On sampling determinantal point processes

Guillaume Gautier

Ph.D. defense

Advisors: Rémi Bardenet, Michal Valko

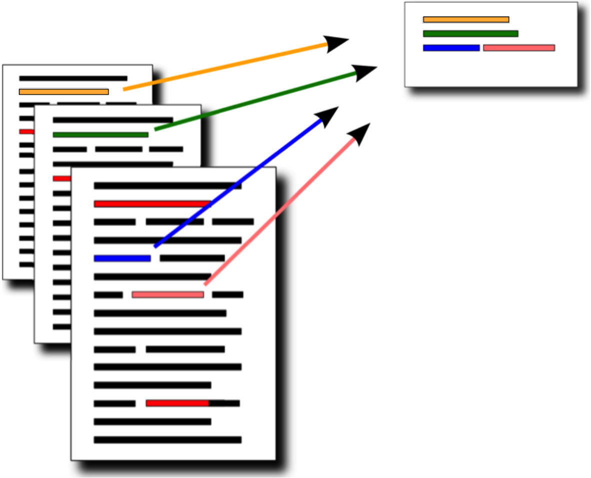
May 19, 2020

guilgautier.github.io

github.com/guilgautier/DPPy

Text summarization

Extract **diverse** sentences of a large corpus to build a representative summary.



Recommendation systems

Two possible sets of answers of an image search engine to the query "bolt".

relevance

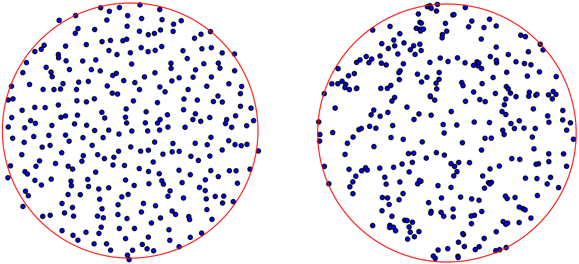
relevance
+
diversity

- Use DPPs to enforce diversity among the recommended items.

Numerical integration

Use random **repulsive** points as quadrature nodes to estimate an integral

$$\int f(x)\mu(\mathrm{d}x) \approx \sum_{n=1}^{N} \omega_n f(x_n).$$

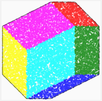


Bardenet and Hardy (2016, 2020)

- Prove faster rate of convergence with DPP points than i.i.d. points.
- Efficient sampler to put theory into practice**?**

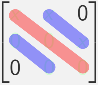
My Ph.D. in a nutshell

Finite projection DPP
Approximate sampling
Linear programming



ICML, 2017

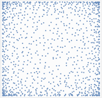
β-ensembles
Gibbs sampling
Random matrices

eigvals 

Submitted, 2020

DPP sampling

Monte Carlo integration
Exact sampling
Random linear system



NeurIPS, 2019

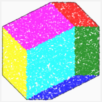
Python toolbox
DPPy
Reproducible research

```
from dppy import *
# [...]
dpp.sample()
```

JMLR-MLOSS, 2019

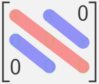
My Ph.D. in a nutshell

Finite projection DPP
Approximate sampling
Linear programming



ICML, 2017

β-ensembles
Gibbs sampling
Random matrices

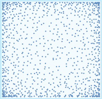
eigvals 

Submitted, 2020

DPP sampling

Continuous setting

Monte Carlo integration
Exact sampling
Random linear system



NeurIPS, 2019

Python toolbox
DPPy
Reproducible research

```
from dppy import *
# [...]
dpp.sample()
```

JMLR-MLOSS, 2019

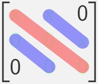
My Ph.D. in a nutshell

Finite setting

Finite projection DPP
Approximate sampling
Linear programming

ICML, 2017

β-ensembles
Gibbs sampling
Random matrices

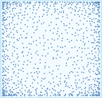
eigvals 

Submitted, 2020

DPP sampling

Continuous setting

Monte Carlo integration
Exact sampling
Random linear system



NeurIPS, 2019

Python toolbox
DPPy
Reproducible research

```
from dppy import *
# [...]
dpp.sample()
```

JMLR-MLOSS, 2019

My Ph.D. in a nutshell

Finite setting

Finite projection DPP
Approximate sampling
Linear programming

ICML, 2017

Continuous setting

β-ensembles
Gibbs sampling
Random matrices

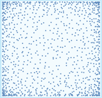
eigvals 

Submitted, 2020

DPP sampling

Continuous setting

Monte Carlo integration
Exact sampling
Random linear system



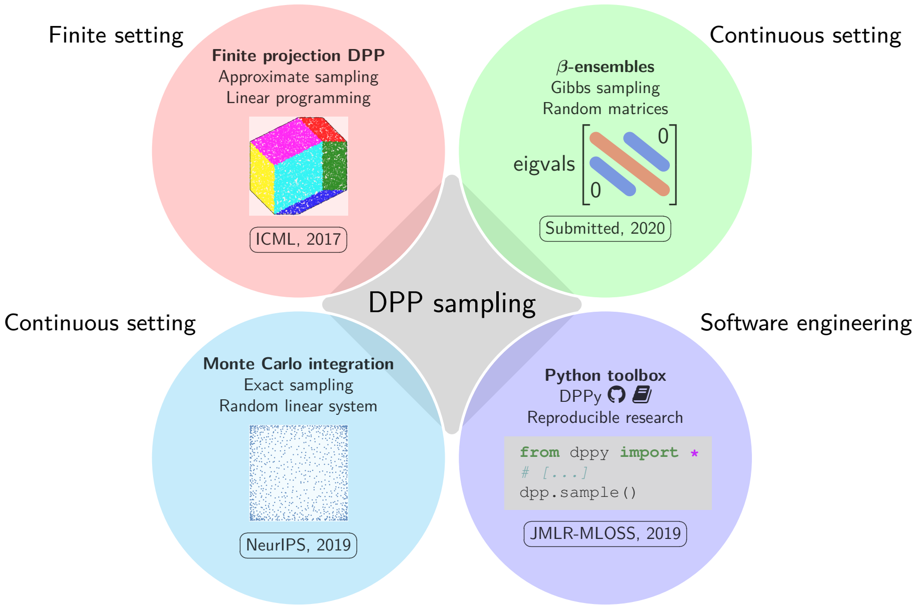
NeurIPS, 2019

Python toolbox
DPPy
Reproducible research

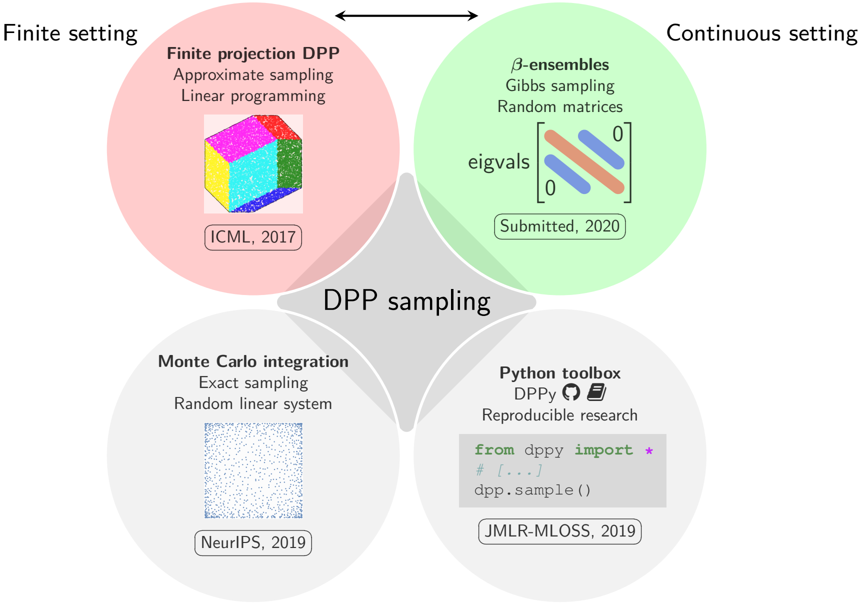
```python
from dppy import *
# [...]
dpp.sample()
```

JMLR-MLOSS, 2019

My Ph.D. in a nutshell



Focus of the presentation

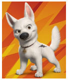


Overview

Some insights on finite DPPs

- Point process

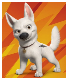
$\mathcal{X} \subset \{$ , , 

, , , $\cdots$,  $\}$

Some insights on finite DPPs

- Point process

$\mathcal{X} \subset \{$ , , 

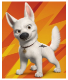
, , , $\cdots$,  $\}$

- Diversity

$\mathbb{P}[\{$ ,  $\} \subset \mathcal{X}] \geq \mathbb{P}[\{$ ,  $\} \subset \mathcal{X}]$

Some insights on finite DPPs

- Point process

$\mathcal{X} \subset \{$ , , , , , $\cdots$,  $\}$

- Diversity

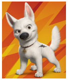
$\mathbb{P}[\{$ ,  $\} \subset \mathcal{X}] \geq \mathbb{P}[\{$ ,  $\} \subset \mathcal{X}]$

- Similarity matrix

$\mathbf{K}$  

Some insights on finite DPPs

- Point process

$\mathcal{X} \subset \{$  ,  ,  ,  ,  , $\cdots$,  $\}$

- Diversity

$\mathbb{P}[\{$  ,  $\} \subset \mathcal{X}] \geq \mathbb{P}[\{$  ,  $\} \subset \mathcal{X}]$

- Similarity matrix

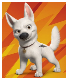
$\mathbf{K}$  

- Inclusion probabilities

$\mathbb{P}[\{$  ,  $\} \subset \mathcal{X}] = \det$ 

Some insights on finite DPPs

- Point process

$\mathcal{X} \subset \{$  $, \cdots ,$  $\}$

- Diversity

$\mathbb{P}[\{$  ,  $\} \subset \mathcal{X}] \geq \mathbb{P}[\{$  ,  $\} \subset \mathcal{X}]$

- Similarity matrix

$\mathbf{K}$ 

- Inclusion probabilities

$\mathbb{P}[\{$  ,  $\} \subset \mathcal{X}] = \det$ 

- Sufficient conditions for existence

$$\mathbf{K}^\top = \mathbf{K} \quad \text{and} \quad 0 \preceq \mathbf{K} \preceq I.$$

Overview

Finite projection DPP

Definition

Consider $\mathbf{K} \in \mathbb{R}^{M \times M}$ such that $\mathbf{K}^\top = \mathbf{K}$ and $\mathbf{K}^2 = \mathbf{K}$.
The point process $\mathcal{X}$ defined by

$$\mathbb{P}[S \subset \mathcal{X}] = \det \mathbf{K}_S, \qquad \forall S \subset \{1, \ldots, M\},$$

is called a projection DPP with kernel $\mathbf{K}$.

Finite projection DPP

Definition

Consider $\mathbf{K} \in \mathbb{R}^{M \times M}$ such that $\mathbf{K}^{\mathsf{T}} = \mathbf{K}$ and $\mathbf{K}^2 = \mathbf{K}$.
The point process $\mathcal{X}$ defined by

$$\mathbb{P}[S \subset \mathcal{X}] = \det \mathbf{K}_S, \qquad \forall S \subset \{1, \dots, M\},$$

is called a projection DPP with kernel $\mathbf{K}$.

- Samples have fixed cardinality

$$N \triangleq |\mathcal{X}| = \operatorname{rank} \mathbf{K}.$$

Finite projection DPP

Definition

Consider $\mathbf{K} \in \mathbb{R}^{M \times M}$ such that $\mathbf{K}^\top = \mathbf{K}$ and $\mathbf{K}^2 = \mathbf{K}$.
The point process $\mathcal{X}$ defined by

$$\mathbb{P}[S \subset \mathcal{X}] = \det \mathbf{K}_S, \qquad \forall S \subset \{1, \ldots, M\},$$

is called a projection DPP with kernel $\mathbf{K}$.

- Samples have fixed cardinality

$$N \triangleq |\mathcal{X}| = \operatorname{rank} \mathbf{K}.$$

- The likelihood reads

$$\mathbb{P}[\mathcal{X} = B] = \det \mathbf{K}_B \; \mathbb{1}_{|B|=N}.$$

Finite projection DPP

Example

Consider the $N \times M$ feature matrix $\mathbf{\Phi} = \left[\phi_1, \ldots, \phi_M\right]$, such that $\underline{\operatorname{rank} \mathbf{\Phi} = N}$, and build the kernel

$$\mathbf{K} = \mathbf{\Phi}^\top [\mathbf{\Phi}\mathbf{\Phi}^\top]^{-1}\mathbf{\Phi}.$$

Finite projection DPP

Example

Consider the $N \times M$ feature matrix $\mathbf{\Phi} = \left[\phi_1, \dots, \phi_M\right]$, such that $\underline{\operatorname{rank} \mathbf{\Phi} = N}$, and build the kernel

$$\mathbf{K} = \mathbf{\Phi}^\top [\mathbf{\Phi}\mathbf{\Phi}^\top]^{-1} \mathbf{\Phi}.$$

- The likelihood reads

$$\mathbb{P}[\mathcal{X} = B] = \frac{(\det \mathbf{\Phi}_{:B})^2}{\det \mathbf{\Phi}\mathbf{\Phi}^\top} \mathbb{1}_{|B|=N}.$$

Finite projection DPP

Example

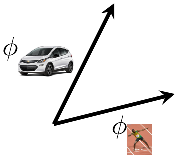
Consider the $N \times M$ feature matrix $\mathbf{\Phi} = \left[\phi_1, \dots, \phi_M\right]$, such that rank $\mathbf{\Phi} = N$, and build the kernel

$$\mathbf{K} = \mathbf{\Phi}^\top [\mathbf{\Phi}\mathbf{\Phi}^\top]^{-1} \mathbf{\Phi}.$$

- The likelihood reads

$$\mathbb{P}[\mathcal{X} = B] = \frac{(\det \mathbf{\Phi}_{:B})^2}{\det \mathbf{\Phi}\mathbf{\Phi}^\top} \, \mathbb{1}_{|B|=N}.$$

- Geometrically, e.g., for $N = 2$,



Finite projection DPP

Example

Consider the $N \times M$ feature matrix $\mathbf{\Phi} = \left[\phi_1, \dots, \phi_M\right]$, such that $\underline{\operatorname{rank} \mathbf{\Phi} = N}$, and build the kernel

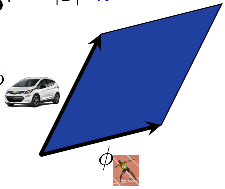
$$\mathbf{K} = \mathbf{\Phi}^{\mathsf{T}} [\mathbf{\Phi}\mathbf{\Phi}^{\mathsf{T}}]^{-1} \mathbf{\Phi}.$$

- The likelihood reads

$$\mathbb{P}[\mathcal{X} = B] = \frac{(\det \mathbf{\Phi}_{:B})^2}{\det \mathbf{\Phi}\mathbf{\Phi}^{\mathsf{T}}} \mathbb{1}_{|B|=N}.$$

- Geometrically, e.g., for $N = 2$,

$\mathbb{P}[\mathcal{X} = \{$ ,  $\}] \propto \text{volume}^2$



Finite projection DPP

Example

Consider the $N \times M$ feature matrix $\mathbf{\Phi} = \left[\phi_1, \ldots, \phi_M\right]$, such that $\underline{\operatorname{rank} \mathbf{\Phi} = N}$, and build the kernel

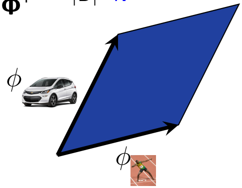
$$\mathbf{K} = \mathbf{\Phi}^{\top}[\mathbf{\Phi}\mathbf{\Phi}^{\top}]^{-1}\mathbf{\Phi}.$$

- The likelihood reads

$$\mathbb{P}[\mathcal{X} = B] = \frac{(\det \mathbf{\Phi}_{:B})^2}{\det \mathbf{\Phi}\mathbf{\Phi}^{\top}} \mathbb{1}_{|B|=N}.$$

- Geometrically, e.g., for $N = 2$,

$\mathbb{P}[\mathcal{X} = \{$  ,  $\}] \propto \text{volume}^2$



- The support is formed by collections of columns of $\mathbf{\Phi}$ forming a basis of $\mathbb{R}^N$,

$$\mathcal{B} \triangleq \{B\,;\ |B| = N, \text{ and } \det \mathbf{\Phi}_{:B} \neq 0\}.$$

Overview

Continuous projection DPP

Definition

Let $\phi_0, \dots, \phi_{N-1}$ be orthonormal functions in $L^2(\mathbb{X}, \mu)$ and

$$K(x, y) = \sum_{k=0}^{N-1} \phi_k(x)\phi_k(y).$$

Take $(x_1, \dots, x_N)$ with joint probability distribution

$$\frac{1}{N!} \det[K(x_i, x_j)]_{i,j=1}^{N} \prod_{n=1}^{N} \mu(\mathrm{d}x_n).$$

Then $\mathcal{X} \triangleq \{x_1, \dots, x_N\} \subset \mathbb{X}$ defines a projection DPP with kernel K.

