A Thinning Analogue of de Finetti’s Theorem

Shannon Starr

University of Rochester

9 Jan 2007
Finite Exchangeability

$$\mu_n \in \mathcal{M}_1(\Omega^n)$$

$$(X_1, \ldots, X_n) \text{ random, } \mu_n\text{-distributed}$$

$$\mu_n \cdot \pi := \text{distribution of } (X_{\pi(1)}, \ldots, X_{\pi(n)})$$
Finite Exchangeability

\( \mu_n \in \mathbf{M}_1(\Omega^n) \)

\((X_1, \ldots, X_n)\) random, \(\mu_n\)-distributed

\(\mu_n \cdot \pi := \text{distribution of } (X_{\pi(1)}, \ldots, X_{\pi(n)})\)

\(\mu_n\) is symmetric if \(\mu_n \cdot \pi = \mu_n\) for all \(\pi \in S_n\).

Let us call this “finite exchangeability”.

Shannon Starr
A Thinning Analogue of de Finetti’s Theorem
Assume $\Omega = \mathcal{X}$ a compact metric space

$\mathcal{X}^\infty := (x_1, x_2, \ldots)$, all $x_i$ in $\mathcal{X}$

$S_\infty := \text{bijective } \pi : \mathbb{Z}_{>0} \rightarrow \mathbb{Z}_{>0}$ s.t. $\#\{i : \pi(i) \neq i\} < \infty$
Infinite Exchangeability

Assume $\Omega = \mathcal{X}$ a compact metric space

$\mathcal{X}^\infty := (x_1, x_2, \ldots)$, all $x_i$ in $\mathcal{X}$

$S_\infty :=$ bijections $\pi: \mathbb{Z}_0^+ \rightarrow \mathbb{Z}_0^+$ s.t. $\#\{i : \pi(i) \neq i\} < \infty$

$\mu_\infty \in M_1(\mathcal{X}^\infty)$

$(X_1, X_2, \ldots)$ random, $\mu_\infty$-distributed

$\mu_\infty \cdot \pi :=$ distribution of $(X_{\pi(1)}, X_{\pi(2)}, \ldots)$. 
Infinite Exchangeability

Assume $\Omega = \mathcal{X}$ a compact metric space

$\mathcal{X}^\infty := (x_1, x_2, \ldots)$, all $x_i$ in $\mathcal{X}$

$S_\infty := \text{bijections } \pi : \mathbb{Z}_> \rightarrow \mathbb{Z}_> \text{ s.t. } \#\{i : \pi(i) \neq i\} < \infty$

$\mu_\infty \in M_1(\mathcal{X}^\infty)$

$(X_1, X_2, \ldots)$ random, $\mu_\infty$-distributed

$\mu_\infty \cdot \pi := \text{distribution of } (X_{\pi(1)}, X_{\pi(2)}, \ldots)$.

$\mu_\infty$ is infinitely exchangeable if $\mu_\infty \cdot \pi = \mu_\infty$ for all $\pi \in S_\infty$. 
Specific Example

\[ \mathcal{X} = \{0, 1\} \]

Bernoulli distribution: \( \beta_p = p \cdot \delta_1 + (1 - p) \delta_0 \)

independent product measure: \( \beta_{p_1} \otimes \beta_{p_2} \otimes \cdots \)
Specific Example

\( \mathcal{X} = \{0, 1\} \)

Bernoulli distribution: \( \beta_p = p \cdot \delta_1 + (1 - p) \delta_0 \)

independent product measure: \( \beta_{p_1} \otimes \beta_{p_2} \otimes \cdots \)

\[
\begin{align*}
X_1, & \quad X_2, & \quad X_3, & \quad X_4, & \quad X_5, \ldots \\
X_3, & \quad X_1, & \quad X_2, & \quad X_4, & \quad X_5, \ldots
\end{align*}
\]

\[
(\beta_{p_1} \otimes \beta_{p_2} \otimes \cdots) \cdot \pi = \beta_{p_{\pi(1)}} \otimes \beta_{p_{\pi(2)}} \otimes \cdots
\]
Specific Example

\[ \mathcal{X} = \{0, 1\} \]

Bernoulli distribution: \( \beta_p = p \cdot \delta_1 + (1 - p)\delta_0 \)

Independent product measure: \( \beta_{p_1} \otimes \beta_{p_2} \otimes \cdots \)

\[
\begin{align*}
X_1, & \quad X_2, & \quad X_3, & \quad X_4, & \quad X_5, & \quad \ldots \\
\downarrow & & & \downarrow & & \\
X_3, & \quad X_1, & \quad X_2, & \quad X_4, & \quad X_5, & \quad \ldots \\
\end{align*}
\]

\[
(\beta_{p_1} \otimes \beta_{p_2} \otimes \cdots) \cdot \pi = \beta_{p_{\pi(1)}} \otimes \beta_{p_{\pi(2)}} \otimes \cdots
\]

So \( \beta_{p^{\otimes \infty}} \) is exchangeable for all \( p \in [0, 1] \).
More general example

Any mixture of i.i.d. product measure is exchangeable.
More general example

Any mixture of i.i.d. product measure is exchangeable.

