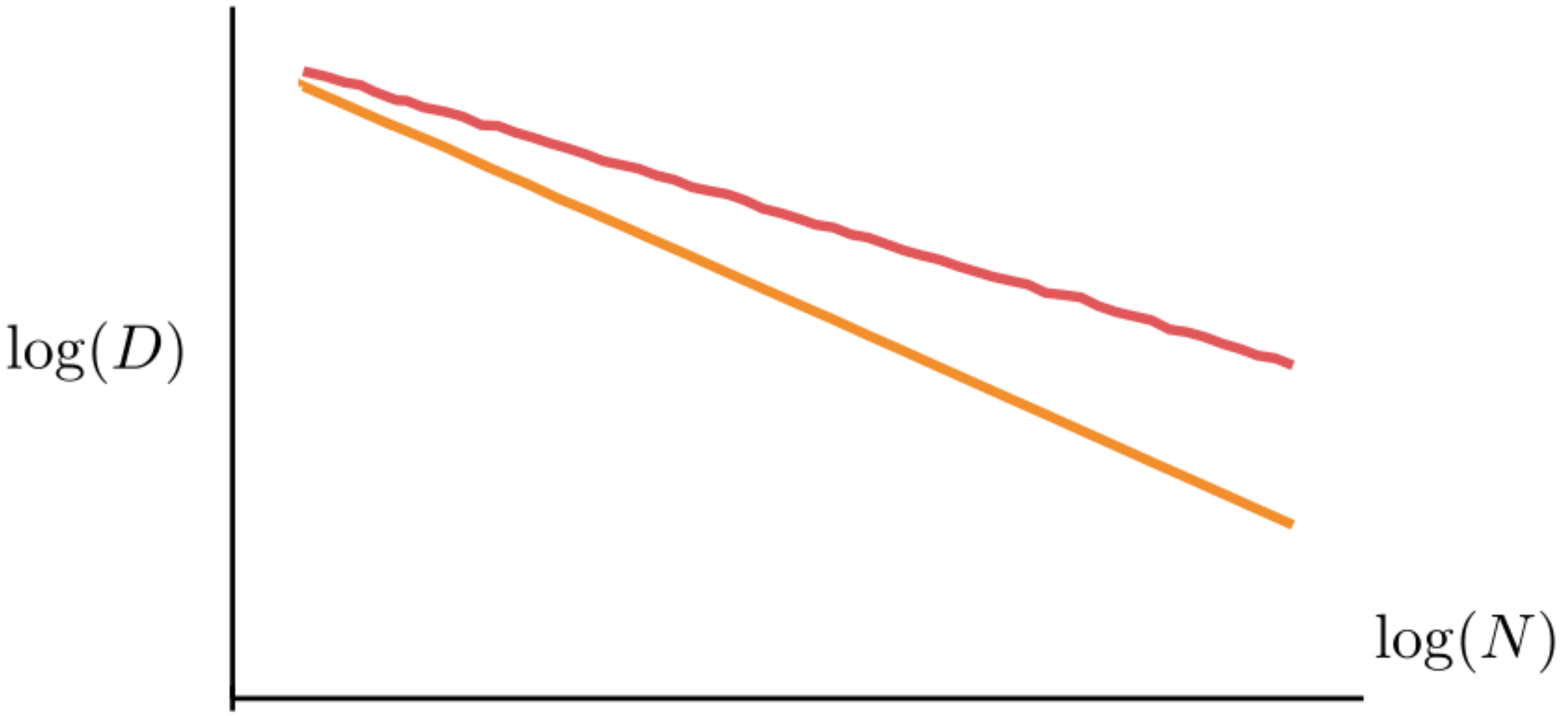


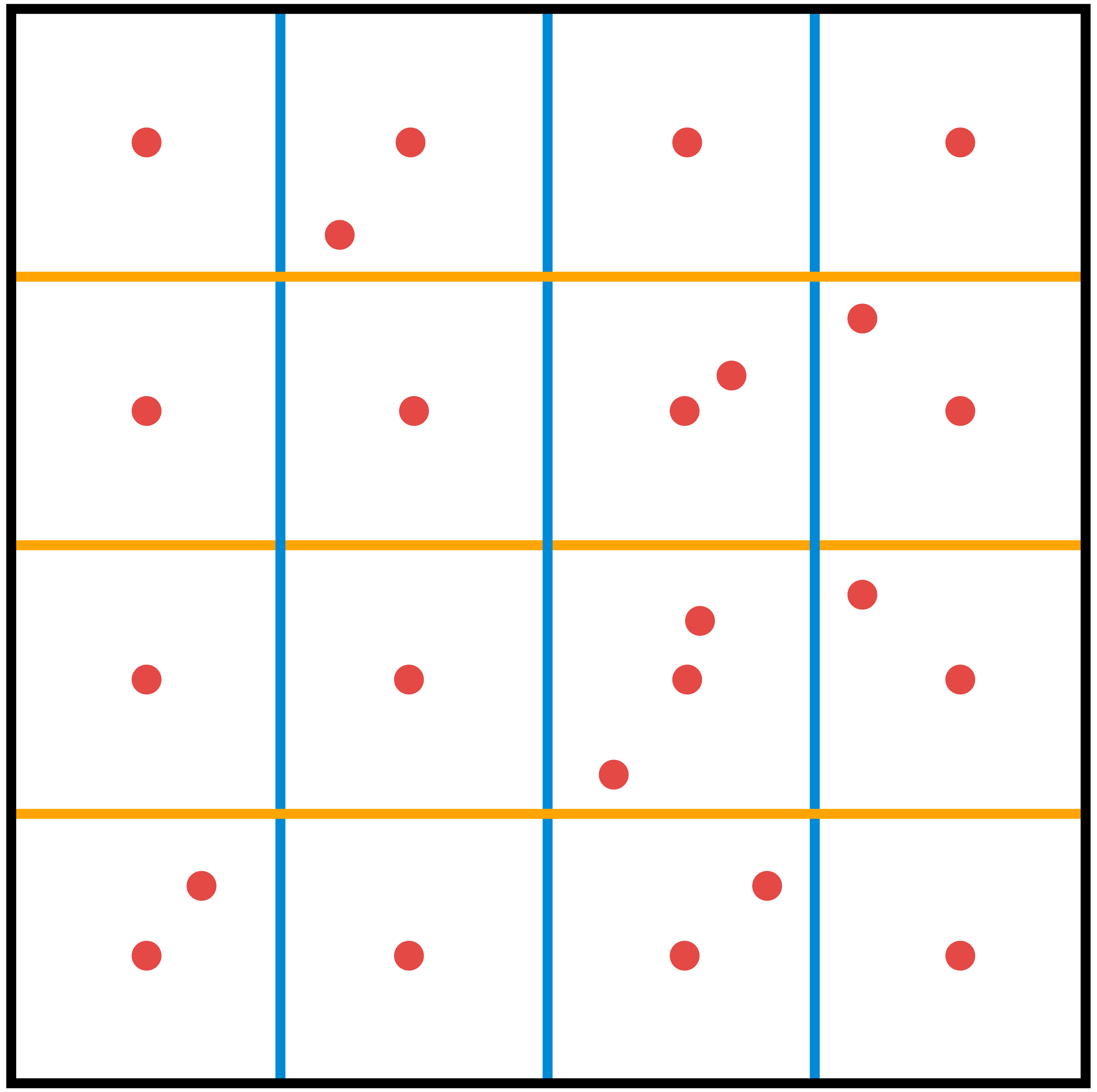
[Sobol]