Continuous projection DPP

Definition

Let $\phi_0, \dots, \phi_{N-1}$ be orthonormal functions in $L^2(\mathbb{X}, \mu)$ and

$$K(x, y) = \sum_{k=0}^{N-1} \phi_k(x) \phi_k(y).$$

Take $(x_1, \dots, x_N)$ with joint probability distribution

$$\frac{1}{N!} \det[K(x_i, x_j)]_{i,j=1}^{N} \prod_{n=1}^{N} \mu(\mathrm{d}x_n).$$

Then $\mathcal{X} \triangleq \{x_1, \dots, x_N\} \subset \mathbb{X}$ defines a projection DPP with kernel K.

Considering

- $\mathbb{X} = \{1, \dots, M\}$,
- $\mu = \sum_{m=1}^{M} \delta_m$,

one recovers the finite case with $\mathbf{K} = \mathbf{\Phi}^\top \mathbf{\Phi}$ and $\mathbf{\Phi}\mathbf{\Phi}^\top = I_N$.

Overview

Sequential sampling using the chain rule

The goal is to generate a random **subset** $\mathcal{X}$, such that

$$\mathbb{P}[\mathcal{X} = \{x_1, \dots, x_N\}] = \det \mathbf{K}_{\{x_1,\dots,x_N\}}.$$

Sequential sampling using the chain rule

The goal is to generate a random **subset** $\mathcal{X}$, such that

$$\mathbb{P}[\mathcal{X} = \{x_1, \dots, x_N\}] = \det \mathbf{K}_{\{x_1,\dots,x_N\}}.$$

(Hough et al., 2006; Gillenwater, 2014)

Consider the eigendecomposition $\mathbf{K} = \mathbf{\Phi}^\top \mathbf{\Phi}$, where $\mathbf{\Phi}\mathbf{\Phi}^\top = I_N$.

- Sample $(x_1, \dots, x_N)$ in the following way

Sequential sampling using the chain rule

The goal is to generate a random **subset** $\mathcal{X}$, such that

$$\mathbb{P}[\mathcal{X} = \{x_1, \dots, x_N\}] = \det \mathbf{K}_{\{x_1,\dots,x_N\}}.$$

(Hough et al., 2006; Gillenwater, 2014)

Consider the eigendecomposition $\mathbf{K} = \mathbf{\Phi}^{\top}\mathbf{\Phi}$, where $\mathbf{\Phi}\mathbf{\Phi}^{\top} = I_N$.

- Sample $(x_1, \dots, x_N)$ in the following way

$$\mathbb{P}[x_1 = x] \quad = \frac{1}{N}\|\phi_x\|^2,$$

Sequential sampling using the chain rule

The goal is to generate a random **subset** $\mathcal{X}$, such that

$$\mathbb{P}[\mathcal{X} = \{x_1, \dots, x_N\}] = \det \mathbf{K}_{\{x_1,\dots,x_N\}}.$$

(Hough et al., 2006; Gillenwater, 2014)

Consider the eigendecomposition $\mathbf{K} = \mathbf{\Phi}^\top \mathbf{\Phi}$, where $\mathbf{\Phi}\mathbf{\Phi}^\top = I_N$.

- Sample $(x_1, \dots, x_N)$ in the following way

$$\begin{aligned} \mathbb{P}[x_1 = x] &= \tfrac{1}{N} \|\phi_x\|^2, \\ \mathbb{P}[x_2 = x | x_1] &= \tfrac{1}{N-1} \operatorname{distance}^2(\phi_x, \operatorname{span}\{\phi_{x_1}\}), \end{aligned}$$

Sequential sampling using the chain rule

The goal is to generate a random **subset** $\mathcal{X}$, such that

$$\mathbb{P}[\mathcal{X} = \{x_1, \dots, x_N\}] = \det \mathbf{K}_{\{x_1,\dots,x_N\}}.$$

(Hough et al., 2006; Gillenwater, 2014)

Consider the eigendecomposition $\mathbf{K} = \mathbf{\Phi}^{\mathsf{T}}\mathbf{\Phi}$, where $\mathbf{\Phi}\mathbf{\Phi}^{\mathsf{T}} = I_N$.

- Sample $(x_1, \dots, x_N)$ in the following way

$$\begin{aligned} \mathbb{P}[x_1 = x] &= \tfrac{1}{N}\|\phi_x\|^2, \\ \mathbb{P}[x_2 = x | x_1] &= \tfrac{1}{N-1}\,\mathrm{distance}^2(\phi_x, \mathrm{span}\{\phi_{x_1}\}), \\ &\vdots \end{aligned}$$

Sequential sampling using the chain rule

The goal is to generate a random **subset** $\mathcal{X}$, such that

$$\mathbb{P}[\mathcal{X} = \{x_1, \dots, x_N\}] = \det \mathbf{K}_{\{x_1,\dots,x_N\}}.$$

(Hough et al., 2006; Gillenwater, 2014)

Consider the eigendecomposition $\mathbf{K} = \mathbf{\Phi}^\top \mathbf{\Phi}$, where $\mathbf{\Phi}\mathbf{\Phi}^\top = I_N$.

- Sample $(x_1, \dots, x_N)$ in the following way

$$\begin{aligned}
\mathbb{P}[x_1 = x] &= \tfrac{1}{N} \|\phi_x\|^2, \\
\mathbb{P}[x_2 = x | x_1] &= \tfrac{1}{N-1} \operatorname{distance}^2(\phi_x, \operatorname{span}\{\phi_{x_1}\}), \\
&\vdots \\
\mathbb{P}[x_N = x | x_{1:N-1}] &= \operatorname{distance}^2\left(\phi_x, \operatorname{span}\left\{\phi_{x_1}, \dots, \phi_{x_{N-1}}\right\}\right).
\end{aligned}$$

Sequential sampling using the chain rule

The goal is to generate a random **subset** $\mathcal{X}$, such that

$$\mathbb{P}[\mathcal{X} = \{x_1, \ldots, x_N\}] = \det \mathbf{K}_{\{x_1,\ldots,x_N\}}.$$

(Hough et al., 2006; Gillenwater, 2014)

Consider the eigendecomposition $\mathbf{K} = \mathbf{\Phi}^\top \mathbf{\Phi}$, where $\mathbf{\Phi}\mathbf{\Phi}^\top = I_N$.

- Sample $(x_1, \ldots, x_N)$ in the following way

$$\begin{aligned}
\mathbb{P}[x_1 = x] &= \tfrac{1}{N} \|\phi_x\|^2, \\
\mathbb{P}[x_2 = x | x_1] &= \tfrac{1}{N-1} \operatorname{distance}^2(\phi_x, \operatorname{span}\{\phi_{x_1}\}), \\
&\vdots \\
\mathbb{P}[x_N = x | x_{1:N-1}] &= \operatorname{distance}^2\left(\phi_x, \operatorname{span}\left\{\phi_{x_1}, \ldots, \phi_{x_{N-1}}\right\}\right).
\end{aligned}$$

- The likelihood of $(x_1, \ldots, x_N)$ reads

$$\mathbb{P}[(x_1, \ldots, x_n)] = \frac{1}{N!} \operatorname{volume}^2\{\phi_{x_1}, \ldots, \phi_{x_N}\} = \frac{1}{N!} \det \mathbf{K}_{\{x_1,\ldots,x_N\}}.$$

Sequential sampling using the chain rule

The goal is to generate a random **subset** $\mathcal{X}$, such that

$$\mathbb{P}[\mathcal{X} = \{x_1, \dots, x_N\}] = \det \mathbf{K}_{\{x_1,\dots,x_N\}}.$$

(Hough et al., 2006; Gillenwater, 2014)

Consider the eigendecomposition $\mathbf{K} = \mathbf{\Phi}^\top \mathbf{\Phi}$, where $\mathbf{\Phi}\mathbf{\Phi}^\top = I_N$.

- Sample $(x_1, \dots, x_N)$ in the following way

$$\begin{aligned} \mathbb{P}[x_1 = x] &= \tfrac{1}{N} \|\phi_x\|^2, \\ \mathbb{P}[x_2 = x | x_1] &= \tfrac{1}{N-1} \operatorname{distance}^2(\phi_x, \operatorname{span}\{\phi_{x_1}\}), \\ &\vdots \\ \mathbb{P}[x_N = x | x_{1:N-1}] &= \operatorname{distance}^2\left(\phi_x, \operatorname{span}\left\{\phi_{x_1}, \dots, \phi_{x_{N-1}}\right\}\right). \end{aligned}$$

- The likelihood of $(x_1, \dots, x_N)$ reads

$$\mathbb{P}[(x_1, \dots, x_n)] = \frac{1}{N!} \operatorname{volume}^2\{\phi_{x_1}, \dots, \phi_{x_N}\} = \frac{1}{N!} \det \mathbf{K}_{\{x_1,\dots,x_N\}}.$$

- The procedure is akin to Gram-Schmidt orthogonalization $\mathcal{O}(MN^2)$.

Illustration of the chain rule ($M = 24, N = 2$)

Text to summarize using $N = 2$ sentences.

But a dream within a dream?
Is all that we see or seem
One from the pitiless wave?
O God! can I not save
Them with a tighter clasp?
O God! can I not grasp
While I weep--while I weep!
Through my fingers to the deep,
How few! yet how they creep
Grains of the golden sand-
And I hold within my hand
Of a surf-tormented shore,
I stand amid the roar
Is but a dream within a dream.
All that we see or seem
Is it therefore the less gone?
In a vision, or in none,
In a night, or in a day,
Yet if hope has flown away
That my days have been a dream;
You are not wrong, who deem
Thus much let me avow-
And, in parting from you now,
Take this kiss upon the brow!

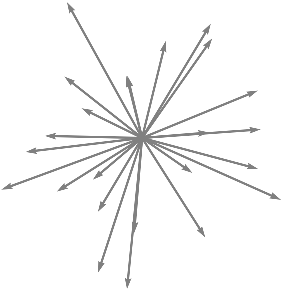
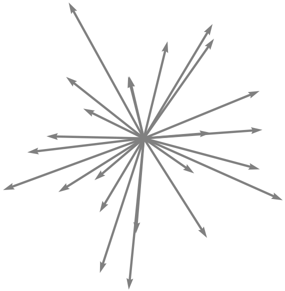


Illustration of the chain rule ($M = 24, N = 2$)

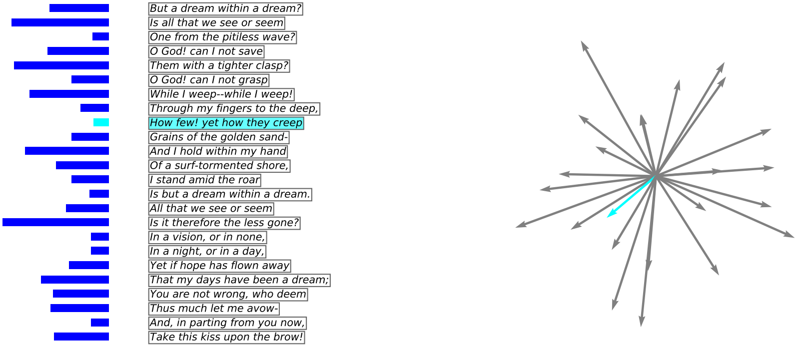
Select the first sentence,



$$\mathbb{P}[x_1 = x] = \frac{1}{2}\|\phi_x\|^2.$$

Illustration of the chain rule ($M = 24, N = 2$)

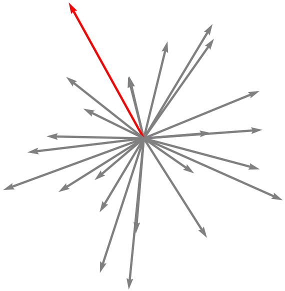
Select the first sentence,



$$\mathbb{P}[x_1 = x] = \frac{1}{2}\|\phi_x\|^2.$$

Illustration of the chain rule ($M = 24, N = 2$)

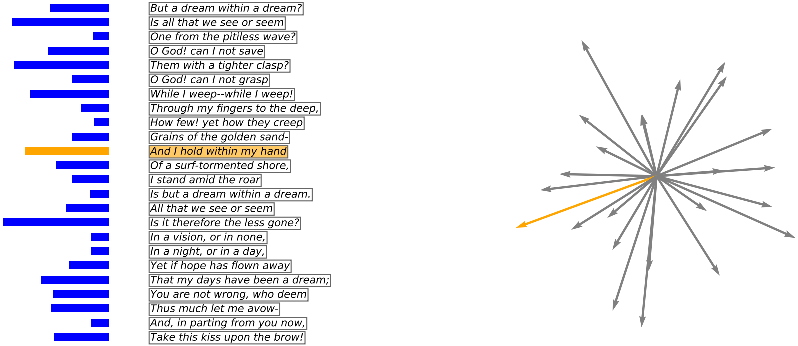
Select the first sentence,



$$\mathbb{P}[x_1 = x] = \frac{1}{2}\|\phi_x\|^2.$$

Illustration of the chain rule ($M = 24, N = 2$)

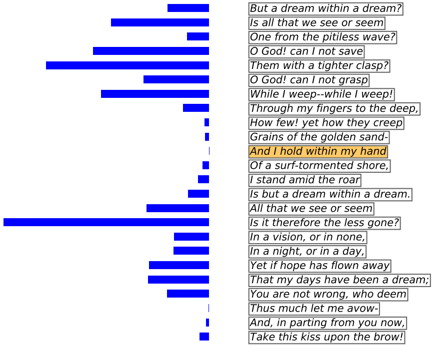
Select the first sentence,

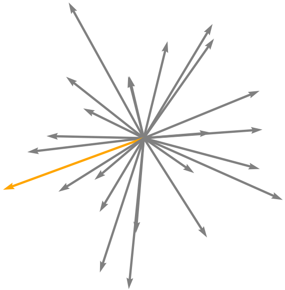


$$\mathbb{P}[x_1 = x] = \frac{1}{2}\|\phi_x\|^2.$$

Illustration of the chain rule ($M = 24, N = 2$)

Select the second sentence,

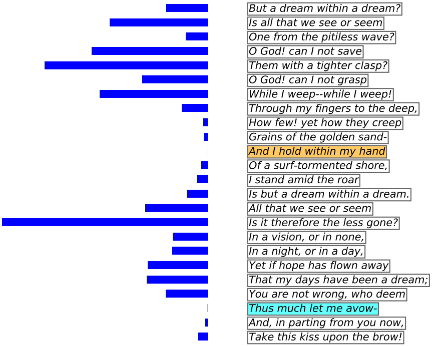


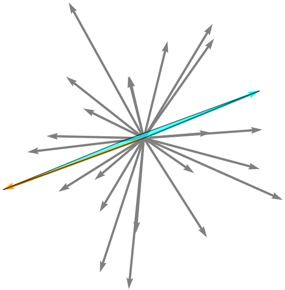


$$\mathbb{P}[x_2 = x \mid x_1] = \text{distance}^2(\phi_x, \text{span}\{\phi_{x_1}\}).$$