Still consider $\mathcal{X} = \{0, 1\}$.

Suppose $P \in M_1([0, 1])$, mixing measure.

Define

$$\mu_\infty(\cdot) = \int_0^1 P(dp) \beta_p^{\otimes \infty}(\cdot)$$
Any mixture of i.i.d. product measure is exchangeable.

Still consider $\mathcal{X} = \{0, 1\}$.

Suppose $P \in \mathbf{M}_1([0, 1])$, mixing measure.

Define

$$\mu_\infty(\cdot) = \int_0^1 P(dp) \beta_p \otimes \infty(\cdot)$$

De Finetti’s theorem says every infinitely exchangeable measure is a mixture of i.i.d. product measures.
De Finetti’s Theorem

Theorem

Suppose $\mu_\infty \in M_1(\mathcal{X}^\infty)$ is exchangeable. Then there is a unique probability measure $P \in M_1(M_1(\mathcal{X}))$ such that

$$\mu_\infty(\cdot) = \int_{M_1(\mathcal{X})} P(d\nu) \nu^\otimes\infty(\cdot)$$
De Finetti’s Theorem

Theorem
Suppose $\mu_\infty \in \mathbf{M}_1(\mathcal{X}^\infty)$ is exchangeable. Then there is a unique probability measure $P \in \mathbf{M}_1(\mathbf{M}_1(\mathcal{X}))$ such that

$$\mu_\infty(\cdot) = \int_{\mathbf{M}_1(\mathcal{X})} P(d\nu) \nu^{\otimes \infty}(\cdot)$$

Note: This also implies the extreme points are i.i.d. product measures.
The right action generalizes to maps other than permutations.

\[ \phi : [m] \rightarrow [n] \] and \( \mu_n \in M_1(\mathcal{X}^n) \sim \mu_n \cdot \phi \in M_1(\mathcal{X}^m) \):
The right action generalizes to maps other than permutations.

\( \phi : [m] \rightarrow [n] \) and \( \mu_n \in \mathbf{M}_1(\mathcal{X}^n) \sim \mu_n \cdot \phi \in \mathbf{M}_1(\mathcal{X}^m) \):

Let \((X_1, \ldots, X_n)\) be \(\mu_n\)-distributed.

\( \mu_n \cdot \phi = \text{distribution of } (X_{\phi(1)}, \ldots, X_{\phi(m)}) \)
Thinning

The right action generalizes to maps other than permutations.

\[ \phi : [m] \to [n] \text{ and } \mu_n \in M_1(\mathcal{X}^n) \sim \mu_n \cdot \phi \in M_1(\mathcal{X}^m) : \]

Let \((X_1, \ldots, X_n)\) be \(\mu_n\)-distributed.

\[ \mu_n \cdot \phi = \text{distribution of } (X_{\phi(1)}, \ldots, X_{\phi(m)}) \]

“Thinning” = linear transformation \(\Theta_{n-1}^n : M_1(\mathcal{X}^n) \to M_1(\mathcal{X}^{n-1})\),

\[ \phi_{n,k} : [n - 1] \to [n]: \]
\[ (\phi_{n,k}(1), \ldots, \phi_{n,k}(n - 1)) = (1, \ldots, k - 1, k + 1, \ldots, n). \]

\[ \mu_n \Theta_{n-1}^n = \frac{1}{n} \sum_{k=1}^{n} \mu_n \cdot \phi_{n,k} , \]
(X_1, \ldots, X_n) \text{ random, } \mu_n\text{-distributed.}
(\(X_1, \ldots, X_n\)) random, \(\mu_n\)-distributed.

\(k \in [n]\) uniform and independent of \((X_1, \ldots, X_n)\).
(X_1, \ldots, X_n) \text{ random, } \mu_n\text{-distributed.}

k \in [n] \text{ uniform and independent of } (X_1, \ldots, X_n).