[Paulin et al 22]

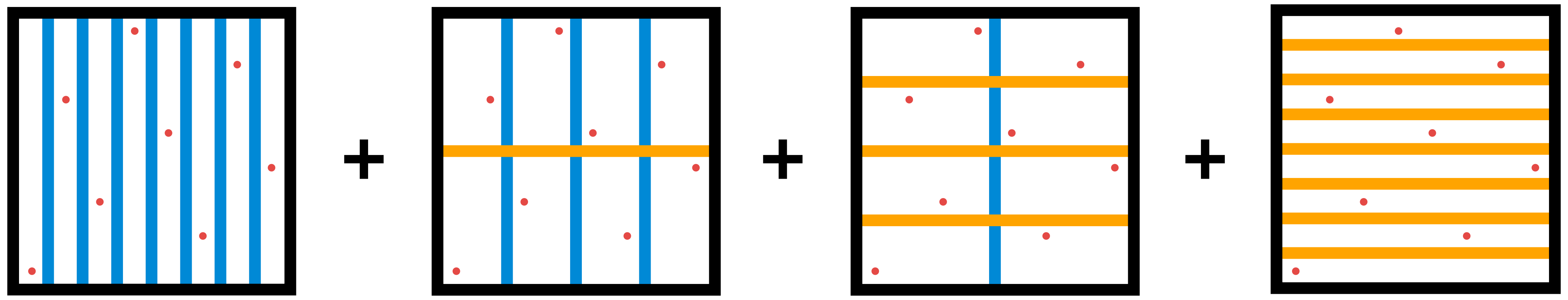
Recap stratified sampling

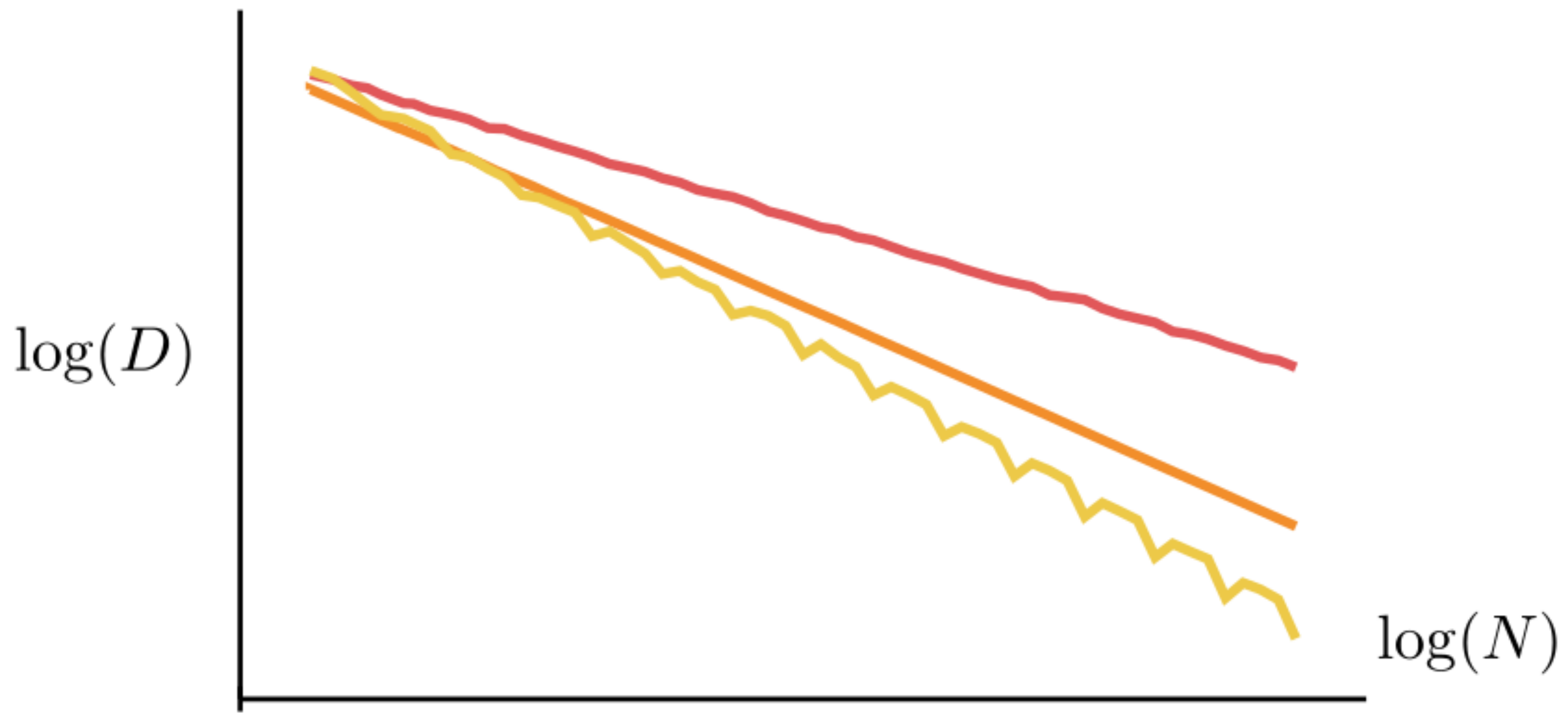


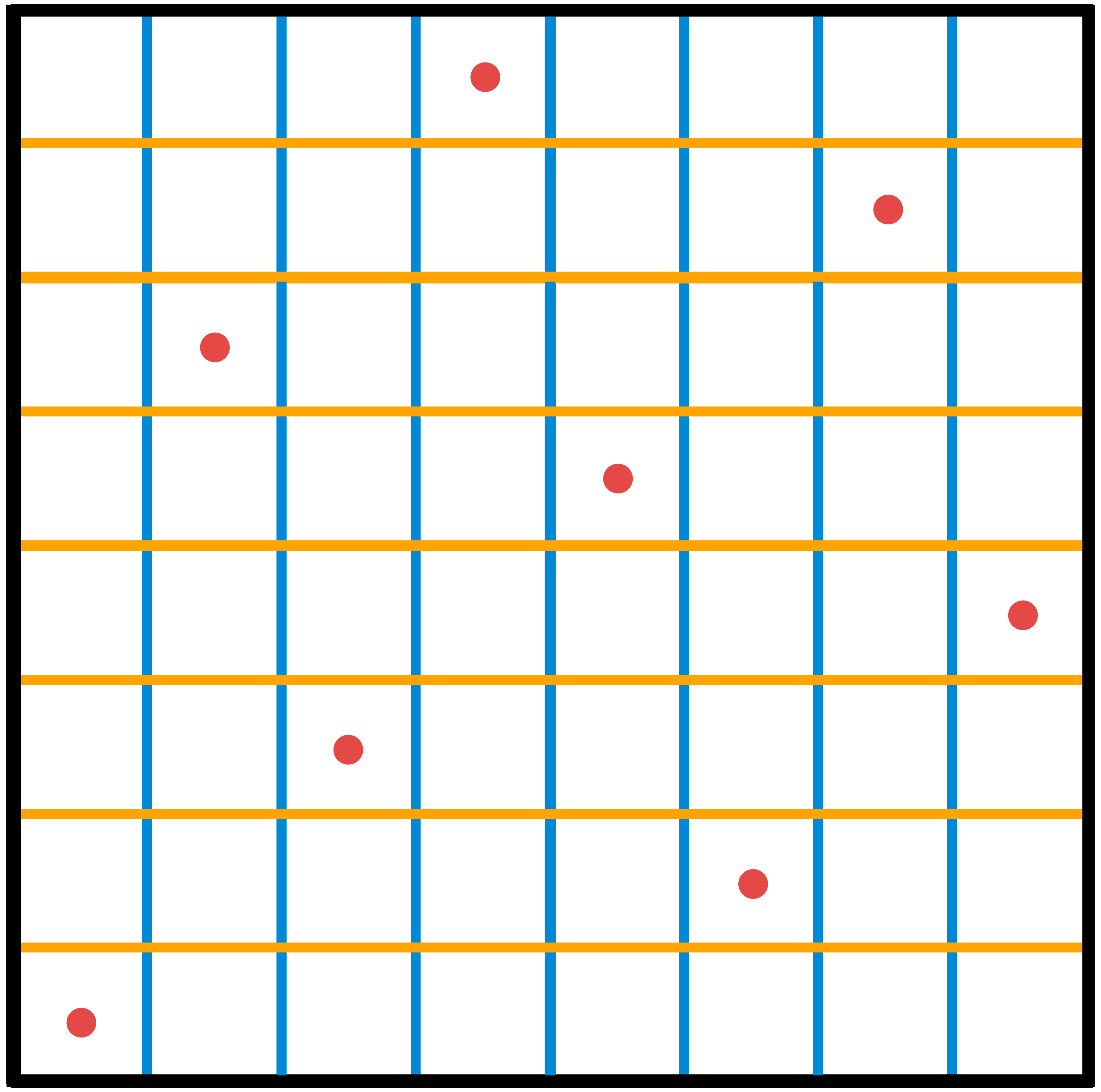


(t,m,s)-nets

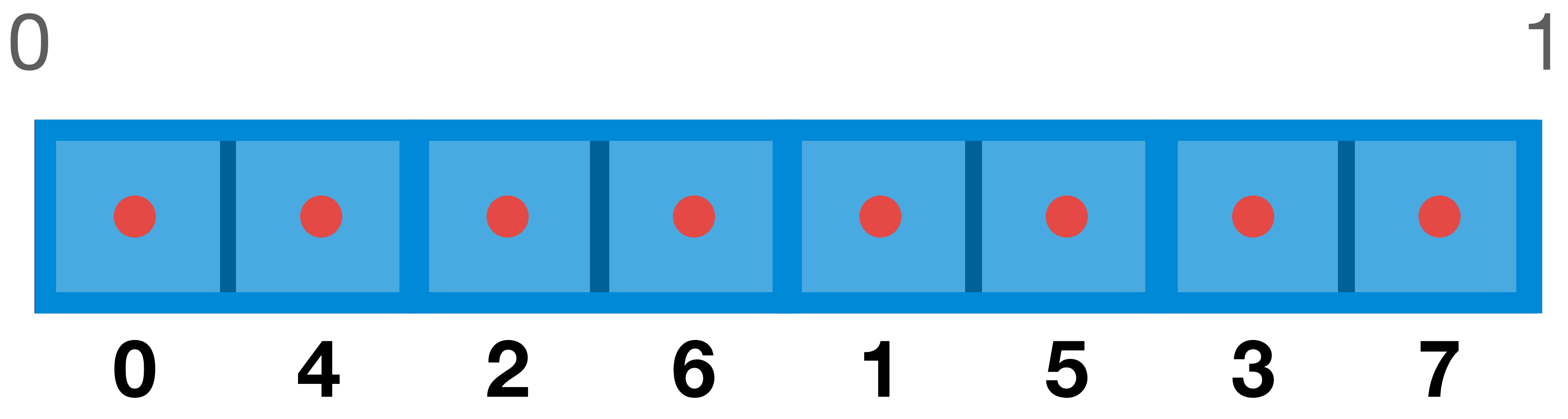
t = 0, m = 3, s = 2, p = 2







Generator Matrices - Algebraic samplers



$$\sigma(a) = b$$
$$\mathbb{N} \to [0,1)$$

$$\begin{bmatrix} c_{1,1} \dots c_{1,m} \\ \vdots \quad\quad \vdots \\ c_{m,1} \dots c_{m,m} \end{bmatrix} \cdot \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix} \leftarrow$$

