(Combinatorial) Characterization of Exponential Families of Lumpable Stochastic Matrices

Geoffrey WOLFER

Tokyo University of Agriculture and Technology / RIKEN AIP

Shun WATANABE

Dec. 18th, 2025

MaxEnt 2025 — University of Auckland, 14–19 Dec. 2025

Outline

1. **Exponential Families of Stochastic Matrices**

   irreducible Markov chains, $\mathfrak{s}$-normalization, maximum entropy principle

2. **Lumpings & Embeddings**

   definition, characterization, operational interpretation

3. **Geometry of Lumpability**

   canonical embedding, foliation of the lumpable family, characterization of e-families

Exponential Families of Stochastic Matrices

Irreducible Markov chains

- Finite space $\mathcal{Y}$. Probability distributions $\mathcal{P}(\mathcal{Y})$.
- $\mathcal{E} \subset \mathcal{Y}^2$ such that the digraph $(\mathcal{Y}, \mathcal{E})$ is **strongly connected**.
- Functions and positive functions over $\mathcal{E}$: $\quad \mathcal{F}(\mathcal{Y}, \mathcal{E}), \quad \mathcal{F}^+(\mathcal{Y}, \mathcal{E})$.
- **Irreducible** row-stochastic matrices over $(\mathcal{Y}, \mathcal{E})$: $\quad \mathcal{W}(\mathcal{Y}, \mathcal{E})$.

  Discrete-time, time-homogeneous Markov chain

$$\mathbb{P}\left(Y_t = y_t \middle| Y_1 = y_1, \dots, Y_{t-1} = y_{t-1}\right) = \mathbb{P}\left(Y_t = y_t \middle| Y_{t-1} = y_{t-1}\right)$$

$$\mathbb{P}_\mu\left(Y_1 = y_1, \dots, Y_n = y_n\right) = \mu(y_1) \prod_{t=1}^{n-1} P(y_t, y_{t+1}),$$

$$(\mu, P) \in (\mathcal{P}(\mathcal{Y}), \mathcal{W}(\mathcal{Y}, \mathcal{E})).$$

- **Stationary distribution**: $\pi P = \pi$.
- **Edge-measure**: $Q(y, y') = \pi(y) P(y, y') = \mathbb{P}_\pi\left(Y_t = y, Y_{t+1} = y'\right)$.

▷ **Probability distributions.**

$$u > 0, \qquad \frac{u}{\|u\|_1}$$

▷ **Stochastic matrices.**

Definition. $\mathfrak{s}$-normalization (Miller, 1961).

When $(\mathcal{Y}, \mathcal{E})$ is strongly connected we define the mapping

$$\mathfrak{s} \colon \mathcal{F}^+(\mathcal{Y}, \mathcal{E}) \to \mathcal{W}(\mathcal{Y}, \mathcal{E})$$

$$\mathfrak{s}(F)(y, y') = \frac{F(y, y') v_F(y')}{\rho_F v_F(y)},$$

with ρ_F and v_F the Perron–Frobenius (PF) root and associated right eigenvector of F.

Maximum Entropy Principle for Markov chains (Csiszár et al., 1987)

Polytope (m-family) generated by the set of linear constraints $\{g_i = c_i\}$,

$$\mathcal{L} = \left\{ P \in \mathcal{W}(\mathcal{Y}, \mathcal{E}) \colon \sum_{(y,y') \in \mathcal{E}} Q(y, y') g_i(y, y') = c_i, \forall i \in [d] \right\} \subset \mathcal{W}(\mathcal{Y}, \mathcal{E}).$$

Let $P \in \mathcal{W}(\mathcal{Y}, \mathcal{E})$, and look at projection (information divergence rate D) onto $\mathcal{L}$,

$$P_e \triangleq \arg\min_{P' \in \mathcal{L}} D\left(P' \middle|\middle| P\right).$$

Minimizer P_e belongs to an "exponential family". For $\lambda \in \mathbb{R}^d$,

$$\tilde{P}_\lambda(y, y') = P(y, y') \exp\left(\sum_{i \in [d]} \lambda^i g_i(y, y') \right),$$