\mu_n \Theta^n_{n-1} \text{ is distribution of } (X_1, \ldots, X_{k-1}, X_{k+1}, \ldots, X_n).
Thinning-invariance

Sequence $\mu = (\mu_1, \mu_2, \ldots)$ is “thinning-invariant” if
$\mu_n \Theta^n_{n-1} = \mu_{n-1}$ all $n > 1$
Thinning-invariance

Sequence $\mu = (\mu_1, \mu_2, \ldots)$ is “thinning-invariant” if
$$\mu_n \Theta_n^{n-1} = \mu_{n-1} \quad \text{all} \ n > 1$$

Call $\mu$ “exchangeable” if:
(1) each $\mu_n \in M_1(\mathcal{X}^n)$ is finitely-exchangeable;
(2) $\mu$ is thinning-invariant.
Sequence $\mu = (\mu_1, \mu_2, \ldots)$ is “thinning-invariant” if $\mu_n \Theta^n_{n-1} = \mu_{n-1}$ all $n > 1$

Call $\mu$ “exchangeable” if:
(1) each $\mu_n \in M_1(\mathcal{X}^n)$ is finitely-exchangeable;
(2) $\mu$ is thinning-invariant.

Because of finite exchangeability, $\mu_n \Theta^n_{n-1}$ is equal-in-distribution to the marginal distribution of $(X_1, \ldots, X_{n-1})$, where $(X_1, \ldots, X_n)$ is $\mu_n$-distributed.
Sequence $\mu = (\mu_1, \mu_2, \ldots)$ is “thinning-invariant” if $\mu_n \Theta_n = \mu_{n-1}$ all $n > 1$.

Call $\mu$ “exchangeable” if:
1. each $\mu_n \in M_1(\mathcal{X}^n)$ is finitely-exchangeable;
2. $\mu$ is thinning-invariant.

Because of finite exchangeability, $\mu_n \Theta_n$ is equal-in-distribution to the marginal distribution of $(X_1, \ldots, X_{n-1})$, where $(X_1, \ldots, X_n)$ is $\mu_n$-distributed.

Kolmogorov’s extension theorem $\Rightarrow \exists$ infinitely exchangeable $\mu_{\infty} \in M_1(\mathcal{X}^\infty)$, such that $\mu_n$ is the marginal distribution of $(X_1, \ldots, X_n)$, where $(X_1, X_2, \ldots)$ is $\mu_{\infty}$-distributed.
Another version of de Finetti’s Theorem

**Theorem**

Suppose $\mu = (\mu_1, \mu_2, \ldots)$ is “exchangeable”. Then there is a unique probability measure $P \in M_1(M_1(\mathcal{X}))$ such that, for each $n > 0$,

$$\mu_n(\cdot) = \int_{M_1(\mathcal{X})} P(d\nu) \nu^\otimes n(\cdot)$$
Main Question

There is a representation theorem for sequences $\mu$ which are thinning-invariant, and such that each $\mu_n$ is finitely exchangeable.

Is there a representation theorem for sequences which are thinning-invariant but not necessarily finitely-exchangeable?
An example of thinning-invariance: Order Statistics

Define $I := [0, 1]$.

$\lambda := $ Borel version of standard Lebesgue measure on $I$

Let $\lambda = I$. 

An example of thinning-invariance: Order Statistics

Define $\mathcal{I} := [0, 1]$. 

$\lambda :=$ Borel version of standard Lebesgue measure on $\mathcal{I}$

Let $\lambda' = \mathcal{I}$.

Let $T_1, \ldots, T_n \in \mathcal{I}$ be i.i.d. $\lambda$-distributed r.v.'s.

Almost surely, there is a unique $\hat{\pi} \in S_n$ such that

$$T_{\hat{\pi}(1)} < \ldots < T_{\hat{\pi}(n)}.$$
An example of thinning-invariance: Order Statistics

Define $\mathcal{I} := [0, 1]$.

$\lambda :=$ Borel version of standard Lebesgue measure on $\mathcal{I}$

Let $\mathcal{X} = \mathcal{I}$.

Let $T_1, \ldots, T_n \in \mathcal{I}$ be i.i.d. $\lambda$-distributed r.v.'s.

Almost surely, there is a unique $\hat{\pi} \in S_n$ such that

$$T_{\hat{\pi}(1)} < \ldots < T_{\hat{\pi}(n)}.$$ 

Let $\mu_n \in M_1(\mathcal{I}^n)$ be the distribution of $(T_{\hat{\pi}(1)}, \ldots, T_{\hat{\pi}(n)})$. 

Shannon Starr  A Thinning Analogue of de Finetti’s Theorem
Define $\mathcal{I} := [0, 1]$.

$\lambda := \text{Borel version of standard Lebesgue measure on } \mathcal{I}$

Let $\mathcal{X} = \mathcal{I}$.

Let $T_1, \ldots, T_n \in \mathcal{I}$ be i.i.d. $\lambda$-distributed r.v.'s.

Almost surely, there is a unique $\hat{\pi} \in S_n$ such that

$$T_{\hat{\pi}(1)} < \ldots < T_{\hat{\pi}(n)}.$$ 

Let $\mu_n \in \mathcal{M}_1(\mathcal{I}^n)$ be the distribution of $(T_{\hat{\pi}(1)}, \ldots, T_{\hat{\pi}(n)})$.