$$a = (a_m \dots a_2 a_1)_p = \sum_1^m a_i p^i$$

$$b = (0.b_1 b_2 \dots b_m)_p = \sum_{i=1}^m b_i \frac{1}{p^i}$$

$$\mathbb{F}_p^m \to \mathbb{F}_p^m$$

Rows of the matrix *encode* strata

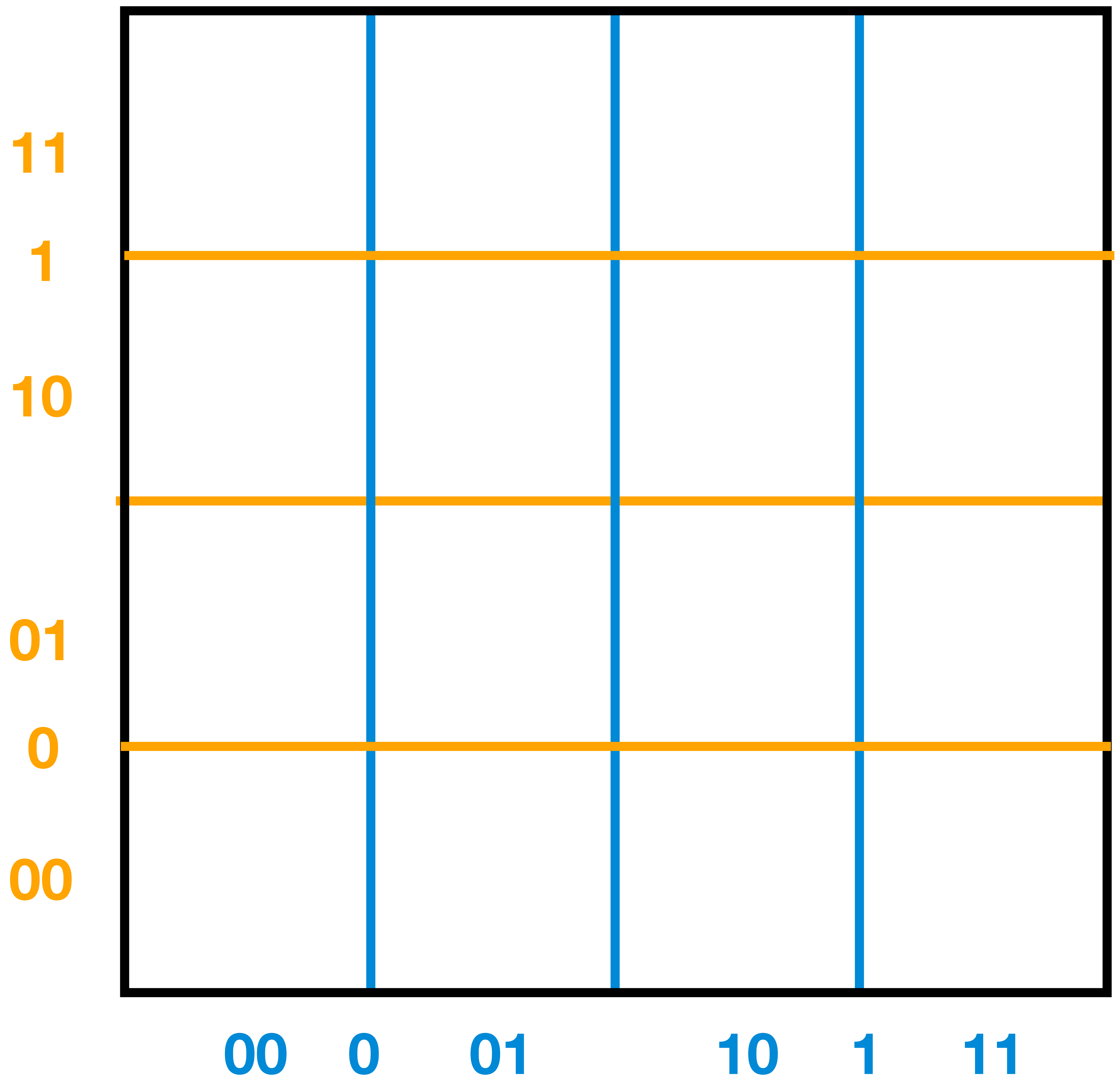
Generator Matrices - Algebraic samplers

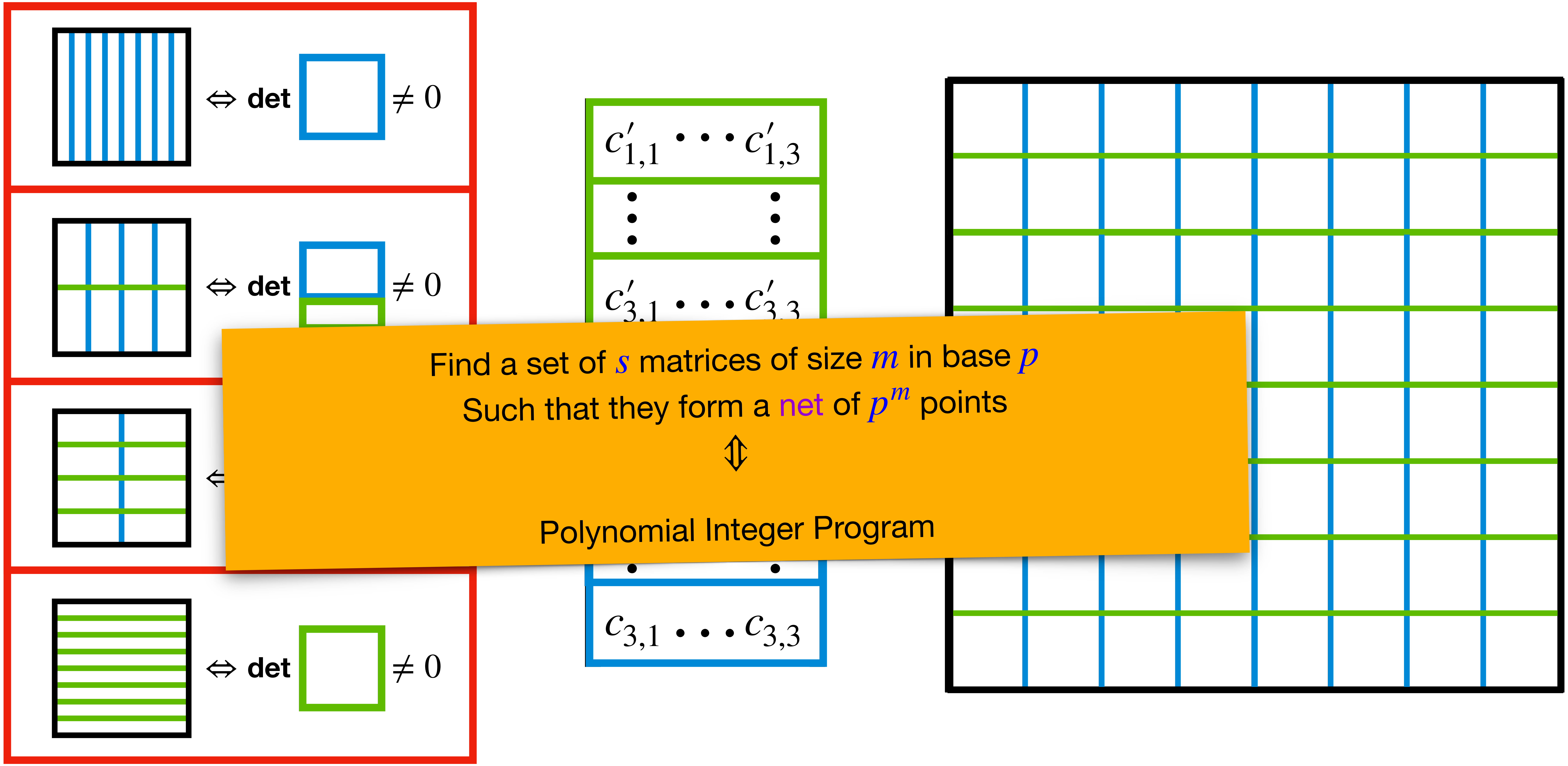
$$\begin{bmatrix} c'_{1,1} & \cdots & c'_{1,3} \\ \vdots & & \vdots \\ c'_{3,1} & \dots & c'_{3,3} \end{bmatrix} \cdot \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} = \begin{bmatrix} b'_1 \\ b'_2 \\ b'_3 \end{bmatrix}$$