Illustration of the chain rule ($M = 24, N = 2$)

Select the second sentence,

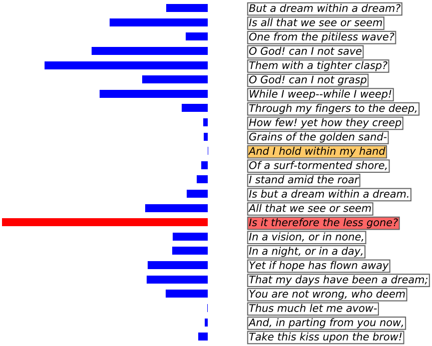


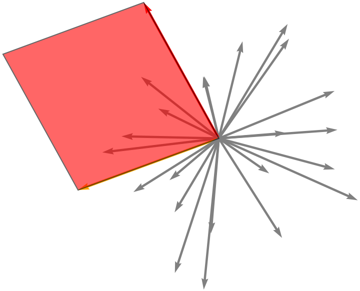


$$\mathbb{P}[x_2 = x \mid x_1] = \text{distance}^2(\phi_x, \text{span}\{\phi_{x_1}\}).$$

Illustration of the chain rule ($M = 24, N = 2$)

Select the second sentence,

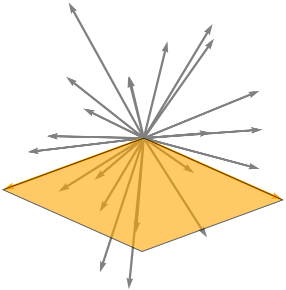




$$\mathbb{P}[x_2 = x \mid x_1] = \text{distance}^2(\phi_x, \text{span}\{\phi_{x_1}\}).$$

Illustration of the chain rule ($M = 24, N = 2$)

Select the second sentence,

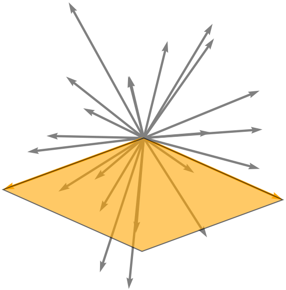


$$\mathbb{P}[x_2 = x \mid x_1] = \text{distance}^2(\phi_x, \text{span}\{\phi_{x_1}\}).$$

Illustration of the chain rule ($M = 24$, $N = 2$)

Output summary.

But a dream within a dream?
Is all that we see or seem
One from the pitiless wave?
O God! can I not save
Them with a tighter clasp?
O God! can I not grasp
While I weep--while I weep!
Through my fingers to the deep,
How few! yet how they creep
Grains of the golden sand-
And I hold within my hand
Of a surf-tormented shore,
I stand amid the roar
Is but a dream within a dream.
All that we see or seem
Is it therefore the less gone?
In a vision, or in none,
In a night, or in a day,
Yet if hope has flown away
That my days have been a dream;
You are not wrong, who deem
Thus much let me avow-
And, in parting from you now,
Take this kiss upon the brow!



$$\mathbb{P}[\mathcal{X} = \{x_1, x_2\}] = \text{volume}^2\{\phi_{x_1}, \phi_{x_2}\} = \det \mathbf{K}_{\{x_1, x_2\}}.$$

Overview

The basis-exchange walk

Metropolis Hastings kernel

(Feder and Mihail, 1992; Anari, Gharan, and Rezaei, 2016; Li, Jegelka, and Sra, 2016; Hermon and Salez, 2019)

- Starting from $B_0 \in \mathcal{B}$.

The basis-exchange walk

Metropolis Hastings kernel

(Feder and Mihail, 1992; Anari, Gharan, and Rezaei, 2016; Li, Jegelka, and Sra, 2016; Hermon and Salez, 2019)

- Starting from $B_0 \in \mathcal{B}$.
- Transitions

$$B \to \widetilde{B} = (B \setminus \{s\}) \cup \{t\},$$

where $s \sim \mathrm{Uniform}(B)$ and $t \sim \mathrm{Uniform}(B^c)$.

The basis-exchange walk

Metropolis Hastings kernel

(Feder and Mihail, 1992; Anari, Gharan, and Rezaei, 2016; Li, Jegelka, and Sra, 2016; Hermon and Salez, 2019)

- Starting from $B_0 \in \mathcal{B}$.
- Transitions

$$B \to \widetilde{B} = (B \setminus \{s\}) \cup \{t\},$$

where $s \sim \mathrm{Uniform}(B)$ and $t \sim \mathrm{Uniform}(B^c)$.

- Acceptance probability (lazy)

$$\frac{1}{2} \min\left(1, \frac{\det \mathbf{K}_{\widetilde{B}}}{\det \mathbf{K}_{B}}\right) = \frac{1}{2} \min\left(1, \frac{(\det \mathbf{\Phi}_{:\widetilde{B}})^2}{(\det \mathbf{\Phi}_{:B})^2}\right).$$

The basis-exchange walk

Metropolis Hastings kernel

(Feder and Mihail, 1992; Anari, Gharan, and Rezaei, 2016; Li, Jegelka, and Sra, 2016; Hermon and Salez, 2019)

- Starting from $B_0 \in \mathcal{B}$.
- Transitions

$$B \to \widetilde{B} = (B \setminus \{s\}) \cup \{t\},$$

  where $s \sim \text{Uniform}(B)$ and $t \sim \text{Uniform}(B^c)$.
- Acceptance probability (lazy)

$$\frac{1}{2} \min\left(1, \frac{\det \mathbf{K}_{\widetilde{B}}}{\det \mathbf{K}_{B}}\right) = \frac{1}{2} \min\left(1, \frac{(\det \mathbf{\Phi}_{:\widetilde{B}})^2}{(\det \mathbf{\Phi}_{:B})^2}\right).$$

- Mixing time

$$\mathcal{O}\left(MN \log\left(\log\left(\frac{1}{\det \mathbf{K}_{B_0}}\right)\right)\right).$$

The basis-exchange walk

Metropolis Hastings kernel

(Feder and Mihail, 1992; Anari, Gharan, and Rezaei, 2016; Li, Jegelka, and Sra, 2016; Hermon and Salez, 2019)

- Starting from $B_0 \in \mathcal{B}$.
- Transitions

$$B \to \widetilde{B} = (B \setminus \{s\}) \cup \{t\},$$

  where $s \sim \text{Uniform}(B)$ and $t \sim \text{Uniform}(B^c)$.
- Acceptance probability (lazy)

$$\frac{1}{2} \min\left(1, \frac{\det \mathbf{K}_{\widetilde{B}}}{\det \mathbf{K}_B}\right) = \frac{1}{2} \min\left(1, \frac{(\det \mathbf{\Phi}_{:\widetilde{B}})^2}{(\det \mathbf{\Phi}_{:B})^2}\right).$$

- Mixing time

$$\mathcal{O}\left(MN \log\left(\log\left(\frac{1}{\det \mathbf{K}_{B_0}}\right)\right)\right).$$

- (naive) Transition cost $\mathcal{O}(N^3)$.

Illustration of the basis-exchange walk ($M = 24, N = 7$)

$B_0 = \{1, 3, 9, 12, 13, 18, 24\}$

But a dream within a dream?
Is all that we see or seem
One from the pitiless wave?
O God! can I not save
Them with a tighter clasp?
O God! can I not grasp
While I weep--while I weep!
Through my fingers to the deep,
How few! yet how they creep
Grains of the golden sand-
And I hold within my hand
Of a surf-tormented shore,
I stand amid the roar
Is but a dream within a dream.
All that we see or seem
Is it therefore the less gone?
In a vision, or in none,
In a night, or in a day,
Yet if hope has flown away
That my days have been a dream;
You are not wrong, who deem
Thus much let me avow-
And, in parting from you now,
Take this kiss upon the brow!

Illustration of the basis-exchange walk ($M = 24, N = 7$)

$B_0 = \{1, 3, 9, 12, 13, 18, 24\}$ $B_1 = (B_0 \setminus \{24\}) \cup \{10\}$

But a dream within a dream?
Is all that we see or seem
One from the pitiless wave?
O God! can I not save
Them with a tighter clasp?
O God! can I not grasp
While I weep--while I weep!
Through my fingers to the deep,
How few! yet how they creep
Grains of the golden sand-
And I hold within my hand
Of a surf-tormented shore,
I stand amid the roar
Is but a dream within a dream.
All that we see or seem
Is it therefore the less gone?
In a vision, or in none,
In a night, or in a day,
Yet if hope has flown away
That my days have been a dream;
You are not wrong, who deem
Thus much let me avow-
And, in parting from you now,
Take this kiss upon the brow!

But a dream within a dream?
Is all that we see or seem
One from the pitiless wave?
O God! can I not save
Them with a tighter clasp?
O God! can I not grasp
While I weep--while I weep!
Through my fingers to the deep,
How few! yet how they creep
Grains of the golden sand-
And I hold within my hand
Of a surf-tormented shore,
I stand amid the roar
Is but a dream within a dream.
All that we see or seem
Is it therefore the less gone?
In a vision, or in none,
In a night, or in a day,
Yet if hope has flown away
That my days have been a dream;
You are not wrong, who deem
Thus much let me avow-
And, in parting from you now,
Take this kiss upon the brow!

Illustration of the basis-exchange walk ($M = 24, N = 7$)

$B_0 = \{1, 3, 9, 12, 13, 18, 24\}$

But a dream within a dream?
Is all that we see or seem
One from the pitiless wave?
O God! can I not save
Them with a tighter clasp?
O God! can I not grasp
While I weep--while I weep!
Through my fingers to the deep,
How few! yet how they creep
Grains of the golden sand-
And I hold within my hand
Of a surf-tormented shore,
I stand amid the roar
Is but a dream within a dream.
All that we see or seem
Is it therefore the less gone?
In a vision, or in none,
In a night, or in a day,
Yet if hope has flown away
That my days have been a dream;
You are not wrong, who deem
Thus much let me avow-
And, in parting from you now,
Take this kiss upon the brow!

$B_1 = (B_0 \setminus \{24\}) \cup \{10\}$

But a dream within a dream?
Is all that we see or seem
One from the pitiless wave?
O God! can I not save
Them with a tighter clasp?
O God! can I not grasp
While I weep--while I weep!
Through my fingers to the deep,
How few! yet how they creep
Grains of the golden sand-
And I hold within my hand
Of a surf-tormented shore,
I stand amid the roar
Is but a dream within a dream.
All that we see or seem
Is it therefore the less gone?
In a vision, or in none,
In a night, or in a day,
Yet if hope has flown away
That my days have been a dream;
You are not wrong, who deem
Thus much let me avow-
And, in parting from you now,
Take this kiss upon the brow!

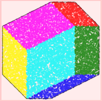
$B_2 = (B_1 \setminus \{10\}) \cup \{22\}$

But a dream within a dream?
Is all that we see or seem
One from the pitiless wave?
O God! can I not save
Them with a tighter clasp?
O God! can I not grasp
While I weep--while I weep!
Through my fingers to the deep,
How few! yet how they creep
Grains of the golden sand-
And I hold within my hand
Of a surf-tormented shore,
I stand amid the roar
Is but a dream within a dream.
All that we see or seem
Is it therefore the less gone?
In a vision, or in none,
In a night, or in a day,
Yet if hope has flown away
That my days have been a dream;
You are not wrong, who deem
Thus much let me avow-
And, in parting from you now,
Take this kiss upon the brow!

Overview

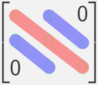
Conceptual shift: sampling by solving randomized linear programs

Finite projection DPP
Approximate sampling
Linear programming



ICML, 2017

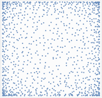
β-ensembles
Gibbs sampling
Random matrices

eigvals 

Submitted, 2020

DPP sampling

Monte Carlo integration
Exact sampling
Random linear system



NeurIPS, 2019

Python toolbox
DPPy
Reproducible research

```
from dppy import *
# [...]
dpp.sample()
```

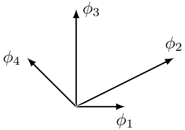
JMLR-MLOSS, 2019

Continuous embedding of the support

The support of finite projection DPPs, characterized by

$$\mathcal{B} \triangleq \{B\,;\, |B| = N, \text{ and } \det \mathbf{\Phi}_{:B} \neq 0\},$$

has the following geometrical representation.

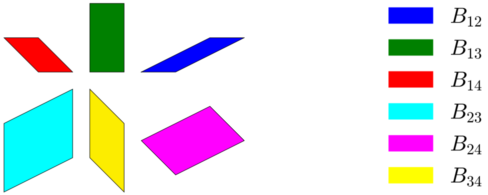


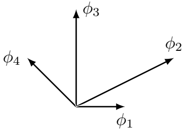
Continuous embedding of the support

The support of finite projection DPPs, characterized by

$$\mathcal{B} \triangleq \{B\,;\,|B| = N,\ \text{and}\ \det \boldsymbol{\Phi}_{:B} \neq 0\},$$

has the following geometrical representation.



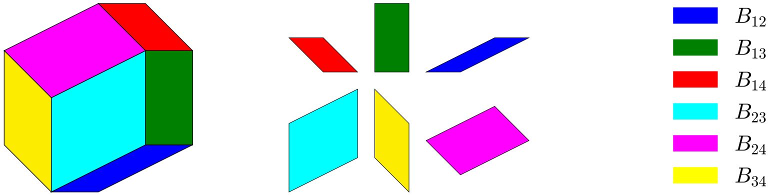


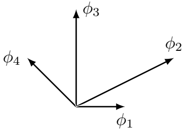
Continuous embedding of the support

The support of finite projection DPPs, characterized by

$$\mathcal{B} \triangleq \{B\,;\, |B| = N, \text{ and } \det \mathbf{\Phi}_{:B} \neq 0\},$$

has the following geometrical representation.



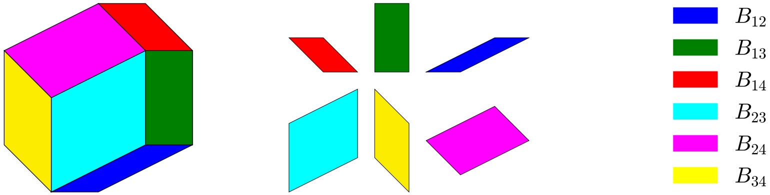


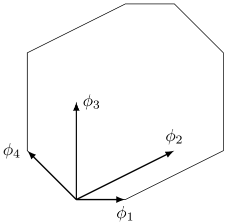
Continuous embedding of the support

The support of finite projection DPPs, characterized by

$$\mathcal{B} \triangleq \{B\,;\,|B| = N,\ \text{and}\ \det \mathbf{\Phi}_{:B} \neq 0\},$$

has the following geometrical representation.





How to identify the tiles?