Then $\mu = (\mu_1, \mu_2, \ldots)$ is thinning-invariant, but not exchangeable.
A more general example

\[ \mathcal{I} = [0, 1] \text{ and } \lambda = \text{Lebesgue measure.} \]
A more general example

\[ \mathcal{I} = [0, 1] \text{ and } \lambda = \text{Lebesgue measure}. \]

\[ \mathbf{M}_1^\lambda(\mathcal{X} \times \mathcal{I}) := \text{all } \alpha \in \mathbf{M}_1(\mathcal{X} \times \mathcal{I}) \text{ such that marginal on } \mathcal{I} \text{ is } \lambda \]
A more general example

\[ \mathcal{I} = [0, 1] \text{ and } \lambda = \text{Lebesgue measure.} \]

\[ \mathbf{M}_1^\lambda(\mathcal{X} \times \mathcal{I}) := \text{all } \alpha \in \mathbf{M}_1(\mathcal{X} \times \mathcal{I}) \text{ such that marginal on } \mathcal{I} \text{ is } \lambda \]

Let \((X_1, T_1), \ldots, (X_n, T_n) \in \mathcal{X} \times \mathcal{I}\) be i.i.d., \(\alpha\)-distributed.
A more general example

\[ I = [0, 1] \text{ and } \lambda = \text{Lebesgue measure}. \]

\[ M_1^\lambda(X \times I) := \text{all } \alpha \in M_1(X \times I) \text{ such that marginal on } I \text{ is } \lambda \]

Let \((X_1, T_1), \ldots, (X_n, T_n) \in X \times I\) be i.i.d., \(\alpha\)-distributed.

Almost surely, \(\exists! \hat{\pi} \in S_n \text{ s.t. } T_{\hat{\pi}(1)} < \ldots < T_{\hat{\pi}(n)}\)
A more general example

\( \mathcal{I} = [0, 1] \) and \( \lambda = \text{Lebesgue measure.} \)

\[ \mathbf{M}_1^\lambda(\mathcal{X} \times \mathcal{I}) := \text{all } \alpha \in \mathbf{M}_1(\mathcal{X} \times \mathcal{I}) \text{ such that marginal on } \mathcal{I} \text{ is } \lambda \]

Let \( (X_1, T_1), \ldots, (X_n, T_n) \in \mathcal{X} \times \mathcal{I} \) be i.i.d., \( \alpha \)-distributed.

Almost surely, \( \exists ! \; \hat{\pi} \in S_n \text{ s.t. } T_{\hat{\pi}(1)} < \ldots < T_{\hat{\pi}(n)} \)

Define \( \mathcal{M}_n(\alpha) \in \mathbf{M}_1(\mathcal{X}^n) \) distribution of \( (X_{\hat{\pi}(1)}, \ldots, X_{\hat{\pi}(n)}) \)
A more general example

\( \mathcal{I} = [0, 1] \) and \( \lambda = \text{Lebesgue measure} \).

\[ M^\lambda_1(\mathcal{X} \times \mathcal{I}) := \text{all } \alpha \in M_1(\mathcal{X} \times \mathcal{I}) \text{ such that marginal on } \mathcal{I} \text{ is } \lambda \]

Let \( (X_1, T_1), \ldots, (X_n, T_n) \in \mathcal{X} \times \mathcal{I} \) be i.i.d., \( \alpha \)-distributed.

Almost surely, \( \exists ! \hat{\pi} \in S_n \) s.t. \( T_{\hat{\pi}(1)} < \ldots < T_{\hat{\pi}(n)} \)

Define \( M_n(\alpha) \in M_1(\mathcal{X}^n) \) distribution of \( (X_{\hat{\pi}(1)}, \ldots, X_{\hat{\pi}(n)}) \)

Define \( M(\alpha) = (M_1(\alpha), M_2(\alpha), \ldots) \).

\( M(\alpha) \) thinning-invariant for every \( \alpha \in M^\lambda_1(\mathcal{X} \times \mathcal{I}) \).
E.g., recovering order statistics

If $\mathcal{X} = \mathcal{I}$ and $\alpha$ is uniform measure on diagonal $\{(t, t) : t \in \mathcal{I}\}$,
E.g., recovering order statistics

If \( \mathcal{X} = \mathcal{I} \) and \( \alpha \) is uniform measure on diagonal \( \{(t, t) : t \in \mathcal{I}\} \),
E.g., recovering order statistics

If $\mathcal{X} = \mathcal{I}$ and $\alpha$ is uniform measure on diagonal $\{(t, t) : t \in \mathcal{I}\}$,
If $\mathcal{X} = \mathcal{I}$ and $\alpha$ is uniform measure on diagonal $\{(t, t) : t \in \mathcal{I}\}$,
Main Result

**Theorem**

Suppose $\mu = (\mu_1, \mu_2, \ldots)$ is thinning-invariant. Then there is a unique probability measure $P \in M_1(M_1^\lambda(\mathcal{X} \times \mathcal{I}))$ such that, for each $n > 0$,

$$\mu_n(\cdot) = \int_{M_1^\lambda(\mathcal{X} \times \mathcal{I})} P(d\alpha) \mathcal{M}_n(\alpha)(\cdot)$$

Proof deferred ...
Mean-Field Models

Let’s start by (re)considering mean-field models.