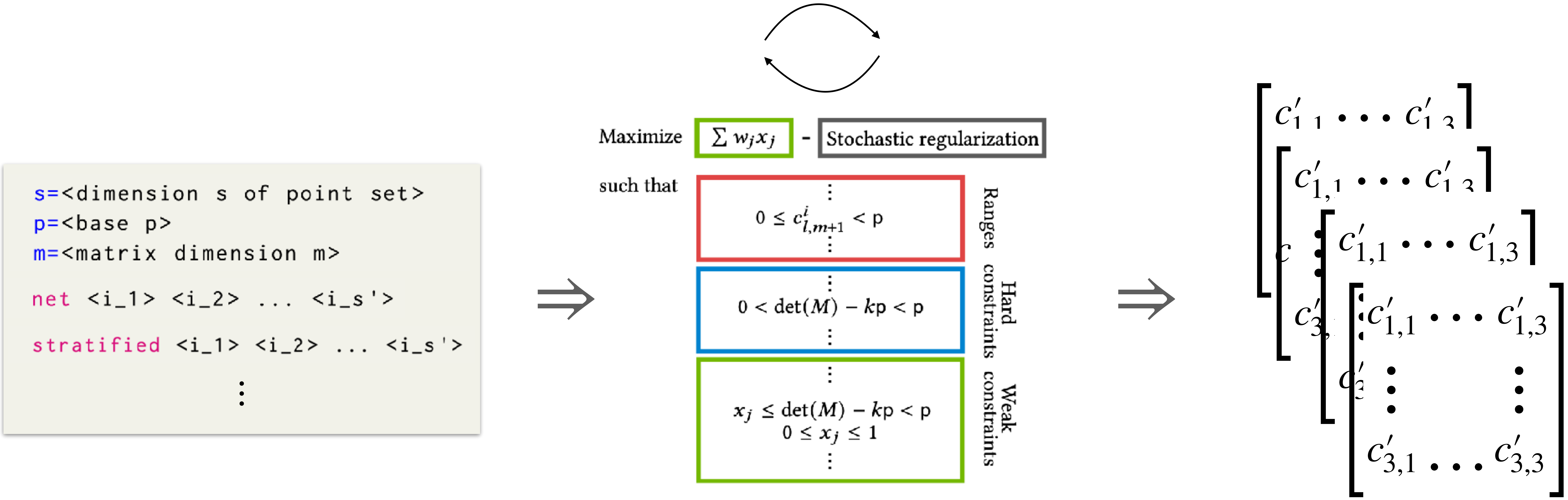
$$\begin{bmatrix} c_{1,1} & \cdots & c_{1,3} \\ & \dots & \end{bmatrix} \quad \begin{bmatrix} a_1 \\ \dots \end{bmatrix} \quad \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$$

Theorem 4.25 [Niederreiter92]

$$\textbf{Stratification} \Leftrightarrow \quad \textbf{det} \begin{bmatrix} c'_{1,1} & \cdots & c'_{1,3} \\ c'_{2,1} & \cdots & c'_{2,3} \\ c_{1,1} & \cdots & c_{1,3} \end{bmatrix} \neq 0$$

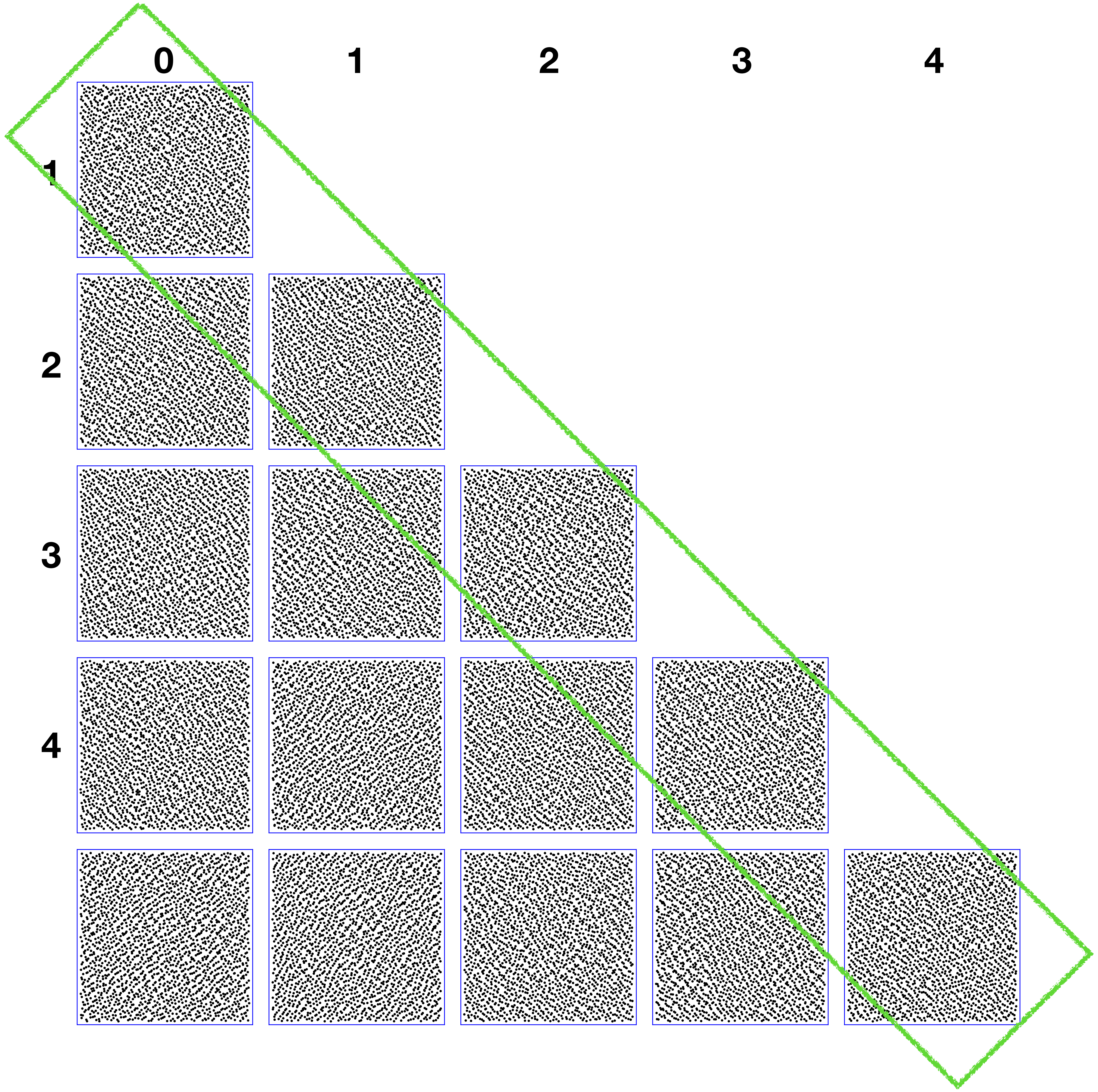



⇔ det ≠ 0
⇔ det ≠ 0
⇔ det ≠ 0
$c'_{1,1} \cdots c'_{1,3}$
$c'_{3,1} \cdots c'_{3,3}$
$c_{3,1} \cdots c_{3,3}$
Find a set of s matrices of size m in base p
Such that they form a net of p^m points
⇕
Polynomial Integer Program