Tiling of a zonotope (Dyer and Frieze, 1994)

Definition (Zonotope)

Let $\mathbf{\Phi} \in \mathbb{R}^{N\times M}$ such that $\underline{\operatorname{rank} \mathbf{\Phi} = N}$

$$\mathcal{Z}(\mathbf{\Phi}) \triangleq \mathbf{\Phi}[0,1]^M.$$

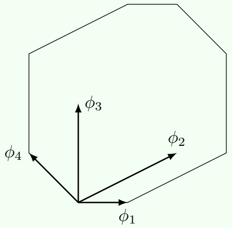
Tiling of a zonotope (Dyer and Frieze, 1994)

Definition (Zonotope)

Let $\mathbf{\Phi} \in \mathbb{R}^{N\times M}$ such that rank $\mathbf{\Phi} = N$

$$\mathcal{Z}(\mathbf{\Phi}) \triangleq \mathbf{\Phi}[0,1]^M.$$

Example ($M = 4, N = 2$)



Tiling of a zonotope (Dyer and Frieze, 1994)

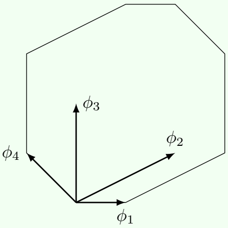
Definition (Zonotope)

Let $\mathbf{\Phi} \in \mathbb{R}^{N\times M}$ such that $\underline{\operatorname{rank}\mathbf{\Phi} = N}$

$$\mathcal{Z}(\mathbf{\Phi}) \triangleq \mathbf{\Phi}[0,1]^M.$$

- Let $x \in \mathcal{Z}(\mathbf{\Phi})$.

Example ($M = 4, N = 2$)



Tiling of a zonotope (Dyer and Frieze, 1994)

Definition (Zonotope)

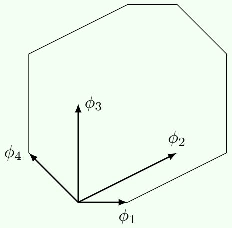
Let $\mathbf{\Phi} \in \mathbb{R}^{N\times M}$ such that $\underline{\operatorname{rank} \mathbf{\Phi} = N}$

$$\mathcal{Z}(\mathbf{\Phi}) \triangleq \mathbf{\Phi}[0,1]^M.$$

- Let $x \in \mathcal{Z}(\mathbf{\Phi})$.
- Solve the linear program (LP)

$$\begin{array}{ll} \min_{y\in\mathbb{R}^M} & c^{\top} y \\ \text{s.t.} & \mathbf{\Phi} y = x \\ & 0 \leq y \leq 1 \end{array}$$

Example ($M = 4, N = 2$)



Tiling of a zonotope (Dyer and Frieze, 1994)

Definition (Zonotope)

Let $\mathbf{\Phi} \in \mathbb{R}^{N \times M}$ such that $\underline{\operatorname{rank} \mathbf{\Phi} = N}$

$$\mathcal{Z}(\mathbf{\Phi}) \triangleq \mathbf{\Phi}[0,1]^M.$$

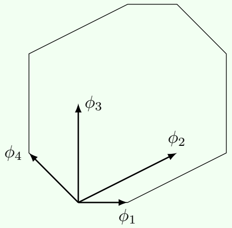
- Let $x \in \mathcal{Z}(\mathbf{\Phi})$.
- Solve the linear program (LP)

$$\begin{array}{ll} \min_{y \in \mathbb{R}^M} & c^\top y \\ \text{s.t.} & \mathbf{\Phi} y = x \\ & 0 \leq y \leq 1 \end{array}$$

- Consider the optimal solution y^* and keep only

$$B_x = \{i\,;\, 0 < y_i^* < 1\} \in \mathcal{B}.$$

Example ($M = 4, N = 2$)



Tiling of a zonotope (Dyer and Frieze, 1994)

Definition (Zonotope)

Let $\mathbf{\Phi} \in \mathbb{R}^{N\times M}$ such that $\underline{\operatorname{rank} \mathbf{\Phi} = N}$

$$\mathcal{Z}(\mathbf{\Phi}) \triangleq \mathbf{\Phi}[0,1]^M.$$

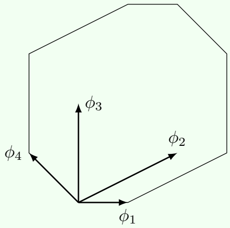
- Let $x \in \mathcal{Z}(\mathbf{\Phi})$.
- Solve the linear program (LP)

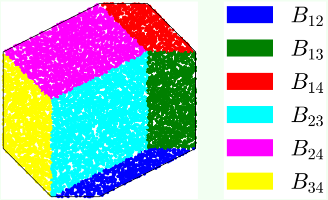
$$\begin{array}{ll} \min_{y\in\mathbb{R}^M} & c^{\top} y \\ \text{s.t.} & \mathbf{\Phi} y = x \\ & 0 \leq y \leq 1 \end{array}$$

- Consider the optimal solution y^* and keep only

$$B_x = \{i\,;\, 0 < y_i^* < 1\} \in \mathcal{B}.$$

Example ($M = 4, N = 2$)





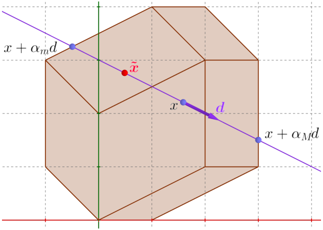
Random walk on zonotope $\overset{(LP)}{\Longrightarrow}$ random walk on tiles

Gautier, Bardenet, and Valko (2017)

Random walk on zonotope $\overset{(LP)}{\Longrightarrow}$ random walk on tiles

Gautier, Bardenet, and Valko (2017)

- Hit-and-run on $\mathcal{Z}(\mathbf{\Phi})$

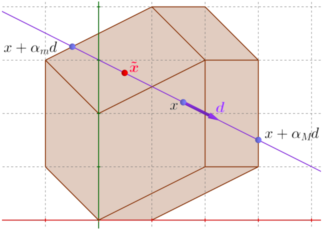


(Lovász and Vempala, 2003; Chen et al., 2018)

Random walk on zonotope $\overset{(LP)}{\Longrightarrow}$ random walk on tiles

Gautier, Bardenet, and Valko (2017)

- Hit-and-run on $\mathcal{Z}(\mathbf{\Phi})$



(Lovász and Vempala, 2003; Chen et al., 2018)

- Random walk on $\mathcal{B}$
  - Solve

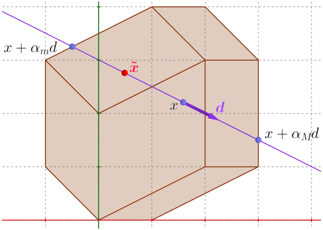
$$\begin{aligned} \min_{y \in \mathbb{R}^M} \quad & c^\top y \\ \text{s.t.} \quad & \mathbf{\Phi} y = x_t \\ & 0 \leq y \leq 1 \end{aligned}$$

  - $B_{x_t} = \{i \,;\, y_i^* \in]0, 1[\}$
  - Markov Chain $(B_{x_t})_{t \in \mathbb{N}}$

Random walk on zonotope $\xrightarrow{(LP)}$ random walk on tiles

Gautier, Bardenet, and Valko (2017)

- Hit-and-run on $\mathcal{Z}(\mathbf{\Phi})$



(Lovász and Vempala, 2003; Chen et al., 2018)

- Target density on $\mathcal{Z}(\mathbf{\Phi})$

$$\pi(x) = \sum_{B \in \mathcal{B}} C_B \times \mathbb{1}_{\mathcal{Z}(\mathbf{\Phi}_{:B})}(x).$$

- Random walk on $\mathcal{B}$
  - Solve

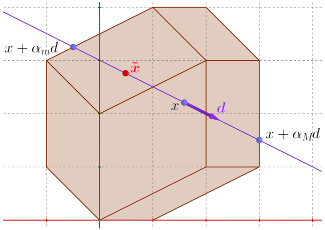
$$\begin{aligned} \min_{y \in \mathbb{R}^M} \quad & c^\top y \\ \text{s.t.} \quad & \mathbf{\Phi} y = x_t \\ & 0 \leq y \leq 1 \end{aligned}$$

  - $B_{x_t} = \{i \,;\, y_i^* \in]0, 1[\}$
  - Markov Chain $(B_{x_t})_{t \in \mathbb{N}}$

Random walk on zonotope $\overset{(LP)}{\Longrightarrow}$ random walk on tiles

Gautier, Bardenet, and Valko (2017)

- Hit-and-run on $\mathcal{Z}(\mathbf{\Phi})$



(Lovász and Vempala, 2003; Chen et al., 2018)

- Target density on $\mathcal{Z}(\mathbf{\Phi})$

$$\pi(x) = \sum_{B \in \mathcal{B}} C_B \times \mathbb{1}_{\mathcal{Z}(\mathbf{\Phi}_{:B})}(x).$$

- Random walk on $\mathcal{B}$
  - Solve

$$\begin{array}{ll} \min_{y \in \mathbb{R}^M} & c^\top y \\ \text{s.t.} & \mathbf{\Phi} y = x_t \\ & 0 \leq y \leq 1 \end{array}$$

  - $B_{x_t} = \{i \,;\, y_i^* \in]0, 1[\}$
  - Markov Chain $(B_{x_t})_{t \in \mathbb{N}}$

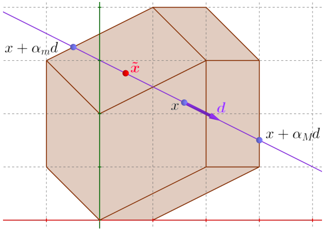
- Limiting distribution on $\mathcal{B}$

$$\mathbb{P}[B_x = B] = C_B \times \left|\det \mathbf{\Phi}_{:B}\right|.$$

Random walk on zonotope $\overset{(LP)}{\Longrightarrow}$ random walk on tiles

Gautier, Bardenet, and Valko (2017)

- Hit-and-run on $\mathcal{Z}(\mathbf{\Phi})$



(Lovász and Vempala, 2003; Chen et al., 2018)

- Target density on $\mathcal{Z}(\mathbf{\Phi})$

$$\pi(x) = \sum_{B \in \mathcal{B}} C_B \times \mathbb{1}_{\mathcal{Z}(\mathbf{\Phi}_{:B})}(x).$$

- Random walk on $\mathcal{B}$
  - Solve

$$\begin{array}{ll} \min_{y \in \mathbb{R}^M} & c^\top y \\ \text{s.t.} & \mathbf{\Phi} y = x_t \\ & 0 \leq y \leq 1 \end{array}$$

  - $B_{x_t} = \{i \,;\, y_i^* \in]0,1[\}$
  - Markov Chain $(B_{x_t})_{t \in \mathbb{N}}$

- Limiting distribution on $\mathcal{B}$

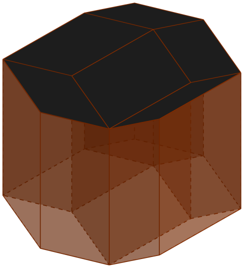
$$\mathbb{P}[B_x = B] = C_B \times |\det \mathbf{\Phi}_{:B}|.$$

How to make $\mathbb{P}[B_x = B] \propto (\det \mathbf{\Phi}_{:B})^2$?

Hit-and-run with acceptance ratio = 1

The target density on $\mathcal{Z}(\mathbf{\Phi})$ is uniform,

$$\pi(x) \propto \sum_{B\in\mathcal{B}} 1 \times \mathbb{1}_{\mathcal{Z}(\mathbf{\Phi}_{:\mathbf{B}})}(x).$$



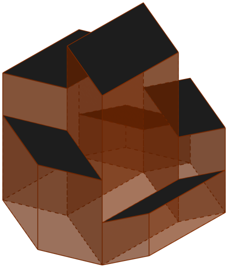
The limiting distribution on $\mathcal{B}$ takes the form

$$\mathbb{P}[B_x = B] \propto 1 \times \left|\det \mathbf{\Phi}_{:B}\right| = \left|\det \mathbf{\Phi}_{:B}\right|^{1}.$$

Hit-and-run with acceptance ratio $= \left|\frac{\det \mathbf{\Phi}_{:\widetilde{B}}}{\det \mathbf{\Phi}_{:B}}\right|$

The target density on $\mathcal{Z}(\mathbf{\Phi})$ is given by

$$\pi(x) \propto \sum_{B\in\mathcal{B}} |\det \mathbf{\Phi}_{:B}| \times \mathbb{1}_{\mathcal{Z}(\mathbf{\Phi}_{:\mathbf{B}})}(x).$$



The limiting distribution on $\mathcal{B}$ takes the form

$$\mathbb{P}[B_x = B] \propto |\det \mathbf{\Phi}_{:B}| \times |\det \mathbf{\Phi}_{:B}| = (\det \mathbf{\Phi}_{:B})^2.$$

Illustration of the zonotope walk ($M = 24, N = 7$)

$B_0 = \{1, 3, 9, 12, 13, 18, 24\}$

But a dream within a dream?
Is all that we see or seem
One from the pitiless wave?
O God! can I not save
Them with a tighter clasp?
O God! can I not grasp
While I weep--while I weep!
Through my fingers to the deep,
How few! yet how they creep
Grains of the golden sand-
And I hold within my hand
Of a surf-tormented shore,
I stand amid the roar
Is but a dream within a dream.
All that we see or seem
Is it therefore the less gone?
In a vision, or in none,
In a night, or in a day,
Yet if hope has flown away
That my days have been a dream;
You are not wrong, who deem
Thus much let me avow-
And, in parting from you now,
Take this kiss upon the brow!

Illustration of the zonotope walk ($M = 24, N = 7$)

$B_0 = \{1, 3, 9, 12, 13, 18, 24\}$ $B_1 = (B_0 \setminus \{1, 9, 13, 18\}) \cup \{6, 8, 14, 17\}$

But a dream within a dream?	*But a dream within a dream?*
Is all that we see or seem	*Is all that we see or seem*
One from the pitiless wave?	***One from the pitiless wave?***
O God! can I not save	*O God! can I not save*
Them with a tighter clasp?	*Them with a tighter clasp?*
O God! can I not grasp	***O God! can I not grasp***
While I weep--while I weep!	*While I weep--while I weep!*
Through my fingers to the deep,	***Through my fingers to the deep,***
How few! yet how they creep	*How few! yet how they creep*
Grains of the golden sand-	*Grains of the golden sand-*
And I hold within my hand	*And I hold within my hand*
Of a surf-tormented shore,	***Of a surf-tormented shore,***
I stand amid the roar	*I stand amid the roar*
Is but a dream within a dream.	***Is but a dream within a dream.***
All that we see or seem	*All that we see or seem*
Is it therefore the less gone?	*Is it therefore the less gone?*
In a vision, or in none,	***In a vision, or in none,***
In a night, or in a day,	*In a night, or in a day,*
Yet if hope has flown away	*Yet if hope has flown away*
That my days have been a dream;	*That my days have been a dream;*
You are not wrong, who deem	*You are not wrong, who deem*
Thus much let me avow-	*Thus much let me avow-*
And, in parting from you now,	*And, in parting from you now,*
Take this kiss upon the brow!	***Take this kiss upon the brow!***

Illustration of the zonotope walk ($M = 24, N = 7$)

$B_0 = \{1, 3, 9, 12, 13, 18, 24\}$

But a dream within a dream?
Is all that we see or seem
One from the pitiless wave?
O God! can I not save
Them with a tighter clasp?
O God! can I not grasp
While I weep--while I weep!
Through my fingers to the deep,
How few! yet how they creep
Grains of the golden sand-
And I hold within my hand
Of a surf-tormented shore,
I stand amid the roar
Is but a dream within a dream.
All that we see or seem
Is it therefore the less gone?
In a vision, or in none,
In a night, or in a day,
Yet if hope has flown away
That my days have been a dream;
You are not wrong, who deem
Thus much let me avow-
And, in parting from you now,
Take this kiss upon the brow!

$B_1 = (B_0 \setminus \{1, 9, 13, 18\}) \cup \{6, 8, 14, 17\}$

But a dream within a dream?
Is all that we see or seem
One from the pitiless wave?
O God! can I not save
Them with a tighter clasp?
O God! can I not grasp
While I weep--while I weep!
Through my fingers to the deep,
How few! yet how they creep
Grains of the golden sand-
And I hold within my hand
Of a surf-tormented shore,
I stand amid the roar
Is but a dream within a dream.
All that we see or seem
Is it therefore the less gone?
In a vision, or in none,
In a night, or in a day,
Yet if hope has flown away
That my days have been a dream;
You are not wrong, who deem
Thus much let me avow-
And, in parting from you now,
Take this kiss upon the brow!