E.g., Curie-Weiss model. \( \mathcal{X} = \{+1, -1\} \), \( H_N : \mathcal{X}^N \rightarrow \mathbb{R} \),

\[
\frac{H_N(\sigma_1, \ldots, \sigma_n)}{N} = -\frac{J}{2} \cdot \frac{\sum_{1 \leq j < k \leq N} \sigma_j \sigma_k}{\binom{N}{2}} - h \cdot \frac{\sum_{k=1}^{N} \sigma_k}{N}
\]

for some fixed \( J, h \in \mathbb{R} \).
Let’s start by (re)considering mean-field models

E.g., Curie-Weiss model. $\mathcal{X} = \{+1, -1\}$, $H_N : \mathcal{X}^N \to \mathbb{R}$,

$$H_N(\sigma_1, \ldots, \sigma_n) = -\frac{J}{2} \cdot \frac{\sum_{1 \leq j < k \leq N} \sigma_j \sigma_k}{\binom{N}{2}} - h \cdot \frac{\sum_{k=1}^{N} \sigma_k}{N}$$

for some fixed $J, h \in \mathbb{R}$.

Ising model on a complete graph $K_N$
Symmetry properties

For $\pi \in S_N$ define left-action $\pi : \mathcal{C}(\mathcal{X}^N) \to \mathcal{C}(\mathcal{X}^N)$,

$$(\pi \cdot f_N)(\sigma_1, \ldots, \sigma_N) := f_N(\sigma_{\pi(1)}, \ldots, \sigma_{\pi(N)}).$$
Symmetry properties

For $\pi \in S_N$ define left-action $\pi : \mathcal{C}(\mathcal{X}^N) \to \mathcal{C}(\mathcal{X}^N)$,

$$(\pi \cdot f_N)(\sigma_1, \ldots, \sigma_N) := f_N(\sigma_{\pi(1)}, \ldots, \sigma_{\pi(N)}).$$

Define $\Theta^N_{N-1} : \mathcal{C}(\mathcal{X}^{N-1}) \to \mathcal{C}(\mathcal{X}^N)$ by

$$(\Theta^N_{N-1}f_{N-1})(\sigma_1, \ldots, \sigma_N) := \frac{1}{N} \sum_{k=1}^N f_{N-1}(\sigma_{\phi_{N,k}(1)}, \ldots, \sigma_{\phi_{N,k}(N-1)}).$$
1. De Finetti's Theorem
2. Thinning-invariance
3. Asymmetric Mean-Field Models

Symmetry properties

For $\pi \in S_N$ define left-action $\pi : \mathcal{C}(\mathcal{X}^N) \to \mathcal{C}(\mathcal{X}^N)$,

$$(\pi \cdot f_N)(\sigma_1, \ldots, \sigma_N) := f_N(\sigma_{\pi(1)}, \ldots, \sigma_{\pi(N)}).$$

Define $\Theta^N_{N-1} : \mathcal{C}(\mathcal{X}^{N-1}) \to \mathcal{C}(\mathcal{X}^N)$ by

$$(\Theta^N_{N-1} f_{N-1})(\sigma_1, \ldots, \sigma_N) := \frac{1}{N} \sum_{k=1}^{N} f_{N-1}(\sigma_{\phi_N,k}(1), \ldots, \sigma_{\phi_N,k}(N-1)).$$

Note: $\langle \mu_N \cdot \pi, f_N \rangle = \langle \mu_N, \pi \cdot f_N \rangle$,

$\langle \mu_N \Theta^N_{N-1}, f_{N-1} \rangle = \langle \mu_N, \Theta^N_{N-1} f_{N-1} \rangle$

where $\langle \mu_N, f_N \rangle := \mathbf{E}^{\mu_N}[f_N]$. 