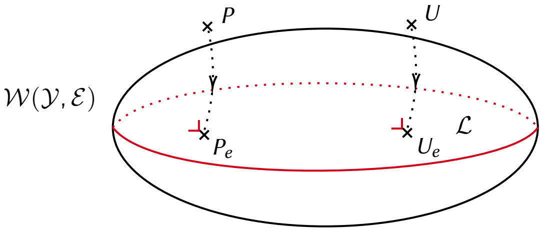
and for $\psi(\lambda)$ log-PF root of $\tilde{P}_\lambda$, $P_e = \mathfrak{s}(\tilde{P}_{\lambda_\star})$ with

$$\lambda^\star = \arg\max_{\lambda \in \mathbb{R}^d} \left\{ \lambda \cdot c - \psi(\lambda) \right\}.$$

Maximum Entropy Principle for Markov chains (Csiszár et al., 1987)

$$\text{Entropy rate: } H(P) \triangleq \lim_{k\to\infty} \frac{1}{k} H(Y_1, Y_2, \dots, Y_k)$$

$U = \mathfrak{s}(\delta_{\mathcal{E}})$, maxentropic in $\mathcal{W}(\mathcal{Y}, \mathcal{E})$, $\qquad U_e \triangleq \arg\min_{P' \in \mathcal{L}} D\left(P' || U\right)$.



$$U_e = \arg\min_{P' \in \mathcal{L}} \left\{ -H(P') - \overbrace{\mathbb{E}_{(Y,Y')\sim Q'}\left[\log U(Y, Y')\right]}^{-\log \rho(\mathfrak{s}(\delta_{\mathcal{E}}))} \right\} = \arg\max_{P' \in \mathcal{L}} H(P').$$

Exponential families of stochastic matrices

Exponential family (e-family) of stochastic matrices (Nagaoka, 2005)

$$\mathcal{V}_e = \left\{ P_\theta \colon \theta = (\theta^1, \dots, \theta^d) \in \mathbb{R}^d \right\} \subset \mathcal{W}(\mathcal{Y}, \mathcal{E}),$$

is e-family with natural parameter θ and dimension d, when there exist a function $K \in \mathcal{F}(\mathcal{Y}, \mathcal{E})$, d linearly independent functions $G_1, \dots, G_d \in \mathcal{G}(\mathcal{Y}, \mathcal{E})$, and functions $R \in \mathbb{R}^{\Theta \times \mathcal{Y}}$, $\psi \in \mathbb{R}^{\Theta}$, such that

$$\log P_\theta(y, y') = K(y, y') + \sum_{i=1}^{d} \theta^i G_i(y, y') + R(\theta, y') - R(\theta, y) - \psi(\theta),$$

$$\text{i.e.} \qquad P_\theta = \mathfrak{s} \circ \exp\left(K + \sum_{i=1}^{d} \theta^i G_i \right),$$

where $\mathcal{G}(\mathcal{Y}, \mathcal{E})$ is the quotient space

$$\mathcal{G}(\mathcal{Y}, \mathcal{E}) \triangleq \mathcal{F}(\mathcal{Y}, \mathcal{E}) / \mathcal{N}(\mathcal{Y}, \mathcal{E}),$$

$$\mathcal{N}(\mathcal{Y}, \mathcal{E}) \triangleq \left\{ N \colon \exists f, c, N(y, y') = f(y') - f(y) + c \right\}.$$

Expectation parameter (♣): $\eta_i(\theta) = \sum_{(y,y') \in \mathcal{E}} Q_\theta(y, y') G_i(y, y')$.

Exponential families of stochastic matrices

Example (Nagaoka, 2005). $\mathcal{W}(\mathcal{Y}, \mathcal{E})$ forms an e-family of dimension $|\mathcal{E}| - |\mathcal{Y}|$.