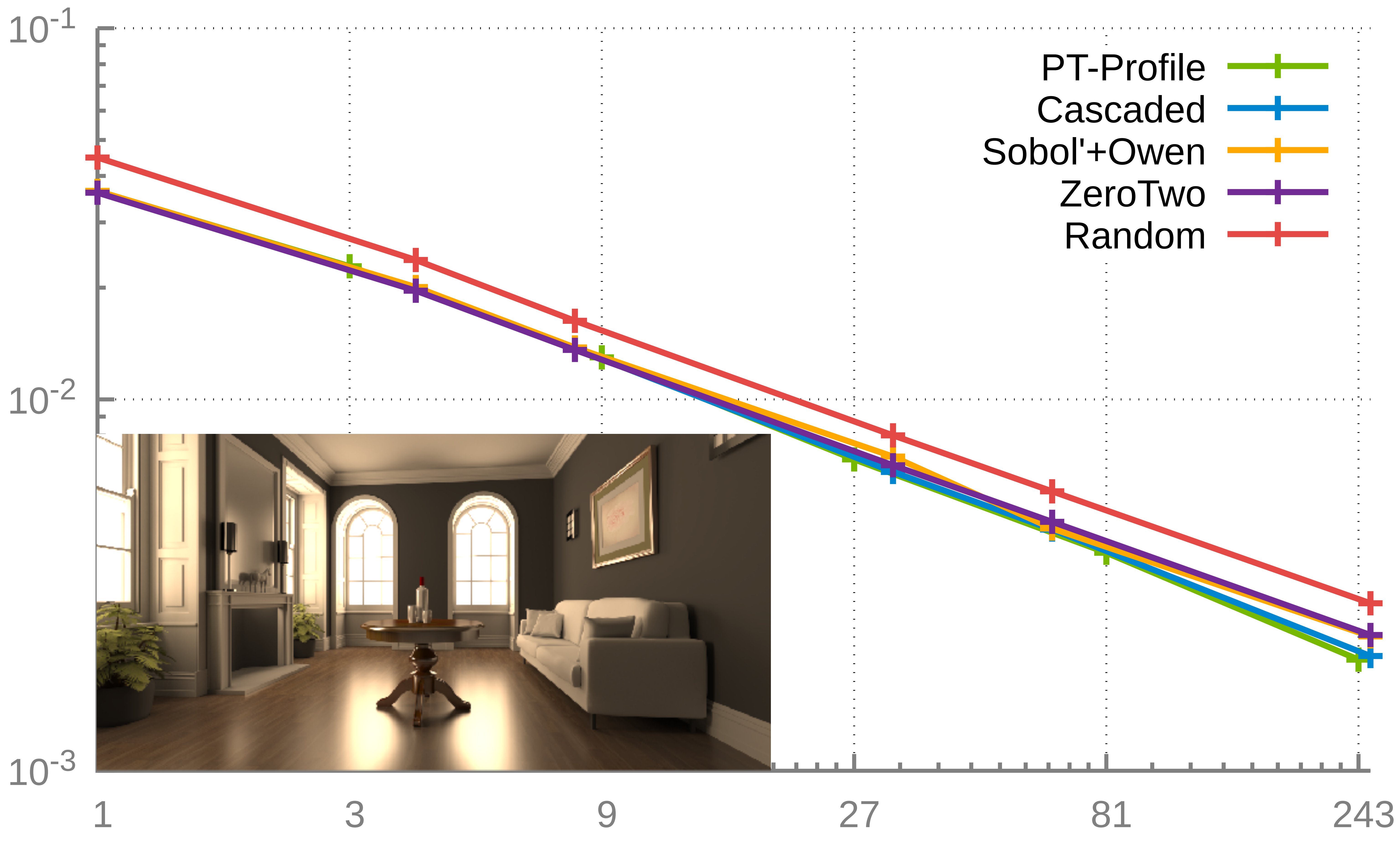
s=<dimension s of point set>
p=<base p>
m=<matrix dimension m>
net <i_1> <i_2> ... <i_s'>
stratified <i_1> <i_2> ... <i_s'>
Maximize
$\sum w_j x_j$
Stochastic regularization
such that
$0 \leq c^i_{l,m+1} < p$
$0 < \det(M) - kp < p$
$x_j \leq \det(M) - kp < p$
$0 \leq x_j \leq 1$
Ranges
Hard constraints
Weak constraints
$c'_{1,1}$ $c'_{1,3}$
$c'_{3,1}$ $c'_{3,3}$

```
#Generic-proj-LDS
s=6
p=3
m=10
net 0 1
net 1 2
net 2 3
net 3 4
net 4 5
weak 1 net 0 2
weak 1 net 0 3
weak 1 net 0 4
weak 1 net 0 5
weak 1 net 1 3
weak 1 net 1 4
weak 1 net 1 5
weak 1 net 2 4
weak 1 net 2 5
weak 1 net 3 5
```



MC Rendering with projective sampler

```
#PT-Profile
p=3
s=6
m=12
net 0 1
net 1 2
net 2 3
net 3 4
net 4 5
weak 1 net u4 0 1 2
weak 1 net u4 1 2 3
weak 1 net u4 2 3 4
weak 1 net u4 3 4 5
weak 1 net u2 0 1 2 3
weak 1 net u2 1 2 3 4
weak 1 net u2 2 3 4 5
weak 1 net u2 0 1 2 3 4
weak 1 net u2 1 2 3 4 5
weak 1 net u4 0 1 2 3 4 5
```