$B_2 = (B_1 \setminus \{3, 8, 12, 17\}) \cup \{7, 10, 15, 23\}$

But a dream within a dream?
Is all that we see or seem
One from the pitiless wave?
O God! can I not save
Them with a tighter clasp?
O God! can I not grasp
While I weep--while I weep!
Through my fingers to the deep,
How few! yet how they creep
Grains of the golden sand-
And I hold within my hand
Of a surf-tormented shore,
I stand amid the roar
Is but a dream within a dream.
All that we see or seem
Is it therefore the less gone?
In a vision, or in none,
In a night, or in a day,
Yet if hope has flown away
That my days have been a dream;
You are not wrong, who deem
Thus much let me avow-
And, in parting from you now,
Take this kiss upon the brow!

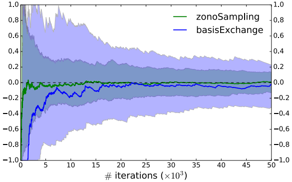
Comparison of the zonotope and basis-exchange walks

Relative error of the estimation of $\mathbb{P}[\{x_1, x_2, x_3\} \subset \mathcal{X}] = \det \mathbf{K}_{\{x_1,x_2,x_3\}}$.

Comparison of the zonotope and basis-exchange walks

Relative error of the estimation of $\mathbb{P}[\{x_1, x_2, x_3\} \subset \mathcal{X}] = \det \mathbf{K}_{\{x_1, x_2, x_3\}}$.

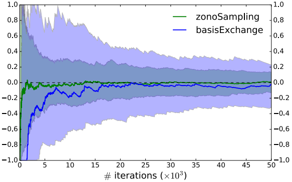
Rel. error / iteration



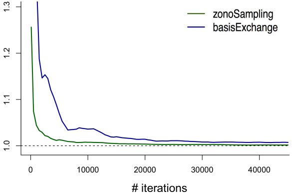
Comparison of the zonotope and basis-exchange walks

Relative error of the estimation of $\mathbb{P}[\{x_1, x_2, x_3\} \subset \mathcal{X}] = \det \mathbf{K}_{\{x_1,x_2,x_3\}}$.

Rel. error / iteration



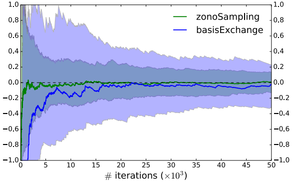
PSRF / iteration



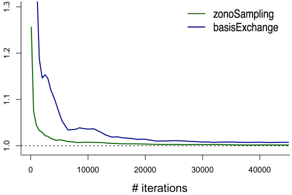
Comparison of the zonotope and basis-exchange walks

Relative error of the estimation of $\mathbb{P}[\{x_1, x_2, x_3\} \subset \mathcal{X}] = \det \mathbf{K}_{\{x_1,x_2,x_3\}}$.

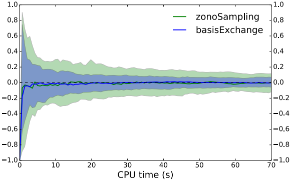
Rel. error / iteration



PSRF / iteration



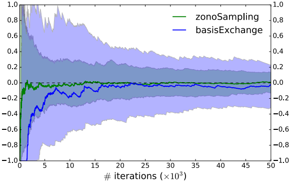
Rel. error / CPU time



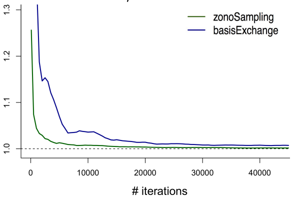
Comparison of the zonotope and basis-exchange walks

Relative error of the estimation of $\mathbb{P}[\{x_1, x_2, x_3\} \subset \mathcal{X}] = \det \mathbf{K}_{\{x_1,x_2,x_3\}}$.

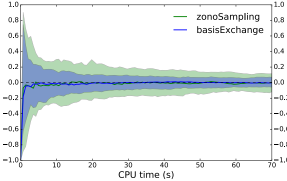
Rel. error / iteration



PSRF / iteration



Rel. error / CPU time

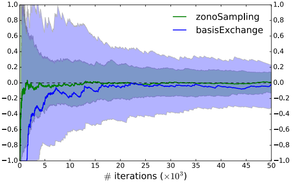


	basis-exchange		zonotope	
Exploration of the support	✓	(lazy)	✓✓✓	
Empirical mixing	✓✓		✓✓✓	
Cost per iteration	✓✓✓	det	✓	det + 3 LPs
Theoretical guarantees	✓✓	poly(M, N)	**?**	poly(M, N)**?**

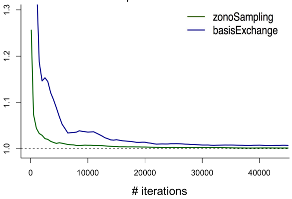
Comparison of the zonotope and basis-exchange walks

Relative error of the estimation of $\mathbb{P}[\{x_1, x_2, x_3\} \subset \mathcal{X}] = \det \mathbf{K}_{\{x_1,x_2,x_3\}}$.

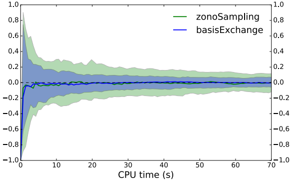
Rel. error / iteration



PSRF / iteration



Rel. error / CPU time



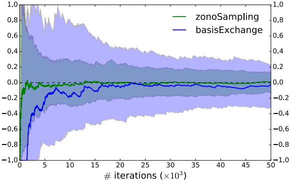
	basis-exchange		zonotope	
Exploration of the support	✓	(lazy)	✓✓✓	
Empirical mixing	✓✓		✓✓✓	
Cost per iteration	✓✓✓	det	✓	det + 3 LPs
Theoretical guarantees	✓✓	poly(M, N)	**?**	poly(M, N)**?**

The zonotope walk is sample efficient.

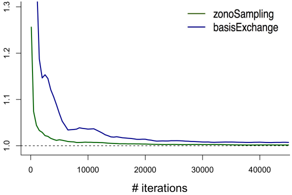
Comparison of the zonotope and basis-exchange walks

Relative error of the estimation of $\mathbb{P}[\{x_1, x_2, x_3\} \subset \mathcal{X}] = \det \mathbf{K}_{\{x_1,x_2,x_3\}}$.

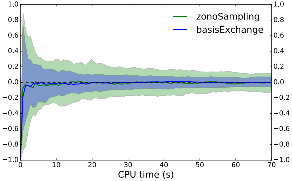
Rel. error / iteration



PSRF / iteration



Rel. error / CPU time



	basis-exchange		zonotope	
Exploration of the support	✓	(lazy)	✓✓✓	
Empirical mixing	✓✓		✓✓✓	
Cost per iteration	✓✓✓	det	✓	det + 3 LPs
Theoretical guarantees	✓✓	poly(M, N)	**?**	poly(M, N)**?**

The zonotope walk is sample efficient.

Can we generalize the idea to the continuous setting?

Overview

Sampling by solving randomized linear programs?

Finite case

Linear Programming (LP)

$$\begin{array}{lrcl} \min_y & c^\top y & & \\ \text{s.t.} & {\varphi_1}^\top y & = & x_1 \\ & & \vdots & \\ & {\varphi_N}^\top y & = & x_N \\ & & 0 \leq y \leq 1 & \end{array}$$

- Unique solution y^* ☺
- Efficient solvers ☺
- "Support" of the solution
  - $|i\ ;\ 0 < y_i^* < 1| = N$ ☺

Sampling by solving randomized linear programs?

Finite case
Linear Programming (LP)

$$\begin{aligned} \min_y \quad & c^\top y \\ \text{s.t.} \quad & {\varphi_1}^\top y = x_1 \\ & \vdots \\ & {\varphi_N}^\top y = x_N \\ & 0 \le y \le 1 \end{aligned}$$

- Unique solution y^* ☺
- Efficient solvers ☺
- "Support" of the solution
  - $|i\ ;\ 0 < y_i^* < 1| = N$ ☺

Continuous case (dimension d)
Linear Semi Infinite Programming (LSIP)

$$\begin{aligned} \min_\nu \quad & \int c(x)\nu(\mathrm{d}x) \\ \text{s.t.} \quad & \int \varphi_1(x)\nu(\mathrm{d}x) = m_1 \\ & \vdots \\ & \int \varphi_N(x)\nu(\mathrm{d}x) = m_N \\ & \text{``}0 \le \mu \le 1\text{''} \end{aligned}$$

- No unique solution ☹
- No efficient solvers ☹
- Structure of the support of solutions
  - $\exists \nu^*$ s.t. $|\operatorname{supp} \nu^*| \le N$.

(Goberna and López, 2014)

Sampling by solving randomized linear programs?

Dimension $d > 1$

For polynomials functions c and φ_n

$$
\begin{aligned}
\min_\nu \quad & \int c(x)\nu(\mathrm{d}x) \\
\text{s.t.} \quad & \int \varphi_1(x)\nu(\mathrm{d}x) = m_1 \\
& \vdots \\
& \int \varphi_N(x)\nu(\mathrm{d}x) = m_N
\end{aligned}
$$

- No unique solution ☹
- Efficient solvers**?** (Lasserre, 2010)
  - hierarchy of SDP relaxations ☺
  - works for small d and N ☹
- Structure of the support of solutions
  - $\exists \nu^*$ s.t. $|\operatorname{supp} \nu^*| \leq N$
  - unstable support extraction ☹

Sampling by solving randomized linear programs?

Dimension $d > 1$

For polynomials functions c and φ_n

$$\begin{array}{llll} \min_\nu & \int c(x)\nu(\mathrm{d}x) & & \\ \text{s.t.} & \int \varphi_1(x)\nu(\mathrm{d}x) & = & m_1 \\ & & \vdots & \\ & \int \varphi_N(x)\nu(\mathrm{d}x) & = & m_N \end{array}$$

- No unique solution ☹
- Efficient solvers**?** (Lasserre, 2010)
  - hierarchy of SDP relaxations ☺
  - works for small d and N ☹
- Structure of the support of solutions
  - $\exists \nu^*$ s.t. $|\operatorname{supp} \nu^*| \leq N$
  - unstable support extraction ☹

Dimension $d = 1$

Truncated moment problem

$$\begin{array}{llll} \min_\nu & \mathbb{E}_\nu\left[X^{2N}\right] & & \\ \text{s.t.} & \mathbb{E}_\nu[X] & = & m_1 \\ & & \vdots & \\ & \mathbb{E}_\nu\left[X^{2N-1}\right] & = & m_{2N-1} \end{array}$$

- Unique solution $\nu^* = \sum_{n=1}^{N} \omega_n \delta_{x_n}$ ☺
- Unstable support extraction ☹

Sampling by solving randomized linear programs?

Finite case
Linear Programming (LP)

$$\begin{aligned} \min_y \quad & c^\top y \\ \text{s.t.} \quad & {\varphi_1}^\top y = x_1 \\ & \vdots \\ & {\varphi_N}^\top y = x_N \\ & 0 \leq y \leq 1 \end{aligned}$$

- Unique solution y^* ☺
- Efficient solvers ☺
- "Support" of the solution
  - $|i \; ; \; 0 < y_i^* < 1| = N$ ☺

Dimension $d = 1$
Truncated moment problem

$$\begin{aligned} \min_\nu \quad & \mathbb{E}_\nu\left[X^{2N}\right] \\ \text{s.t.} \quad & \mathbb{E}_\nu[X] = m_1 \\ & \vdots \\ & \mathbb{E}_\nu\left[X^{2N-1}\right] = m_{2N-1} \end{aligned}$$

- Unique solution $\nu^* = \sum_{n=1}^{N} \omega_n \delta_{x_n}$ ☺
- Unstable support extraction ☹

How to randomize the moment constraints s.t. $\{x_1, \dots, x_N\} \sim$ target DPP?

Sampling by solving randomized linear programs?

Finite case
Linear Programming (LP)

$$\begin{aligned} \min_y \quad & c^\top y \\ \text{s.t.} \quad & {\varphi_1}^\top y = x_1 \\ & \vdots \\ & {\varphi_N}^\top y = x_N \\ & 0 \leq y \leq 1 \end{aligned}$$

- Unique solution y^* ☺
- Efficient solvers ☺
- "Support" of the solution
  - $|i \; ; \; 0 < y_i^* < 1| = N$ ☺

Dimension $d = 1$
Truncated moment problem

$$\begin{aligned} \min_\nu \quad & \mathbb{E}_\nu \left[X^{2N}\right] \\ \text{s.t.} \quad & \mathbb{E}_\nu[X] = m_1 \\ & \vdots \\ & \mathbb{E}_\nu \left[X^{2N-1}\right] = m_{2N-1} \end{aligned}$$

- Unique solution $\nu^* = \sum_{n=1}^{N} \omega_n \delta_{x_n}$ ☺
- Unstable support extraction ☹

$(\omega_n), (x_n)$ define a quadrature rule
(**RyBo15**; Dette and Studden, 1997)

~~How to randomize the moment constraints s.t. $\{x_1, \dots, x_N\} \sim$ target DPP?~~

Sampling by computing the eigenvalues of random tridiagonal matrices

Dimension $d = 1$
Truncated moment problem

$$\begin{aligned} \min_\nu \quad & \mathbb{E}_\nu\left[X^{2N}\right] \\ \text{s.t.} \quad & \mathbb{E}_\nu[X] &=& \quad m_1 \\ & &\vdots& \\ & \mathbb{E}_\nu\left[X^{2N-1}\right] &=& \quad m_{2N-1} \end{aligned}$$

- Unique solution $\nu^* = \sum_{n=1}^{N} \omega_n \delta_{x_n}$
- $(\omega_n), (x_n)$ define a quadrature rule
  $\int p \, \mathrm{d}\mu = \sum \omega_n p(x_n), \ \deg p \leq 2N-1.$

Sampling by computing the eigenvalues of random tridiagonal matrices

Dimension $d = 1$
Truncated moment problem

$$\begin{array}{lll} \min_\nu & \mathbb{E}_\nu\left[X^{2N}\right] & \\ \text{s.t.} & \mathbb{E}_\nu[X] = & m_1 \\ & \vdots & \\ & \mathbb{E}_\nu\left[X^{2N-1}\right] = & m_{2N-1} \end{array}$$

- Unique solution $\nu^* = \sum_{n=1}^{N} \omega_n \delta_{x_n}$
- $(\omega_n), (x_n)$ define a quadrature rule
  $\int p \, \mathrm{d}\mu = \sum \omega_n p(x_n), \ \deg p \leq 2N-1.$

Reparametrize ν^* via the 3-terms recurrence relation $\perp$ polynomials encoded by

$$J_{\mathbf{a},\mathbf{b}} \triangleq \begin{bmatrix} a_1 & \sqrt{b_1} & & (0) \\ \sqrt{b_1} & a_2 & \ddots & \\ & \ddots & \ddots & \sqrt{b_{N-1}} \\ (0) & & \sqrt{b_{N-1}} & a_N \end{bmatrix}$$

- $\{x_1, \dots, x_N\} = \text{eigvals } J_{\mathbf{a},\mathbf{b}}$
- Computational cost $\mathcal{O}(N^2)$

Sampling by computing the eigenvalues of random tridiagonal matrices

Dimension $d = 1$
Truncated moment problem

$$
\begin{array}{llcl}
\min_\nu & \mathbb{E}_\nu\left[X^{2N}\right] & & \\
\text{s.t.} & \mathbb{E}_\nu[X] & = & m_1 \\
 & & \vdots & \\
 & \mathbb{E}_\nu\left[X^{2N-1}\right] & = & m_{2N-1}
\end{array}
$$

- Unique solution $\nu^* = \sum_{n=1}^{N} \omega_n \delta_{x_n}$
- $(\omega_n), (x_n)$ define a quadrature rule
  $\int p \, \mathrm{d}\mu = \sum \omega_n p(x_n), \ \deg p \leq 2N-1.$

Reparametrize ν^* via the 3-terms recurrence relation $\perp$ polynomials encoded by

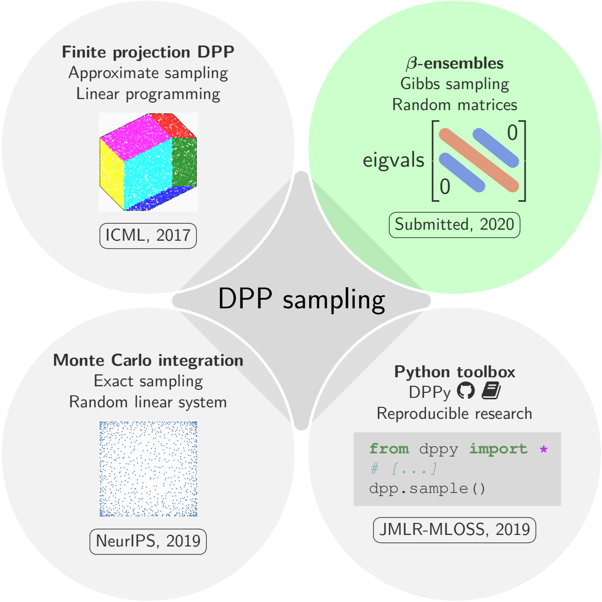
$$
J_{\mathbf{a},\mathbf{b}} \triangleq \begin{bmatrix}
a_1 & \sqrt{b_1} & & (0) \\
\sqrt{b_1} & a_2 & \ddots & \\
 & \ddots & \ddots & \sqrt{b_{N-1}} \\
(0) & & \sqrt{b_{N-1}} & a_N
\end{bmatrix}
$$

- $\{x_1, \dots, x_N\} = \text{eigvals } J_{\mathbf{a},\mathbf{b}}$
- Computational cost $\mathcal{O}(N^2)$

Sampling 1D continuous projection DPPs may be cheaper than sampling finite projection DPPs?!