Example. Parametrization of $\mathcal{W}(\mathcal{Y}, \mathcal{Y}^2)$ proposed by Ito and Amari (1988).
With $\mathcal{Y} \cong [m]$, pick $y_\star \in \mathcal{Y}$, and write,

$$\begin{aligned}
\log P(y, y') = &\sum_{i=1, i \neq y_\star}^{m} \log \frac{P(y_\star, i) P(i, y_\star)}{P(y_\star, y_\star) P(y_\star, y_\star)} \delta_i(y') \\
&+ \sum_{i=1, i \neq y_\star}^{m} \sum_{j=1, j \neq y_\star}^{m} \log \frac{P(i, j) P(y_\star, y_\star)}{P(y_\star, j) P(i, y_\star)} \delta_i(y) \delta_j(y') \\
&+ \log P(y, y_\star) - \log P(y', y_\star) + \log P(y_\star, y_\star).
\end{aligned}$$

Basis is given by

$$\begin{aligned}
g_i &= 1^\intercal \delta_i, && i \in [m], i \neq y_\star \\
g_{ij} &= \delta_i^\intercal \delta_j, && i, j \in [m], i, j \neq y_\star
\end{aligned}$$

and the parameters are

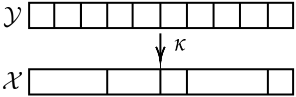
$$\theta^i = \log \frac{P(y_\star, i) P(i, y_\star)}{P(y_\star, y_\star) P(y_\star, y_\star)}, \qquad \theta^{ij} = \log \frac{P(i, j) P(y_\star, y_\star)}{P(y_\star, j) P(i, y_\star)}.$$

Lumpings & Embeddings

Notation & Lumping map

- $\mathcal{Y}$ a **finite** alphabet.
- $(\mathcal{Y}, \mathcal{E})$ be a **strongly connected digraph** with vertex set $\mathcal{Y}$ and edge set $\mathcal{E} \subset \mathcal{Y}^2$.
- $\mathcal{F}(\mathcal{Y}, \mathcal{E}) \cong \mathbb{R}^{\mathcal{E}}$, $\mathcal{F}^+(\mathcal{Y}, \mathcal{E}) \cong \mathbb{R}_+^{\mathcal{E}}$.

- $\mathcal{X}$ another **finite** alphabet with $|\mathcal{X}| \leq |\mathcal{Y}|$.
- $\kappa \colon \mathcal{Y} \to \mathcal{X}$ **surjective** map.
- $\mathcal{Y} = \biguplus_{x \in \mathcal{X}} \mathcal{S}_x$ where for $x \in \mathcal{X}$, $\mathcal{S}_x \triangleq \kappa^{-1}(\{x\})$.



- $\mathcal{D} \triangleq \kappa(\mathcal{E}) \triangleq \{(\kappa(y), \kappa(y')) \colon (y, y') \in \mathcal{E}\} \subset \mathcal{X}^2$.
- $(\mathcal{Y}, \mathcal{E})$ strongly connected $\implies$ $(\mathcal{X}, \mathcal{D})$ strongly connected.

κ maps the graph $(\mathcal{Y}, \mathcal{E})$ to the graph $(\mathcal{X}, \mathcal{D})$

Stationary Markov chain $\{Y_t\}_{t\in\mathbb{N}} \sim P \in \mathcal{W}(\mathcal{Y}, \mathcal{E})$.

Data-processing:

$$\kappa(Y_1), \kappa(Y_2), \kappa(Y_3), \ldots$$

- State space compression (storage, interpretability, ...).
- Process $\{\kappa(Y_t)\}_{t\in\mathbb{N}}$ is **not necessarily a Markov chain**, (Burke and Rosenblatt, 1958; Rogers and Pitman, 1981).
- Whenever Markovian, we say that the MC (or P) is κ-lumpable.

Characterization of κ-lumpability (Kemeny and Snell, 1983).

$P \in \mathcal{W}(\mathcal{Y}, \mathcal{E})$ is κ-lumpable if and only if when for any $(x, x') \in \mathcal{D}$ and any $y_1, y_2 \in \mathcal{S}_x$,

$$\sum_{y' \in \mathcal{S}_{x'}} P(y_1, y') = \sum_{y' \in \mathcal{S}_{x'}} P(y_2, y').$$

Lumpable family

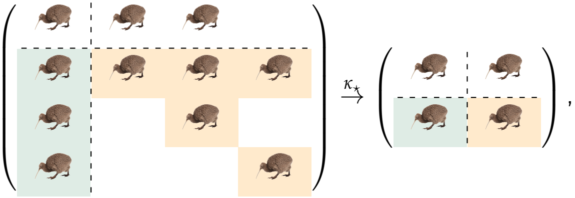
$$\text{Lumpable family: } \mathcal{W}_\kappa(\mathcal{Y}, \mathcal{E})$$
$$\kappa_\star \colon \mathcal{W}_\kappa(\mathcal{Y}, \mathcal{E}) \to \mathcal{W}(\mathcal{X}, \mathcal{D})$$

When P is κ-lumpable,

$$\kappa_\star P(x, x') \triangleq P(y, \mathcal{S}_{x'}), y \in \mathcal{S}_x.$$

Visual representation upon relabeling (support, blocks and rows).