Shannon Starr
A Thinning Analogue of de Finetti's Theorem
Symmetry properties

For $\pi \in S_N$ define left-action $\pi : \mathcal{C}(\mathcal{X}^N) \to \mathcal{C}(\mathcal{X}^N)$,

$$(\pi \cdot f_N)(\sigma_1, \ldots, \sigma_N) := f_N(\sigma_{\pi(1)}, \ldots, \sigma_{\pi(N)}) .$$

Define $\Theta^N_{N-1} : \mathcal{C}(\mathcal{X}^{N-1}) \to \mathcal{C}(\mathcal{X}^N)$ by

$$(\Theta^N_{N-1}f_{N-1})(\sigma_1, \ldots, \sigma_N) := \frac{1}{N} \sum_{k=1}^{N} f_{N-1}(\sigma_{\phi_N,k(1)}, \ldots, \sigma_{\phi_N,k(N-1)})$$

Note: $\langle \mu_N \cdot \pi, f_N \rangle = \langle \mu_N, \pi \cdot f_N \rangle$,
$\langle \mu_N \Theta^N_{N-1}, f_{N-1} \rangle = \langle \mu_N, \Theta^N_{N-1}f_{N-1} \rangle$

where $\langle \mu_N, f_N \rangle := E_{\mu_N}[f_N]$.

$\forall \pi \in S_N, \pi \cdot H_N = H_N$. Finite-exchangeability.
$\forall N > 2, \Theta^N_{N-1}H_{N-1} = [(N-1)/N]H_N$. Thinning-invariance.
Thermodynamic Quantities

Let $\nu_0 \in M_1(\mathcal{X})$ be a-priori measure, e.g. uniform.

Partition function

$$Z_N(\beta) = \int_{\mathcal{X}^N} e^{-\beta H_N} \, d\nu_0 \otimes N$$

Free energy

$$F_N(\beta) = -\frac{1}{\beta} \log(Z_N(\beta))$$

“Pressure”

$$p_N(\beta) = \frac{1}{N} \log(Z_N(\beta)).$$

Boltzmann-Gibbs measure $\mu^*_{N,\beta} \in M_1(\mathcal{X}^N)$ s.t.

$$\frac{d\mu^*_{N,\beta}}{d\nu_0 \otimes N} = Z_N(\beta)^{-1} e^{-\beta H_N}.$$
Gibbs Variational Principle

For $\mu_N \in M_1(\mathcal{X}^N)$, define

$$G_N(\mu_N; \beta) := \frac{1}{N} \left( S_N(\mu_N|\nu_0^\otimes N) - \beta E^{\mu_N}[H_N] \right)$$

where $S_N(\mu_N|\nu_N)$ is relative entropy

$$S_N(\mu_N|\nu_N) = - \int_{\mathcal{X}^N} \log \left( \frac{d\mu_N}{d\nu_N} \right) d\mu_N$$

and $E^{\mu_N}[\cdot]$ is expectation.
Gibbs Variational Principle

For $\mu_N \in \mathbf{M}_1(\mathcal{X}^N)$, define

$$G_N(\mu_N; \beta) := \frac{1}{N} \left( S_N(\mu_N|\nu_0^\otimes) - \beta E^{\mu_N}[H_N] \right)$$

where $S_N(\mu_N|\nu_N)$ is relative entropy

$$S_N(\mu_N|\nu_N) = -\int_{\mathcal{X}^N} \log \left( \frac{d\mu_N}{d\nu_N} \right) d\mu_N$$

and $E^{\mu_N}[\cdot]$ is expectation.

$$p_N(\beta) = \max_{\mu_N} G_N(\mu_N; \beta)$$

and the unique argmax is Boltzmann-Gibbs distribution $\mu^*_N,\beta$.
Consequences of symmetry

\[ p_N(\beta) = \max_{\mu_N} G_N(\mu_N; \beta) = G_N(\mu^*_N, \beta; \beta), \]

\[ \frac{d\mu^*_{N,\beta}}{d\nu_0}(\cdot) = Z_N(\beta)^{-1} e^{-\beta H_N(\cdot)} \]
Consequences of symmetry

\[ p_N(\beta) = \max_{\mu_N} G_N(\mu_N; \beta) = G_N(\mu^*_N, \beta; \beta), \]

\[ \frac{d\mu^*_N, \beta}{d\nu_0^N} (\cdot) = Z_N(\beta)^{-1} e^{-\beta H_N(\cdot)} \]

Since \( \pi \cdot H_N = H_N \), know \( \mu^*_N, \beta = \mu^*_N, \beta \cdot \pi \).

So, defining \( M^e_1(\mathcal{X}^N) \) the finite exchangeable measures,

\[ p_N(\beta) = \max_{\mu_N \in M^e_1(\mathcal{X}^N)} G_N(\mu_N; \beta). \]
If we further restrict to i.i.d. product measures, we get a simpler formula.
If we further restrict to i.i.d. product measures, we get a simpler formula.