Overview

Convert a random matrix analysis tool to a computational tool



Definition (β-ensemble)

Let $(x_1, \dots, x_N)$ with distribution proportional to

$$\left| \prod_{i<j} (x_j - x_i) \right|^\beta \prod_{n=1}^{N} e^{-V(x_n)} \mathrm{d}x_n,$$

then $\mathcal{X} = \{x_1, \dots, x_N\}$ is called a β-ensemble with potential V.

- Repulsion characterized by $\prod_{i<j}(x_j - x_i) = \det\left[x_j^{i-1}\right]_{i,j=1}^{N}$.
- Strength of the repulsion parametrized by $\beta > 0$ (inverse temperature).

Definition (β-ensemble)

Let $(x_1, \dots, x_N)$ with distribution proportional to

$$\left|\prod_{i<j}(x_j - x_i)\right|^\beta \prod_{n=1}^{N} e^{-V(x_n)} \mathrm{d}x_n,$$

then $\mathcal{X} = \{x_1, \dots, x_N\}$ is called a β-ensemble with potential V.

- Repulsion characterized by $\prod_{i<j}(x_j - x_i) = \det\left[x_j^{i-1}\right]_{i,j=1}^{N}$.
- Strength of the repulsion parametrized by $\beta > 0$ (inverse temperature).

Example ($\beta = 2$, corresponds to a projection DPP)

- $\mu(\mathrm{d}x) = e^{-V(x)}\mathrm{d}x$.
- $K(x, y) = \sum_{k=0}^{N-1} p_k(x)p_k(y), \quad p_k, p_\ell \perp$ polynomials w.r.t. μ.

Classical β-ensembles and random matrix models

Name	Potential $V(x)$	Support
Hermite	$\frac{1}{2\sigma^2}(x-\mu)^2$	$\mathbb{R}$
Laguerre	$-(k-1)\log(x)+\frac{1}{\theta}x$	$]0,\infty[$
Jacobi	$-(a-1)\log(x)-(b-1)\log(1-x)$	$]0,1[$

Classical β-ensembles and random matrix models

Name	Potential $V(x)$	Support
Hermite	$\frac{1}{2\sigma^2}(x-\mu)^2$	$\mathbb{R}$
Laguerre	$-(k-1)\log(x)+\frac{1}{\theta}x$	$]0,\infty[$
Jacobi	$-(a-1)\log(x)-(b-1)\log(1-x)$	$]0,1[$

$\beta(=1,2,4)$**-ensembles as the eigenvalue distribution of random matrices.**

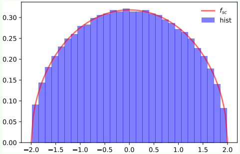
Classical β-ensembles and random matrix models

Name	Potential $V(x)$	Support
Hermite	$\frac{1}{2\sigma^2}(x-\mu)^2$	$\mathbb{R}$
Laguerre	$-(k-1)\log(x)+\frac{1}{\theta}x$	$]0,\infty[$
Jacobi	$-(a-1)\log(x)-(b-1)\log(1-x)$	$]0,1[$

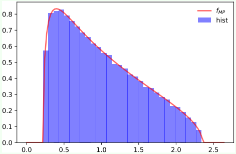
$\beta(=1,2,4)$**-ensembles as the eigenvalue distribution of random matrices.**

Example ($\beta=2$ and $X\sim$ standard complex Gaussian matrix)

- $X\in\mathbb{C}^{N\times N}$
- eigvals$(X+X^{\mathsf{H}})\sim$ Hermite



- $X\in\mathbb{C}^{N\times M}$
- eigvals$(XX^{\mathsf{H}})\sim$ Laguerre



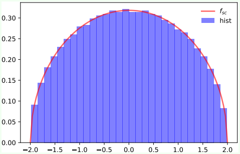
Classical β-ensembles and random matrix models

Name	Potential $V(x)$	Support
Hermite	$\frac{1}{2\sigma^2}(x-\mu)^2$	$\mathbb{R}$
Laguerre	$-(k-1)\log(x) + \frac{1}{\theta}x$	$]0,\infty[$
Jacobi	$-(a-1)\log(x) - (b-1)\log(1-x)$	$]0,1[$

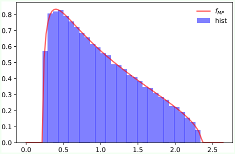
$\beta(=1,2,4)$**-ensembles as the eigenvalue distribution of random matrices.**

Example ($\beta = 2$ and $X \sim$ standard complex Gaussian matrix)

- $X \in \mathbb{C}^{N\times N}$
- eigvals$(X + X^{\mathsf{H}}) \sim$ Hermite



- $X \in \mathbb{C}^{N\times M}$
- eigvals$(XX^{\mathsf{H}}) \sim$ Laguerre



Random matrix models grant $\mathcal{O}(N^3)$ exact samplers!

Classical β-ensembles and random tridiagonal models

Dumitriu and Edelman (2002) and Killip and Nenciu (2004) derived equivalent random tridiagonal models, valid for $\beta > 0$,

$$\operatorname{eigvals}\begin{bmatrix} a_1 & \sqrt{b_1} & & (0) \\ \sqrt{b_1} & a_2 & \ddots & \\ & \ddots & \ddots & \sqrt{b_{N-1}} \\ (0) & & \sqrt{b_{N-1}} & a_N \end{bmatrix}.$$

Classical β-ensembles and random tridiagonal models

Dumitriu and Edelman (2002) and Killip and Nenciu (2004) derived equivalent random tridiagonal models, valid for $\beta > 0$,

$$\operatorname{eigvals} \begin{bmatrix} a_1 & \sqrt{b_1} & & (0) \\ \sqrt{b_1} & a_2 & \ddots & \\ & \ddots & \ddots & \sqrt{b_{N-1}} \\ (0) & & \sqrt{b_{N-1}} & a_N \end{bmatrix}.$$

Example (Hermite ensemble, $\beta > 0$, $V(x) = \frac{1}{2\sigma^2}(x-\mu)^2$)

Consider independent $a_n \sim \mathcal{N}\left(\mu, \sigma^2\right)$, and $b_n \sim \Gamma\left(\frac{\beta}{2}(N-n), \sigma^2\right)$.

Classical β-ensembles and random tridiagonal models

Dumitriu and Edelman (2002) and Killip and Nenciu (2004) derived equivalent random tridiagonal models, valid for $\beta > 0$,

$$\text{eigvals}\begin{bmatrix} a_1 & \sqrt{b_1} & & (0) \\ \sqrt{b_1} & a_2 & \ddots & \\ & \ddots & \ddots & \sqrt{b_{N-1}} \\ (0) & & \sqrt{b_{N-1}} & a_N \end{bmatrix}.$$

Example (Hermite ensemble, $\beta > 0$, $V(x) = \frac{1}{2\sigma^2}(x-\mu)^2$)

Consider independent $a_n \sim \mathcal{N}\left(\mu, \sigma^2\right)$, and $b_n \sim \Gamma\left(\frac{\beta}{2}(N-n), \sigma^2\right)$.

Random tridiagonal models grant $\mathcal{O}(N^2)$ exact samplers!

Classical β-ensembles and random tridiagonal models

Dumitriu and Edelman (2002) and Killip and Nenciu (2004) derived equivalent random tridiagonal models, valid for $\beta > 0$,

$$\operatorname{eigvals}\begin{bmatrix} a_1 & \sqrt{b_1} & & (0) \\ \sqrt{b_1} & a_2 & \ddots & \\ & \ddots & \ddots & \sqrt{b_{N-1}} \\ (0) & & \sqrt{b_{N-1}} & a_N \end{bmatrix}.$$

Example (Hermite ensemble, $\beta > 0$, $V(x) = \frac{1}{2\sigma^2}(x-\mu)^2$)

Consider independent $a_n \sim \mathcal{N}\left(\mu, \sigma^2\right)$, and $b_n \sim \Gamma\left(\frac{\beta}{2}(N-n), \sigma^2\right)$.

Random tridiagonal models grant $\mathcal{O}(N^2)$ exact samplers!

Extend tridiagonal models to more general β-ensembles ?

How to randomize the entries of the tridiagonal matrix?

Proposition (Krishnapur et. al, 2016)

Consider the random tridiagonal matrix $J_{\mathbf{a},\mathbf{b}}$ where the entries have joint density

$$\propto e^{-\operatorname{Tr} V(J_{\mathbf{a},\mathbf{b}})} \prod_{n=1}^{N-1} b_n^{\frac{\beta}{2}(N-n)-1}.$$

Then, the eigenvalues of $J_{\mathbf{a},\mathbf{b}}$ have joint density

$$\propto \left| \prod_{i<j} (x_j - x_i) \right|^{\beta} \prod_{n=1}^{N} e^{-V(x_n)}.$$

How to randomize the entries of the tridiagonal matrix?

Proposition (Krishnapur et. al, 2016)

Consider the random tridiagonal matrix $J_{\mathbf{a},\mathbf{b}}$ where the entries have joint density

$$\propto e^{-\operatorname{Tr} V(J_{\mathbf{a},\mathbf{b}})} \prod_{n=1}^{N-1} b_n^{\frac{\beta}{2}(N-n)-1}.$$

Then, the eigenvalues of $J_{\mathbf{a},\mathbf{b}}$ have joint density

$$\propto \left| \prod_{i<j} (x_j - x_i) \right|^{\beta} \prod_{n=1}^{N} e^{-V(x_n)}.$$

Gautier, Bardenet, and Valko (2020 - arXiv)

- Provide simple and clean proof. .

How to randomize the entries of the tridiagonal matrix?

Proposition (Krishnapur et. al, 2016)

Consider the random tridiagonal matrix $J_{\mathbf{a},\mathbf{b}}$ where the entries have joint density

$$\propto e^{-\operatorname{Tr} V(J_{\mathbf{a},\mathbf{b}})} \prod_{n=1}^{N-1} b_n^{\frac{\beta}{2}(N-n)-1}.$$

Then, the eigenvalues of $J_{\mathbf{a},\mathbf{b}}$ have joint density

$$\propto \left| \prod_{i<j} (x_j - x_i) \right|^{\beta} \prod_{n=1}^{N} e^{-V(x_n)} \prod_{n=1}^{N} \omega_n^{\frac{\beta}{2}-1}.$$

Gautier, Bardenet, and Valko (2020 - arXiv)

- Provide simple and clean proof, starting from $\mu^* = \sum_{n=1}^{N} \omega_n \delta_{x_n}$.

How to randomize the entries of the tridiagonal matrix?

Proposition (Krishnapur et. al, 2016)

Consider the random tridiagonal matrix $J_{\mathbf{a},\mathbf{b}}$ where the entries have joint density

$$\propto e^{-\operatorname{Tr} V(J_{\mathbf{a},\mathbf{b}})} \prod_{n=1}^{N-1} b_n^{\frac{\beta}{2}(N-n)-1}.$$

Then, the eigenvalues of $J_{\mathbf{a},\mathbf{b}}$ have joint density

$$\propto \left| \prod_{i<j} (x_j - x_i) \right|^{\beta} \prod_{n=1}^{N} e^{-V(x_n)} \prod_{n=1}^{N} \omega_n^{\frac{\beta}{2}-1}.$$

Gautier, Bardenet, and Valko (2020 - arXiv)

- Provide simple and clean proof, starting from $\mu^* = \sum_{n=1}^{N} \omega_n \delta_{x_n}$.
- Extend Krishnapur's result to unify the treatment of classical β-ensembles.

How to randomize the entries of the tridiagonal matrix?

Proposition (Krishnapur et. al, 2016)

Consider the random tridiagonal matrix $J_{\mathbf{a},\mathbf{b}}$ where the entries have joint density

$$\propto e^{-\operatorname{Tr} V(J_{\mathbf{a},\mathbf{b}})} \prod_{n=1}^{N-1} b_n^{\frac{\beta}{2}(N-n)-1}.$$

Then, the eigenvalues of $J_{\mathbf{a},\mathbf{b}}$ have joint density

$$\propto \left| \prod_{i<j} (x_j - x_i) \right|^{\beta} \prod_{n=1}^{N} e^{-V(x_n)} \prod_{n=1}^{N} \omega_n^{\frac{\beta}{2}-1}.$$

Gautier, Bardenet, and Valko (2020 - arXiv)

- Provide simple and clean proof, starting from $\mu^* = \sum_{n=1}^{N} \omega_n \delta_{x_n}$.
- Extend Krishnapur's result to unify the treatment of classical β-ensembles.
- Perform empirical study of tridiagonal models for polynomial potential V.