$\mathcal{Y} = \{0, 1, 2, 3\}, \mathcal{X} = \{a, b\}, \mathcal{S}_a = \{0\}, \mathcal{S}_b = \{1, 2, 3\}.$

$\mathcal{W}_\kappa(\mathcal{Y}, \mathcal{E}) \sim$ 

Markov embeddings of Markov chains (W. and Watanabe, 2024)

Let $\kappa \colon \mathcal{Y} \to \mathcal{X}$ inducing the partition $\biguplus_{x \in \mathcal{X}} \mathcal{S}_x = \mathcal{Y}$.

Definition. κ-compatible Markov embedding.

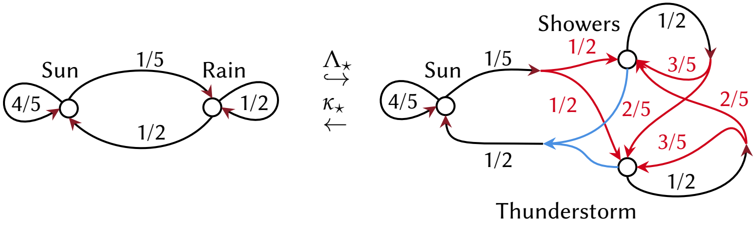
$$\Lambda_\star \colon \mathcal{W}(\mathcal{X}, \mathcal{D}) \to \mathcal{W}_\kappa(\mathcal{Y}, \mathcal{E})$$
$$\text{where } \Lambda_\star P(y, y') = P(\kappa(y), \kappa(y'))\Lambda(y, y'), \forall (y, y') \in \mathcal{E},$$

- $\Lambda \in \mathcal{F}^+(\mathcal{Y}, \mathcal{E})$.
- $\forall y \in \mathcal{Y}, x' \in \mathcal{X}, (\kappa(y), x') \in \mathcal{D} \implies (\Lambda(y, y'))_{y' \in \mathcal{S}_{x'}} \in \mathcal{P}(\mathcal{S}_{x'})$.

$$\Lambda = \left(\begin{array}{c:c:c:c:c} W_{1,1} & W_{1,2} & \cdots & & W_{1,|\mathcal{X}|} \\ \hdashline \vdots & & & & \vdots \\ \hdashline W_{x,1} & \cdots & W_{x,x'} & \cdots & W_{x,|\mathcal{X}|} \\ \hdashline \vdots & & & & \vdots \\ \hdashline W_{|\mathcal{X}|,1} & & \cdots & & W_{|\mathcal{X}|,|\mathcal{X}|} \end{array}\right).$$

Can verify: $\Lambda_\star P \in \mathcal{W}_\kappa(\mathcal{Y}, \mathcal{E})$,
$\kappa_\star \Lambda_\star P = P$ (right-inverse).

Example: weather model



$$P = \begin{pmatrix} 4/5 & 1/5 \\ 1/2 & 1/2 \end{pmatrix}, \qquad \Lambda = \left(\begin{array}{c:cc} 1 & 1/2 & 1/2 \\ \hdashline 1 & 3/5 & 2/5 \\ 1 & 2/5 & 3/5 \end{array}\right),$$

$$\Lambda_\star P = \left(\begin{array}{c:cc} 4/5 & 1/10 & 1/10 \\ \hdashline 1/2 & 3/10 & 1/5 \\ 1/2 & 1/5 & 3/10 \end{array}\right).$$