Suppose \( \nu \in \mathcal{M}_1(\mathcal{X}) \) and \( \mu_n = \nu \otimes^N \). Then

\[
S_N(\mu_N|\nu_0 \otimes^N) = S_N(\nu \otimes^N|\nu_0 \otimes^N) = N \cdot S_1(\nu|\nu_0).
\]
If we further restrict to i.i.d. product measures, we get a simpler formula.

Suppose $\nu \in M_1(\mathcal{X})$ and $\mu_n = \nu^\otimes N$. Then

$$S_N(\mu_N|\nu_0^\otimes N) = S_N(\nu^\otimes N|\nu_0^\otimes N) = N \cdot S_1(\nu|\nu_0).$$

Also,

$$E^{\nu^\otimes N}[H_N] = N \cdot \varphi(\nu),$$

where

$$\varphi(\nu) = -\frac{J}{2}(E^{\nu}[\sigma])^2 - hE^{\nu}[\sigma].$$
If we further restrict to i.i.d. product measures, we get a simpler formula. Suppose \( \nu \in \mathcal{M}_1(\mathcal{X}) \) and \( \mu_n = \nu^\otimes N \). Then

\[
S_N(\mu_N|\nu^\otimes N) = S_N(\nu^\otimes N|\nu^\otimes N) = N \cdot S_1(\nu|\nu_0).
\]

Also,

\[
E_{\nu^\otimes N}[H_N] = N \cdot \varphi(\nu),
\]

where

\[
\varphi(\nu) = -\frac{J}{2}(E_{\nu}[\sigma])^2 - hE_{\nu}[\sigma].
\]

So \( G_N(\nu^\otimes N; \beta) = S_1(\nu|\nu_0) - \beta \varphi(\nu) \).
Thermodynamic Limit: One method of solution

“Gap equation”

\[ p(\beta) := \lim_{N \to \infty} p_N(\beta) = \max_{\nu \in M_1(\mathcal{X})} \left( S_1(\nu|\nu_0) - \varphi(\nu) \right). \]
Thermodynamic Limit: One method of solution

"Gap equation"

\[ p(\beta) := \lim_{N \to \infty} p_N(\beta) = \max_{\nu \in M_1(\mathcal{X})} \left( S_1(\nu | \nu_0) - \varphi(\nu) \right). \]

Heuristic idea: Given a convex function on a simplex, it attains its maximum value at an extreme point.
Thermodynamic Limit : One method of solution


“Gap equation”

\[ p(\beta) := \lim_{N \to \infty} p_N(\beta) = \max_{\nu \in M_1(X)^\infty} \left( S_1(\nu|\nu_0) - \varphi(\nu) \right). \]

Heuristic idea: Given a convex function on a simplex, it attains its maximum value at an extreme point.

For any finite \( N \), \( G_N(\cdot; \beta) \) is concave not convex, so cannot restrict to extreme points.
Thermodynamic Limit: One method of solution

“Gap equation”

\[
p(\beta) := \lim_{N \to \infty} p_N(\beta) = \max_{\nu \in \mathcal{M}_1(\mathcal{X})} \left( S_1(\nu|\nu_0) - \varphi(\nu) \right).
\]

Heuristic idea: Given a convex function on a simplex, it attains its maximum value at an extreme point.

For any finite \( N \), \( G_N(\cdot; \beta) \) is concave not convex, so cannot restrict to extreme points.

But the nonlinear part, entropy, is “almost convex”.

\( N \to \infty \) limit of \( G_N(\cdot; \beta) \) is linear/affine, when restricted to infinitely exchangeable measures and extreme points are product measures.
Consider a sequence of Hamiltonians $H_N : \mathcal{X}^N \to \mathbb{R}$, for $N \geq n$, s.t.

$$\frac{1}{N-1} \Theta^N_{N-1} H_{N-1} = \frac{1}{N} H_N$$

for all $N > n$. But do not assume finite exchangeability.
Consider a sequence of Hamiltonians $H_N : \mathcal{X}^N \rightarrow \mathbb{R}$, for $N \geq n$, s.t.

$$\frac{1}{N-1} \Theta_{N-1}^N H_{N-1} = \frac{1}{N} H_N$$

for all $N > n$. But do not assume finite exchangeability.

Similar result to Fannes, Spohn and Verbeure’s holds, adapted to simplex of thinning-invariant measures

$$p(\beta) = \max_{\alpha \in M_1^\lambda(\mathcal{X} \times I)} \left( S(\alpha|\nu_0 \otimes \lambda) - \beta \mathbb{E}^{M_n(\alpha)}[H_n/n] \right)$$
An inversion of $\pi \in S_n$ is a pair $i < j$ s.t. $\pi(i) > \pi(j)$.

\[
\begin{array}{cccccc}
1 & 2 & 3 & 4 & 5 \\
\downarrow & \downarrow & \downarrow & \downarrow & \downarrow \\
1 & 2 & 3 & 4 & 5
\end{array}
\]
Example: Mallows Model

An inversion of $\pi \in S_n$ is a pair $i < j$ s.t. $\pi(i) > \pi(j)$.