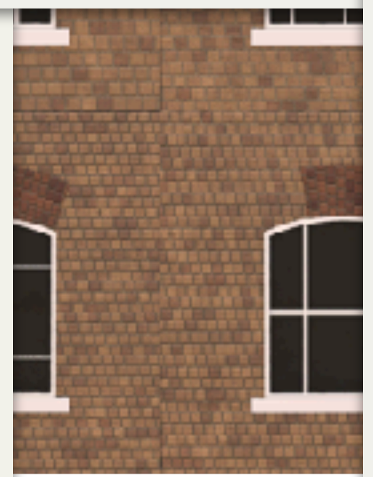
Experiment design

```
s=4
p=3
m=17
# particular stratifications were
# lack of solutions
from 3 to 5 stratified 0 1
from 7 to 9 stratified 0 1
from 11 to 13 stratified 0
from 15 stratified 0 1 2 3
weak 1 net 0 1 2 3
```

```
#Generic-OA
s=9
p=3
m=10
from 3 stratified 0 1 2
from 3 stratified 1 2 3
from 3 stratified 2 3 4
from 3 stratified 3
from 3 stratified 4
from 3 stratified 5
from 3 stratified 6
from 5 weak 1 strati
```

```
#Gener
s=6
p=3
m=10
net 0
net 1
net 2
net 3
net 4 5
weak 1 net 0 2
weak 1 net 0 3
weak 1 net 0 4
weak 1 net 0 5
weak 1 net 1 3
weak 1 net 1 4
weak 1 net 1 5
weak 1 net 2 4
weak 1 net 2 5
weak 1 net 3 5
```

Good Sam



```
#Mixed
s=10
p=3
m=10
net 0 1
net 1 2
net 2 3
net 3 4
net 4 5
from 3 stratified 0 1 2
from 4 stratified 1 2 3
from 3 stratified 2 3 4
from 4 stratified 3 4 5
from 4 to 6 stratified 0 1 2 3 4
from 4 to 6 stratified 1 2 3 4 5
```

es

```
s=7
p=2
m=5
# (Brick-Amount-X, Brick-Amount-Y)
net 0 1
# (Window-Brace-Amount-X, Window-Brace-Width-X,
#  Window-Brace-Amount-Y, Window-Brace-Width-Y)
weak 1 net 3 4 5 6
# Overall uniformity, including Brick-Lintel-Wi
weak 2 net 0 1 2 3 4 5 6
```

Conclusion

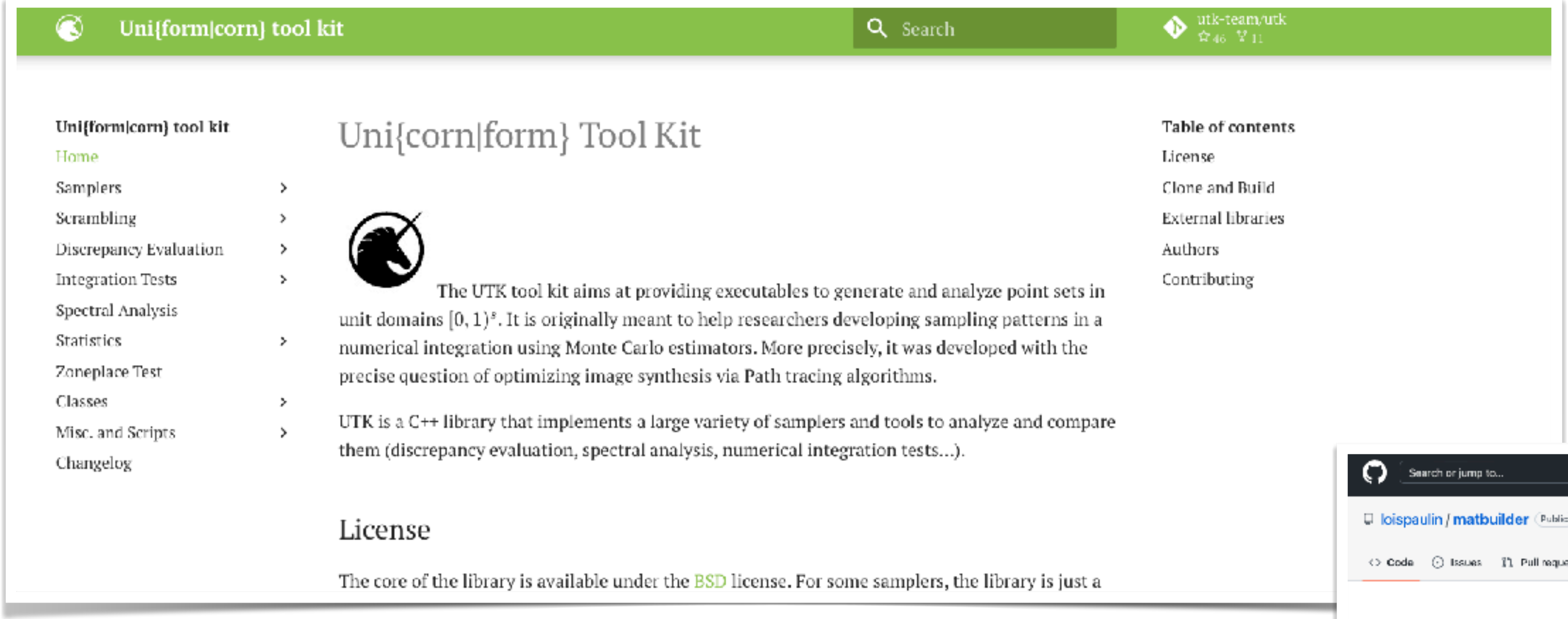
Correlated Samplers in Computer Graphics

- Practical constraints: variance reduction in MC rendering, high dimension, fast samplers, controllable subspaces, adaptive