Geometry of Lumpability

Classification of established families — Markov-centric properties

Manifold	m-family	e-family	Dimension	Reference
$\mathcal{W}(\mathcal{Y}, \mathcal{E})$	○	○	$\lvert\mathcal{E}\rvert - \lvert\mathcal{Y}\rvert$	Nagaoka (2005)
$\mathcal{W}_{\mathrm{iid}}(\mathcal{Y}, \mathcal{Y}^2)$	×	○	$\lvert\mathcal{Y}\rvert - 1$	
$\mathcal{W}_{\mathrm{bis}}(\mathcal{Y}, \mathcal{Y}^2)$	○	×	$(\lvert\mathcal{Y}\rvert - 1)^2$	Hayashi and Watanabe (2016)
$\mathcal{W}_{\mathrm{rev}}(\mathcal{Y}, \mathcal{E})$	○	○	$(\lvert\mathcal{E}\rvert + \lvert\ell(\mathcal{E})\rvert)/2 - 1$	W. and Watanabe (2021)
$\mathcal{W}_{\mathrm{sym}}(\mathcal{Y}, \mathcal{Y}^2)$	○	×	$\lvert\mathcal{Y}\rvert\,(\lvert\mathcal{Y}\rvert - 1)/2$	
$\mathcal{W}_{\kappa}(\mathcal{Y}, \mathcal{E})$	?	?	?	W. and Watanabe (2024) Watanabe and W. (2024)

Observation.

$\mathcal{W}_{\kappa}(\mathcal{Y}, \mathcal{E})$ is generally neither an e-family, nor an m-family.

...but we can decompose the family into simpler structures.

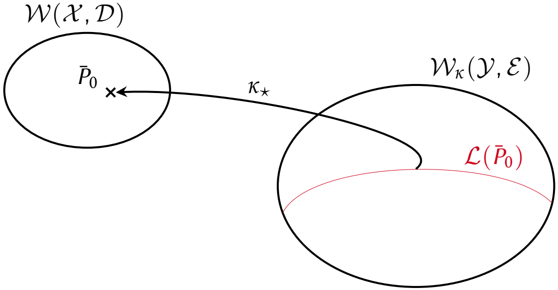
Let $\bar{P}_0 \in \mathcal{W}(\mathcal{X}, \mathcal{D})$, and

$$\mathcal{L}(\bar{P}_0) \triangleq \{P \in \mathcal{W}_\kappa(\mathcal{Y}, \mathcal{E}) : \kappa_\star P = \bar{P}_0\},$$

Lemma.

$\mathcal{L}(\bar{P}_0)$ forms an m-family in $\mathcal{W}(\mathcal{Y}, \mathcal{E})$, with

$$\dim \mathcal{L}(\bar{P}_0) = |\mathcal{E}| - \sum_{(x,x') \in \mathcal{D}} |\mathcal{S}_x|.$$



Exponential family of embedding at some prescribed origin $P_\odot$

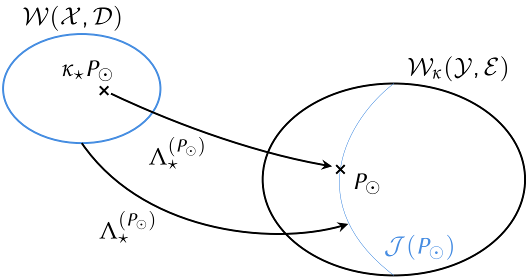
Definition. Canonical embedding.

Let $P \in \mathcal{W}_\kappa(\mathcal{Y}, \mathcal{E})$. There exists a unique $\Lambda_\star^{(P)} \colon \mathcal{W}(\mathcal{X}, \mathcal{D}) \to \mathcal{W}_\kappa(\mathcal{Y}, \mathcal{E})$ satisfying

$$P = \Lambda_\star^{(P)} \kappa_\star P.$$

Let $P_\odot \in \mathcal{W}_\kappa(\mathcal{Y}, \mathcal{E})$, and $\Lambda_\star^{(P_\odot)}$ canonical embedding.

$$\mathcal{J}(P_\odot) \triangleq \left\{ \Lambda_\star^{(P_\odot)} \bar{P} \colon \bar{P} \in \mathcal{W}(\mathcal{X}, \mathcal{D}) \right\} \subset \mathcal{W}_\kappa(\mathcal{Y}, \mathcal{E}).$$

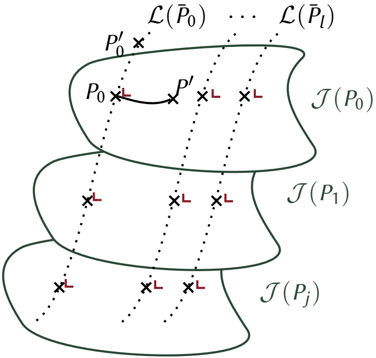


Lemma.