Tridiagonal models for polynomial potentials V

When degree $V = 2$,

- $(a_n), (b_n)$ are independent ☺
- have easy-to-sample distribution ☺

Example ($V(x) = \frac{1}{2\sigma^2}(x-\mu)^2$)

$$a_n \sim \mathcal{N}(\mu, \sigma^2), \; b_n \sim \Gamma\left(\frac{\beta}{2}(N-n), \sigma^2\right).$$

Tridiagonal models for polynomial potentials V

When degree $V = 2$,

- $(a_n), (b_n)$ are independent ☺
- have easy-to-sample distribution ☺

Example ($V(x) = \frac{1}{2\sigma^2}(x-\mu)^2$)

$$a_n \sim \mathcal{N}(\mu, \sigma^2), \; b_n \sim \Gamma\Big(\frac{\beta}{2}(N-n), \sigma^2\Big).$$

When degree $V > 2$,

- $(a_n), (b_n)$ are **not** independent ☹
- but have short range interaction ☺

Tridiagonal models for polynomial potentials V

When degree $V = 2$,

- $(a_n), (b_n)$ are independent ☺
- have easy-to-sample distribution ☺

Example ($V(x) = \frac{1}{2\sigma^2}(x-\mu)^2$)

$$a_n \sim \mathcal{N}(\mu, \sigma^2), \; b_n \sim \Gamma\left(\frac{\beta}{2}(N-n), \sigma^2\right).$$

When degree $V > 2$,

- $(a_n), (b_n)$ are **not** independent ☹
- but have short range interaction ☺

Example ($V(x) = g_4 x^4 + g_2 x^2$)

$a_n \mid \mathbf{a}_{\setminus n}, \mathbf{b}$

$$\sim \exp\left[-\left(g_4 a_n^4 + a_n^2[g_2 + 4g_4(b_{n-1} + b_n)] + 4g_4 a_n(a_{n-1}b_{n-1} + a_{n+1}b_n)\right)\right],$$

$b_n \mid \mathbf{a}, \mathbf{b}_{\setminus n}$

$$\sim b_n^{\frac{\beta}{2}(N-n)-1} \exp\left[-2\left(g_4 b_n^2 + b_n\left[g_2 + 2g_4({a_n}^2 + a_n a_{n+1} + {a_{n+1}}^2 + b_{n-1} + b_{n+1})\right]\right)\right].$$

Tridiagonal models for polynomial potentials V

When degree $V = 2$,

- $(a_n), (b_n)$ are independent ☺
- have easy-to-sample distribution ☺

Example $\left(V(x) = \frac{1}{2\sigma^2}(x-\mu)^2\right)$

$$a_n \sim \mathcal{N}(\mu, \sigma^2), \; b_n \sim \Gamma\left(\frac{\beta}{2}(N-n), \sigma^2\right).$$

When degree $V > 2$,

- $(a_n), (b_n)$ are **not** independent ☹
- but have short range interaction ☺

Example $\left(V(x) = g_4 x^4 + g_2 x^2\right)$

$a_n \mid \mathbf{a}_{\setminus n}, \mathbf{b}$

$$\sim \exp\left[-\left(g_4 a_n^4 + a_n^2[g_2 + 4g_4(b_{n-1} + b_n)] + 4g_4 a_n(a_{n-1}b_{n-1} + a_{n+1}b_n)\right)\right],$$

$b_n \mid \mathbf{a}, \mathbf{b}_{\setminus n}$

$$\sim b_n^{\frac{\beta}{2}(N-n)-1} \exp\left[-2\left(g_4 b_n^2 + b_n\left[g_2 + 2g_4({a_n}^2 + a_n a_{n+1} + {a_{n+1}}^2 + b_{n-1} + b_{n+1})\right]\right)\right].$$

This suggests a Gibbs sampling strategy!

Combining tridiagonal models with Gibbs sampling

Target: β-ensembles with potentials of the form

$$V(x) = g_6 x^6 + \cancel{g_5 x^5} + g_4 x^4 + g_3 x^3 + g_2 x^2 + g_1 x.$$

Combining tridiagonal models with Gibbs sampling

Target: β-ensembles with potentials of the form

$$V(x) = g_6 x^6 + \cancel{g_5 x^5} + g_4 x^4 + g_3 x^3 + g_2 x^2 + g_1 x.$$

- Systematic scan Gibbs sampler

$$\begin{aligned}
&\textbf{for } t = 1 \textbf{ to } T \\
&\quad \textbf{for } n = 1 \textbf{ to } N \\
&\quad\quad \text{sample } a_n \mid \mathbf{a}_{\setminus n}, \mathbf{b} \\
&\quad\quad \text{sample } b_n \mid \mathbf{a}, \mathbf{b}_{\setminus n} \textbf{ if } n < N \\
&\quad \{x_1^t, \dots, x_N^t\} = \text{eigvals } J_{\mathbf{a},\mathbf{b}}
\end{aligned}$$

Combining tridiagonal models with Gibbs sampling

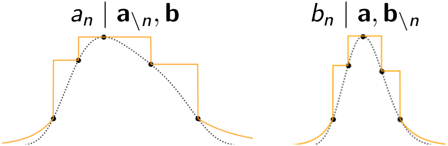
Target: β-ensembles with potentials of the form

$$V(x) = g_6 x^6 + \cancel{g_5 x^5} + g_4 x^4 + g_3 x^3 + g_2 x^2 + g_1 x.$$

- Systematic scan Gibbs sampler

  for $t = 1$ **to** T
  for $n = 1$ **to** N
  sample $a_n \mid \mathbf{a}_{\setminus n}, \mathbf{b}$
  sample $b_n \mid \mathbf{a}, \mathbf{b}_{\setminus n}$ **if** $n < N$
  $\{x_1^t, \ldots, x_N^t\} = \text{eigvals}\, J_{\mathbf{a},\mathbf{b}}$

- Exact sampling of log-concave conditionals (Devroye, 2012).
  - e.g., $V(x) = \frac{1}{4}x^4$.



Combining tridiagonal models with Gibbs sampling

Target: β-ensembles with potentials of the form

$$V(x) = g_6 x^6 + \cancel{g_5 x^5} + g_4 x^4 + g_3 x^3 + g_2 x^2 + g_1 x.$$

- Systematic scan Gibbs sampler

  for $t = 1$ **to** T
    for $n = 1$ **to** N
      sample $a_n \mid \mathbf{a}_{\setminus n}, \mathbf{b}$
      sample $b_n \mid \mathbf{a}, \mathbf{b}_{\setminus n}$ **if** $n < N$
    $\{x_1^t, \dots, x_N^t\} = \operatorname{eigvals} J_{\mathbf{a},\mathbf{b}}$

- Exact sampling of log-concave conditionals (Devroye, 2012).
  - e.g., $V(x) = \frac{1}{4}x^4$.
- Metropolis-Hastings kernel (MALA) for **non** log-concave conditionals.
  - e.g., $V(x) = \frac{1}{6}x^6$.

Combining tridiagonal models with Gibbs sampling

Target: β-ensembles with potentials of the form

$$V(x) = g_6 x^6 + \cancel{g_5 x^5} + g_4 x^4 + g_3 x^3 + g_2 x^2 + g_1 x.$$

- Systematic scan Gibbs sampler

```
for t = 1 to T
   for n = 1 to N
      sample a_n | a_\n, b
      sample b_n | a, b_\n if n < N
   {x_1^t, ..., x_N^t} = eigvals J_a,b
```

- Exact sampling of log-concave conditionals (Devroye, 2012).
  - e.g., $V(x) = \frac{1}{4}x^4$.
- Metropolis-Hastings kernel (MALA) for **non** log-concave conditionals.
  - e.g., $V(x) = \frac{1}{6}x^6$.

How does it perform?

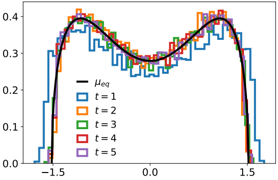
Monitoring of the empirical convergence

Convergence of the empirical marginal distribution to the equilibrium measure.

$$\widehat{\mu}_N^t = \frac{1}{N} \sum_{n=1}^{N} \delta_{x_n^t} \xrightarrow[N,t\to\infty]{} \mu_{\text{eq}}.$$

- $V(x) = \frac{1}{4}x^4$, exact sampling of the conditionals.

$N = 20$

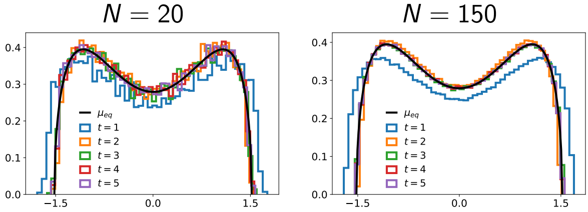


Monitoring of the empirical convergence

Convergence of the empirical marginal distribution to the equilibrium measure.

$$\widehat{\mu}_N^t = \frac{1}{N}\sum_{n=1}^{N} \delta_{x_n^t} \xrightarrow[N,t\to\infty]{} \mu_{\text{eq}}.$$

- $V(x) = \frac{1}{4}x^4$, exact sampling of the conditionals.

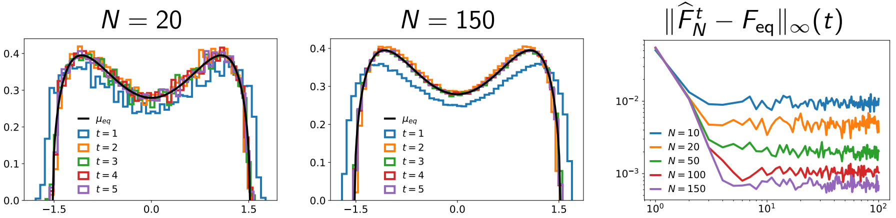


Monitoring of the empirical convergence

Convergence of the empirical marginal distribution to the equilibrium measure.

$$\widehat{\mu}_N^t = \frac{1}{N} \sum_{n=1}^{N} \delta_{x_n^t} \xrightarrow[N,t\to\infty]{} \mu_{\text{eq}}.$$

- $V(x) = \frac{1}{4}x^4$, exact sampling of the conditionals.

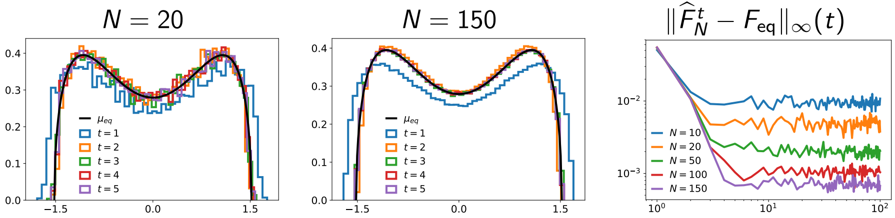


Monitoring of the empirical convergence

Convergence of the empirical marginal distribution to the equilibrium measure.

$$\widehat{\mu}_N^t = \frac{1}{N}\sum_{n=1}^{N} \delta_{x_n^t} \xrightarrow[N,t\to\infty]{} \mu_{\text{eq}}.$$

- $V(x) = \frac{1}{4}x^4$, exact sampling of the conditionals.



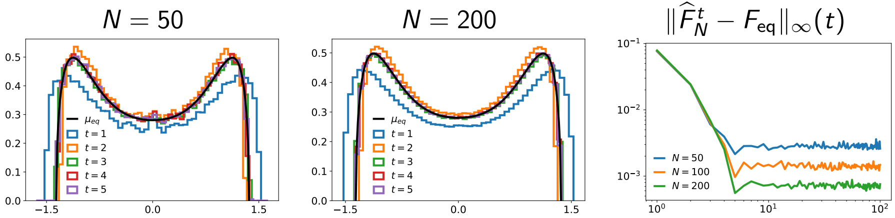
- Good adequation with the theory.
- Empirical convergence within $t \leq 10$ Gibbs passes, **only!**

Monitoring of the empirical convergence

Convergence of the empirical marginal distribution to the equilibrium measure.

$$\widehat{\mu}_N^t = \frac{1}{N}\sum_{n=1}^{N} \delta_{x_n^t} \xrightarrow[N,t\to\infty]{} \mu_{\text{eq}}.$$

- $V(x) = \frac{1}{6}x^6$, approximate sampling of the conditionals.



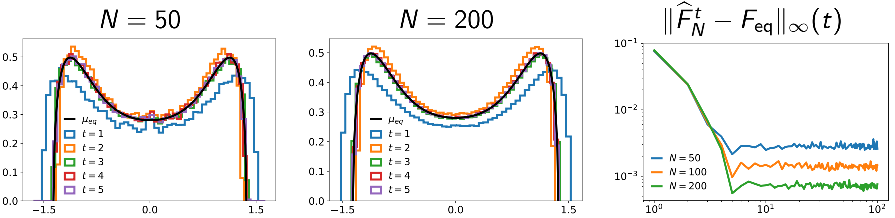
- Good adequation with the theory.
- Empirical convergence within $t \leq 10$ Gibbs passes, **only!**

Monitoring of the empirical convergence

Convergence of the empirical marginal distribution to the equilibrium measure.

$$\widehat{\mu}_N^t = \frac{1}{N}\sum_{n=1}^{N} \delta_{x_n^t} \xrightarrow[N,t\to\infty]{} \mu_{\text{eq}}.$$

- $V(x) = \frac{1}{6}x^6$, approximate sampling of the conditionals.



- Good adequation with the theory.
- Empirical convergence within $t \leq 10$ Gibbs passes, **only!**

Supports the $\mathcal{O}(\log(N))$ mixing time conjecture of Krishnapur et. al (2016).

Overview

My Ph.D. in a nutshell

DPP sampling

Finite setting

Finite projection DPP
Approximate sampling
Linear programming

ICML, 2017

Continuous setting

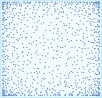
β-ensembles
Gibbs sampling
Random matrices

eigvals 

Submitted, 2020

Continuous setting

Monte Carlo integration
Exact sampling
Random linear system



NeurIPS, 2019

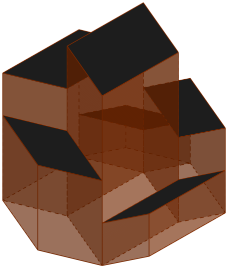
Software engineering

Python toolbox
DPPy
Reproducible research

```python
from dppy import *
# [...]
dpp.sample()
```

JMLR-MLOSS, 2019

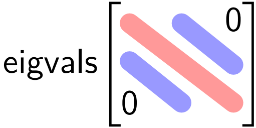
Zonotope sampling for finite projection DPPs

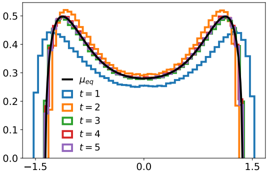


- New perspective on finite projection DPPs.
- Combination of geometry, Markov chains and linear programming.
- Approximate sampler involving randomized linear programs.
- More efficient exploration of the state space.

ICML, 2017

Tridiagonal models for sampling β-ensembles





- Unified treatment of tridiagonal models for the classical β-ensembles.
- Combination of a Gibbs sampler with calculation of eigenvalues.
- Very fast empirical convergence supporting the $\mathcal{O}(\log(N))$ mixing time conjecture.

Submitted to an international journal, 2020

Monte Carlo integration with DPPs

Let $\{x_1, \dots, x_N\} \sim \mathrm{DPP}(K, \mu)$, where $K(x,y) = \sum_{k=0}^{N-1} \phi_k(x)\phi_k(y)$.

$$\int f(x)\mu(\mathrm{d}x) \approx \sum_{n=1}^{N} \omega_n f(x_n),$$

- Shed light on the estimator of Ermakov and Zolotukhin (1960)
  - involving a randomized linear system
  - provide new simple proofs of its properties

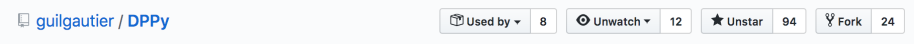
$$\mathbb{V}\mathrm{ar} = \|f\|^2 - \sum_{k=0}^{N-1} \langle f, \phi_k \rangle^2.$$

- Numerical comparison with the estimator of Bardenet and Hardy (2020)
- Tailored implementation of the chain rule.

Adapt the kernel K to the basis where f has a smooth/sparse expansion.

NeurIPS, 2019

DPPy: DPP sampling with Python




```python
from dppy import *
# [...]
dpp.sample()
```

- Open source toolbox.
- Implementation of exact and approximate samplers.
- Extensive documentation.

JMLR-MLOSS, 2019

Overview

Open questions

- Zonotope
  - prove a bound on the mixing time.
  - extend the LP idea for continuous DPPs.
- β-ensembles
  - prove the $\mathcal{O}(\log(N))$ mixing time for the Gibbs sampler.
  - extend tridiagonal models for multivariate β-ensembles.
- Efficient sampler for continuous projection DPPs ($d > 1$)?
- Avoid kernel eigendecomposition for sampling non-projection DPPs?

Perspectives

- Find a good reparametrization of DPPs where
  - complex interaction structure vanishes.
  - efficient sampling can be performed.
- Continuous extension of sampling by solving linear programs.
- Sampling by coupling the target DPP with another process.
  - Decreusefond, Flint, and Low (2013), Launay, Galerne, and Desolneux (2018), and Dereziński, Calandriello, and Valko (2019).
- Continue developing the DPPy toolbox.