For $q > 0$ Mallows model of random permutations

$$P_{n,q}(\{\pi\}) = Z_n(q)^{-1} q^{\#\text{inv}(\pi)},$$

$\pi \in S_n$. 
Hamiltonian

\[ \mathcal{X} = \mathcal{I} = [0, 1]; \quad \nu_0 = \lambda; \]

\[ H_N(x_1, \ldots, x_N) = N \cdot \frac{\sum_{1 \leq i < j \leq N} 1_{x_i > x_j}}{\binom{N}{2}}. \]
1. De Finetti’s Theorem
2. Thinning-invariance
3. Asymmetric Mean-Field Models

Hamiltonian

\[ \mathcal{X} = \mathcal{I} = [0, 1]; \nu_0 = \lambda; \]

\[ H_N(x_1, \ldots, x_N) = N \cdot \frac{\sum_{1 \leq i < j \leq N} 1_{x_i > x_j}}{\binom{N}{2}}. \]

Boltzmann-Gibbs measure \( \mu_{N, \beta}^* \in M_1(\mathcal{X}^N) \) concentrates on subset s.t. all \( x_1, \ldots, x_N \) distinct.
Hamiltonian

\[ \mathcal{X} = \mathcal{I} = [0, 1]; \quad \nu_0 = \lambda; \]

\[ H_N(x_1, \ldots, x_N) = N \cdot \frac{\sum_{1 \leq i < j \leq N} 1_{x_i > x_j}}{\binom{N}{2}}. \]

Boltzmann-Gibbs measure \( \mu^*_{N, \beta} \in \mathcal{M}_1(\mathcal{X}^N) \) concentrates on subset s.t. all \( x_1, \ldots, x_N \) distinct.

Marginal on the ordering permutation \( \hat{\pi} \) s.t.

\[ x_{\hat{\pi}(1)} < \ldots < x_{\hat{\pi}(N)}, \]

is \( P_{N,q} \) for \( q = e^{-\beta/N} \). Weak asymmetry.
Solving the “Gap” equation

\[ \alpha \in M_1^\lambda (X \times I), \]
\[ \alpha (dx \otimes dt) = f(t, x) \, dx \, dt, \]

satisfies, both marginals equal to \( \lambda \) and

\[ \frac{\partial^2}{\partial t \partial x} \log (f(t, x)) = 2\beta f(t, x). \]

Hyperbolic 2D Liouville’s equation. Well-known PDE. Integrable. Arises when looking for 2D metric of constant negative curvature.
Solving the “Gap” equation

\[ \alpha \in \mathcal{M}^\lambda_1(\mathcal{X} \times \mathcal{I}), \]

\[ \alpha(dx \otimes dt) = f(t, x) \, dx \, dt, \]

satisfies, both marginals equal to \( \lambda \) and

\[ \frac{\partial^2}{\partial t \partial x} \log(f(t, x)) = 2\beta f(t, x). \]

Hyperbolic 2D Liouville’s equation. Well-known PDE. Integrable. Arises when looking for 2D metric of constant negative curvature.

Unique solution with these boundary conditions.
Conclusion for Mallows model

Let $\hat{\pi} \in S_N$ be distributed by Mallows model $P_{N,q}$ for $q = e^{-\beta/N}$. 
Conclusion for Mallows model

Let $\hat{\pi} \in S_N$ be distributed by Mallows model $P_{N,q}$ for $q = e^{-\beta/N}$.

Define random “empirical” measure

$$\hat{\alpha}_{N,\beta} = \frac{1}{N} \sum_{i=1}^{N} \delta \left( \frac{i}{N}, \frac{\hat{\pi}(i)}{N} \right).$$
Conclusion for Mallows model

Let \( \hat{\pi} \in S_N \) be distributed by Mallows model \( P_{N,q} \) for \( q = e^{-\beta/N} \).

Define random “empirical” measure

\[
\hat{\alpha}_{N,\beta} = \frac{1}{N} \sum_{i=1}^{N} \delta\left( \frac{i}{N}, \frac{\hat{\pi}(i)}{N} \right).
\]

Then in weak topology, \( \hat{\alpha}_{N,\beta}(dx \otimes dt) \Rightarrow f_\beta(t, x) \, dx \, dt \), where:

\[
f_\beta(t, x) = \frac{2\beta \sinh(\beta/2)}{\left[ 2e^{\beta/4} \cosh \left( \frac{\beta}{2} [x - t] \right) - 2e^{-\beta/4} \cosh \left( \frac{\beta}{2} [x + t - 1] \right) \right]^2}
\]