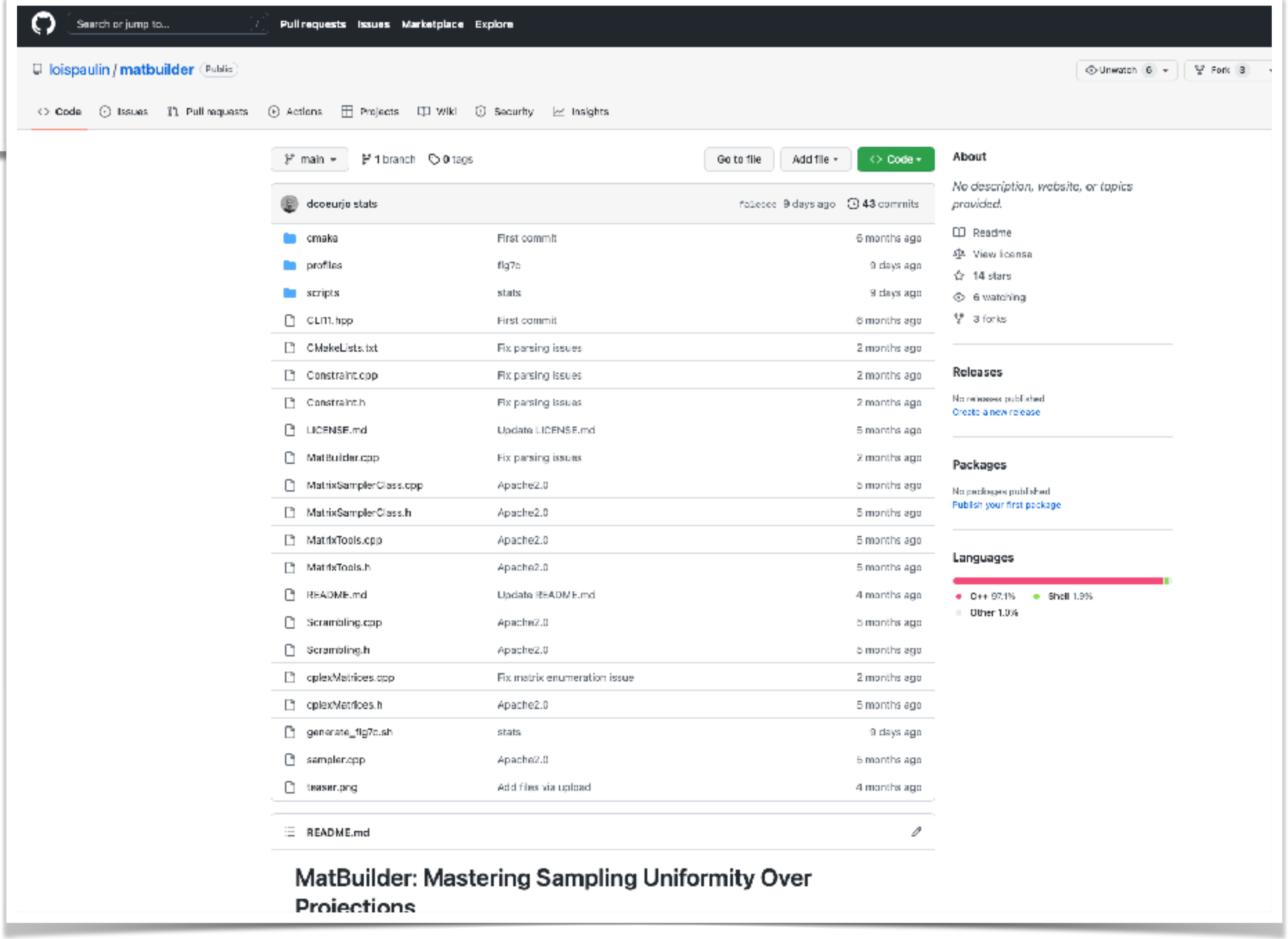
- Various theoretical approaches: stochastic point processes, spectral analysis, discrepancy, discrete mathematics, optimization...

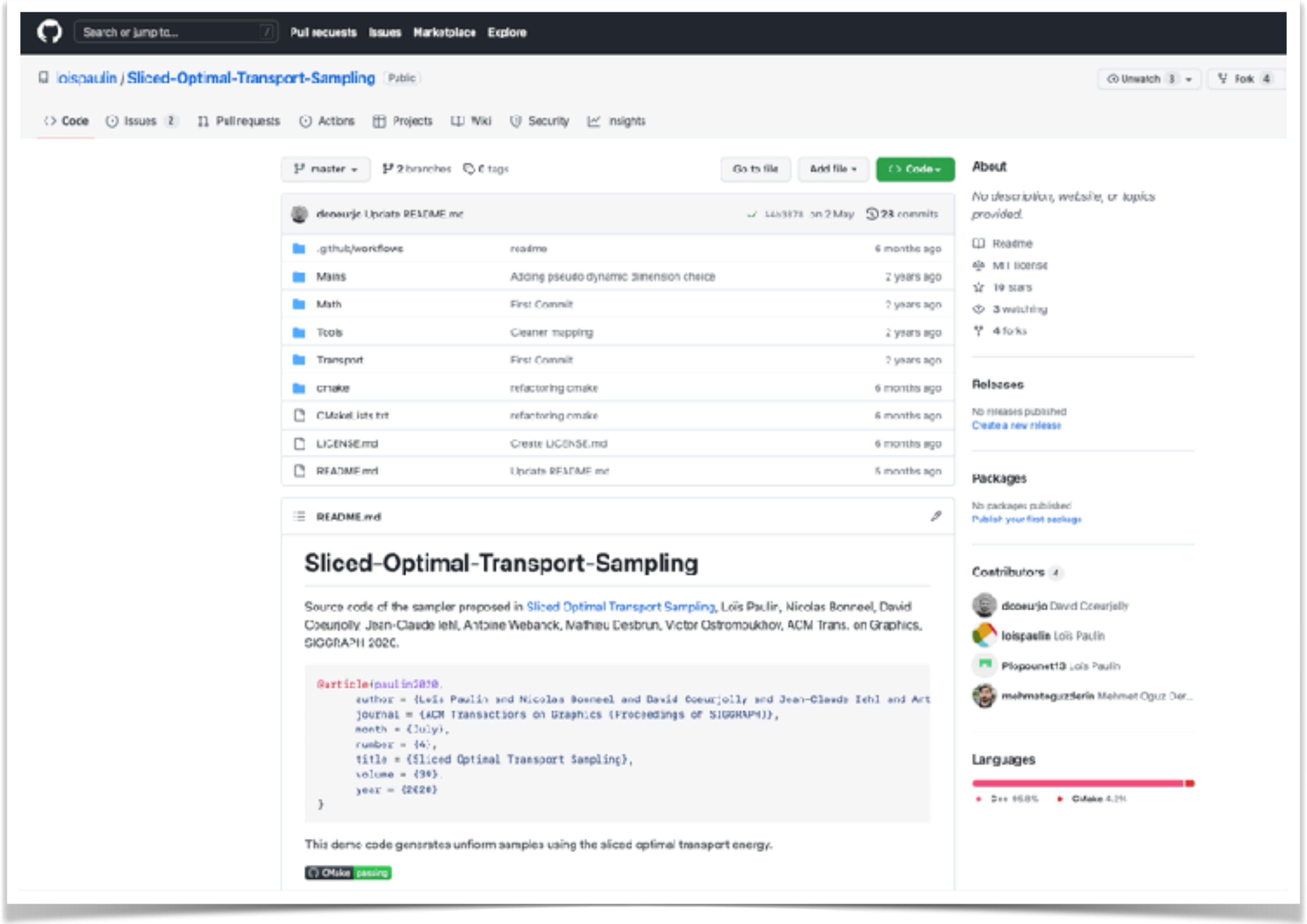
- Future works:
    - Still room between (projective) optimized LDS and blue noise properties (for low sample counts)
    - Many possible improvements in MC Rendering
        - By focusing on special classes of integrands
        - By carefully integrating with advanced techniques (canonical vs. actual integration domain, —resampled— importance sampling, splitting, control variates, screen space diffusion...)

https://utk-team.github.io/utk/



https://github.com/loispaulin/matbuilder





https://github.com/loispaulin/Sliced-Optimal-Transport-Sampling