Thank you!
Ευχαριστώ!
Merci!


My PhD story
COMMENT GÉNÉRER DES RECOMMANDATIONS DIVERSES
From dppy import *
det K_S
NOUVEAUX ALGORITHMES
CONFÉRENCES
Var[î] = ||f||² - Σ ⟨f|φk⟩²
ARTICLES SCIENTIFIQUES
RECHERCHE REPRODUCTIBLE
GUIllaume
MISSIONS
Monsieur Jeanne

References [1]

Anari, N., S. O. Gharan, and A. Rezaei. 2016. *Monte Carlo Markov Chain Algorithms for Sampling Strongly Rayleigh Distributions and Determinantal Point Processes.* In Conference on Learning Theory (COLT). arXiv:1602.05242. (see slides 49, 50, 51, 52, 53).

Bardenet, R., and A. Hardy. 2020. *Monte Carlo with Determinantal Point Processes.* Annals of Applied Probability. arXiv:1605.00361. (see slides 4, 133, 148).

Chen, Y., R. Dwivedi, M. Wainwright, and Y. Bin. 2018. *Fast MCMC Sampling Algorithms on Polytopes.* Journal of Machine Learning Research. (see slides 70, 71, 72, 73, 74).

Decreusefond, L., I. Flint, and K. C. Low. 2013. *Perfect Simulation of Determinantal Point Processes.* ArXiv e-prints. arXiv:1311.1027. (see slide 137).

References [2]

Dereziński, M., D. Calandriello, and M. Valko. 2019. *Exact sampling of determinantal point processes with sublinear time preprocessing.* In Advances in Neural Information Processing Systems (NeurIPS), edited by H. W. Garnett, H. Larochelle, A. Beygelzimer, F. D'Alché-Buc, E. Fox, and R. Garnett. Vancouver, Canada: Curran Associates, Inc. arXiv:1905.13476. (see slide 137).

Dette, H., and W. J. Studden. 1997. *The theory of canonical moments with applications in statistics, probability, and analysis.* Wiley. (see slide 93).

Devroye, L. 2012. *A note on generating random variables with log-concave densities.* Technical report. (see slides 119, 120, 121, 122).

Dumitriu, I., and A. Edelman. 2002. *Matrix models for beta ensembles.* Journal of Mathematical Physics. arXiv:math-ph/0206043. (see slides 105, 106, 107, 108).

References [3]

Dyer, M., and A. Frieze. 1994. *Random walks, totally unimodular matrices, and a randomised dual simplex algorithm.* Mathematical Programming. (see slides 63, 64, 65, 66, 67, 68).

Ermakov, S. M., and V. G. Zolotukhin. 1960. *Polynomial Approximations and the Monte-Carlo Method.* Theory of Probability and Its Applications. (see slides 133, 149, 150).

Feder, T., and M. Mihail. 1992. *Balanced matroids.* Proceedings of the twenty-fourth annual ACM. (see slides 49, 50, 51, 52, 53).

Gautier, G., R. Bardenet, and M. Valko. 2017. *Zonotope hit-and-run for efficient sampling from projection DPPs.* In International Conference on Machine Learning (ICML). arXiv:1705.10498. (see slides 5, 6, 7, 8, 9, 10, 58, 69, 70, 71, 72, 73, 74, 98, 130, 131).

———. 2019. *On two ways to use determinantal point processes for Monte Carlo integration.* In Advances in Neural Information Processing Systems (NeurIPS). (see slides 5, 6, 7, 8, 9, 10, 58, 98, 130, 133).

References [4]

Gautier, G., R. Bardenet, and M. Valko. 2020. *Fast sampling from β-ensembles.* ArXiv e-prints. arXiv:2003.02344. (see slides 5, 6, 7, 8, 9, 10, 58, 98, 110, 111, 112, 113, 130, 132).

Gautier, G., G. Polito, R. Bardenet, and M. Valko. 2019. *DPPy: DPP Sampling with Python.* Journal of Machine Learning Research - Machine Learning Open Source Software (JMLR-MLOSS). arXiv:1809.07258. (see slides 5, 6, 7, 8, 9, 10, 58, 98, 130, 134).

Gillenwater, J. 2014. *Approximate inference for determinantal point processes.* PhD dissertation, University of Pennsylvania. (see slides 30, 31, 32, 33, 34, 35, 36, 37).

Goberna, M. A., and M. A. López. 2014. *Post-Optimal Analysis in Linear Semi-Infinite Optimization.* Springer, New York, NY. (see slide 89).

Hermon, J., and J. Salez. 2019. *Modified log-Sobolev inequalities for strong-Rayleigh measures.* arXiv:1902.02775. (see slides 49, 50, 51, 52, 53).

References [5]

Hough, J. B., M. Krishnapur, Y. Peres, and B. Virág. 2006. *Determinantal Processes and Independence.* In Probability Surveys. arXiv:math/0503110. (see slides 30, 31, 32, 33, 34, 35, 36, 37).

Killip, R., and I. Nenciu. 2004. *Matrix models for circular ensembles.* International Mathematics Research Notices. arXiv:math/0410034. (see slides 105, 106, 107, 108, 152).

Krishnapur, M., B. Rider, and B. Virág. 2016. *Universality of the Stochastic Airy Operator.* Communications on Pure and Applied Mathematics. arXiv:arXiv:1306.4832. (see slides 109, 110, 111, 112, 113, 123, 124, 125, 126, 127, 128).

Lasserre, J.-B. 2010. *Moments, positive polynomials and their applications.* Imperial College Press. (see slides 90, 91).

Launay, C., B. Galerne, and A. Desolneux. 2018. *Exact Sampling of Determinantal Point Processes without Eigendecomposition.* ArXiv e-prints. arXiv:1802.08429. (see slide 137).

References [6]

Li, C., S. Jegelka, and S. Sra. 2016. *Fast Mixing Markov Chains for Strongly Rayleigh Measures, DPPs, and Constrained Sampling.* In Advances in Neural Information Processing Systems (NIPS). Barcelona, Spain. arXiv:1608.01008. (see slides 49, 50, 51, 52, 53).

Lovász, L., and S. Vempala. 2003. *Hit-and-Run is Fast and Fun.* Technical report. Microsoft Research. (see slides 70, 71, 72, 73, 74).

Luenberger, D. G., and Y. Ye. 2016. *Linear and Nonlinear Programming.*

L-ensembles and k-DPPs

Definition (L-ensemble)

Let $\mathbf{L} \succeq 0$. The point process defined by

$$\mathbb{P}[\mathcal{X} = S] = \frac{\det \mathbf{L}_S}{\det(I + \mathbf{L})},$$

is called an L-ensemble. It is a DPP with kernel $\mathbf{K} = \mathbf{L}(I + \mathbf{L})^{-1}$.

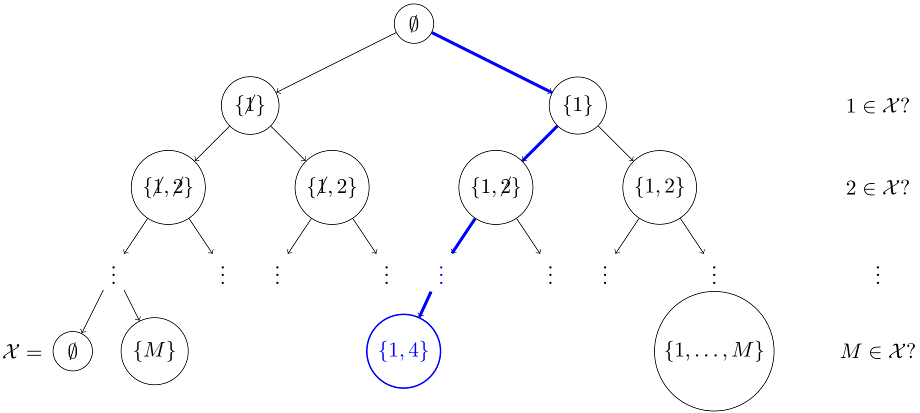
Definition (k-DPP)

Let $\mathbf{L} \succeq 0$ and $k \in \mathbb{N}^*$. The point process defined by

$$\mathbb{P}[\mathcal{X} = S] \propto \det \mathbf{L}_S \; \mathbb{1}_{|S|=k}.$$

is called a k-DPP.

Chain rule on sets



BH estimator and the multivariate Jacobi ensemble

Natural unbiased estimator of $\int_{\mathbb{X}} f(x)\mu(\mathrm{d}x)$

$$\boxed{\widehat{I}_N^{\mathrm{BH}}(f) = \sum_{n=1}^{N} \frac{f(x_n)}{K(x_n, x_n)}}$$

- Bardenet and Hardy (2020) show fast CLT, for f essentially $\mathcal{C}^1$

$$\sqrt{N^{1+1/d}}\left(\widehat{I}_N^{\mathrm{BH}}(f) - \int_{[-1,1]^d} f(x)\,\omega(x)\mathrm{d}x\right) \xrightarrow[N\to\infty]{\text{law}} \mathcal{N}\left(0, \mathbf{\Omega}_{f,\omega}^2\right),$$

with $\mathbf{\Omega}_{f,\omega}^2 \triangleq \frac{1}{2} \sum_{k\in\mathbb{N}^d} (k_1 + \cdots + k_d)\ \mathcal{F}\left[\frac{f\,\omega}{\omega_{\mathrm{eq}}}\right](k)^2$

Theorem (Ermakov and Zolotukhin, 1960)

$$f = \sum_{\ell=0}^{M-1} \langle f, \phi_\ell \rangle \phi_\ell, \quad M \in \mathbb{N} \cup \{\infty\}$$

1. Sample $\{x_1, \dots, x_N\} \sim \mathrm{DPP}(\mu, K)$ with $K(x, y) = \sum_{k=0}^{N-1} \phi_k(x) \phi_k(y)$
2. Random linear system

$$\begin{bmatrix} \phi_0(x_1) & \dots & \phi_{N-1}(x_1) \\ \vdots & & \vdots \\ \phi_0(x_N) & \dots & \phi_{N-1}(x_N) \end{bmatrix} \begin{bmatrix} y_0 \\ \vdots \\ y_{N-1} \end{bmatrix} = \begin{bmatrix} f(x_1) \\ \vdots \\ f(x_N) \end{bmatrix}$$

- $\mathbb{E}[y_k] = \langle f, \phi_k \rangle = \int f(x) \phi_k(x) \mu(\mathrm{d}x)$
- $\mathbb{V}\mathrm{ar}[y_k] = \|f\|^2 - \sum_{\ell=0}^{N-1} \langle f, \phi_\ell \rangle^2 = \sum_{\ell=N}^{M-1} \langle f, \phi_\ell \rangle^2 = 0 \quad \text{if } M \leq N$
- $\mathbb{C}\mathrm{ov}[y_j, y_k] = 0, j \neq k$

Ermakov and Zolotukhin (1960) estimator

For constant ϕ_0, e.g., multivariate Jacobi ensemble,

$$\mathbb{E}[y_0] = \phi_0 \int_{\mathbb{X}} f(x)\mu(\mathrm{d}x)$$

A direct application of EZ theorem yields

$$\widehat{I}_N^{\mathrm{EZ}}(f) \triangleq \frac{y_0}{\phi_0} = \sqrt{\mu([-1,1]^d)}\ \frac{\det \mathbf{\Phi}_{\phi_0,f}(x_{1:N})}{\det \mathbf{\Phi}(x_{1:N})}$$

as an unbiased estimator of $\int f(x)\mu(\mathrm{d}x)$

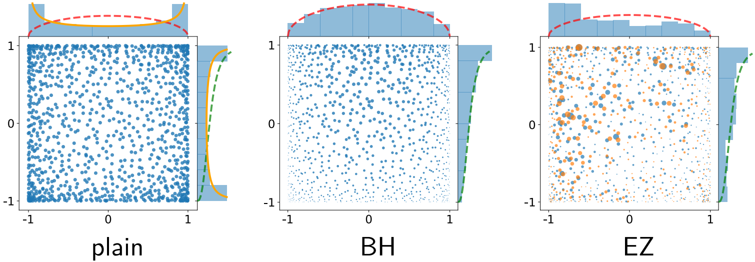
Using $\|\phi_0\| = 1$ and Cramer's rule

$$\mathbf{\Phi}_{\phi_0,f} = \begin{bmatrix} f(x_1) & \dots & \psi_{N-1}(x_1) \\ \vdots & & \vdots \\ f(x_N) & \dots & \psi_{N-1}(x_N) \end{bmatrix} \quad \mathbf{\Phi} = \begin{bmatrix} \phi_0(x_1) & \dots & \phi_{N-1}(x_1) \\ \vdots & & \vdots \\ \phi_0(x_N) & \dots & \phi_{N-1}(x_N) \end{bmatrix}$$

Comparison weights ω_n BH-EZ

$$\int_{\mathbb{X}} f(x)\mu(\mathrm{d}x) \approx \widehat{I}_N = \sum_{n=1}^{N} \omega_n(x_1, \ldots, x_N) f(x_n)$$

- weights ω_n

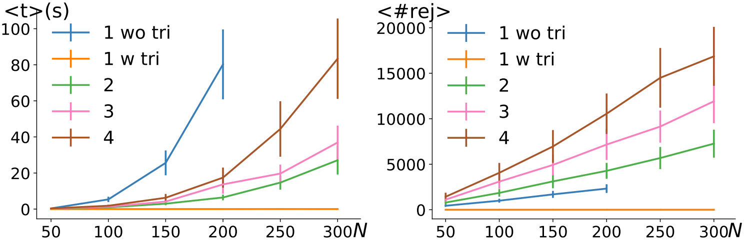


- Non-asymptotic variance

$$\mathbb{V}\mathrm{ar}\left[\widehat{I}_N^{\mathrm{BH}}\right] = \frac{1}{2}\int_{\mathbb{X}^2} \left(\frac{f(x)}{K(x,x)} - \frac{f(y)}{K(y,y)}\right)^2 K(x,y)^2 \mu(\mathrm{d}x)\mu(\mathrm{d}y)$$

$$\mathbb{V}\mathrm{ar}\left[\widehat{I}_N^{\mathrm{EZ}}\right] = \|f\|^2 - \sum_{\ell=0}^{N-1} \langle f, \phi_\ell \rangle^2$$

Timings



(a) ⟨time⟩ to get one sample (b) ⟨#rejections⟩

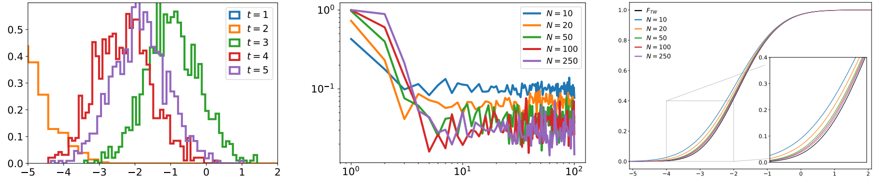
Figure 1: The colors and numbers correspond to the dimension. $a_i, b_i = -1/2$. For $d = 1$, the tridiagonal model (tri) of Killip and Nenciu (BH, 2004) offers tremendous savings, without it is cheaper to get a sample in larger dimension. The number of rejections grows as $N \log(N) 2^d$.

Monitoring of the empirical convergence ($\lambda_{\max}, \beta = 2$)

Convergence of the distribution of the largest eigenvalue to Tracy-Widom.

$$\text{rescaled } \lambda^t_{\max} \xrightarrow[N,t\to\infty]{\text{law}} \mathrm{TW}_2\,.$$

- $V(x) = \frac{1}{4}x^4$, #indepedent runs $= 10^3$.



- $V(x) = \frac{1}{6}x^6$, #indepedent runs $= 10^3$.