$\mathcal{J}(P_\odot)$ forms an e-family in $\mathcal{W}(\mathcal{Y}, \mathcal{E})$, with

$$\dim \mathcal{J}(P_\odot) = |\mathcal{D}| - |\mathcal{X}|.$$

Foliated manifold of lumpable stochastic matrices



Theorem.

For any fixed $\bar{P}_0 \in \mathcal{W}(\mathcal{X}, \mathcal{D})$,

$$\mathcal{W}_\kappa(\mathcal{Y}, \mathcal{E}) = \biguplus_{P \in \mathcal{L}(\bar{P}_0)} \mathcal{J}(P).$$

$$\dim \mathcal{W}_\kappa(\mathcal{Y}, \mathcal{E}) = |\mathcal{E}| - \sum_{(x,x') \in \mathcal{D}} |\mathcal{S}_x| + |\mathcal{D}| - |\mathcal{X}| .$$

Characterization problem.

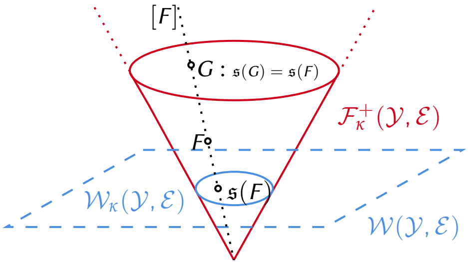
When does $\mathcal{W}_\kappa(\mathcal{Y}, \mathcal{E})$ form an e-family?

single leaf only? conditions on κ and $(\mathcal{Y}, \mathcal{E})$? algorithmic considerations?

Lumpable cone

▷ **To analyze the properties of $\mathcal{W}_\kappa(\mathcal{Y},\mathcal{E})$, we look at the relaxation $\mathcal{F}_\kappa^+(\mathcal{Y},\mathcal{E})$.**

$$\mathcal{F}_\kappa(\mathcal{Y},\mathcal{E}) \triangleq \{\text{lumpable functions (not nec. stochastic)}\},$$
$$\mathcal{F}_\kappa^+(\mathcal{Y},\mathcal{E}) \triangleq \{\text{positive lumpable cone}\}.$$



Commutativity.

$\mathfrak{s}$-normalization preserves κ-lumpability and

$$\begin{tikzcd}
\mathcal{F}_\kappa^+(\mathcal{Y},\mathcal{E}) \arrow[r, "\kappa_\star"] \arrow[d, "\mathfrak{s}"'] & \mathcal{F}^+(\mathcal{X},\mathcal{D}) \arrow[d, "\mathfrak{s}"] \\
\mathcal{W}_\kappa(\mathcal{Y},\mathcal{E}) \arrow[r, "\kappa_\star"] & \mathcal{W}(\mathcal{X},\mathcal{D}).
\end{tikzcd}$$

Merging blocks

- **To analyze the properties of $\mathcal{F}_\kappa^+(\mathcal{Y},\mathcal{E})$, we will look at the structure of the blocks.**
- **Merging rows in block $(x,x') \in \mathcal{D}$:**

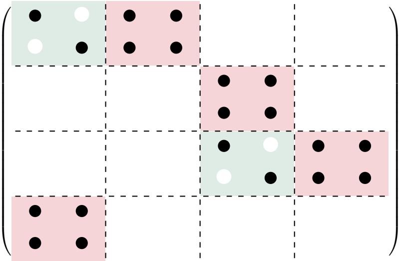
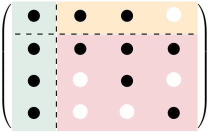
$$\mathcal{M}_{x,x'} \triangleq \left\{ y \in \mathcal{S}_x \colon \exists y_1', y_2' \in \mathcal{S}_{x'}, y_1' \neq y_2', (y,y_1'), (y,y_2') \in \mathcal{E} \right\}.$$

- **Merging blocks:**

$$\mathcal{M} \triangleq \left\{ (x,x') \in \mathcal{D} \colon \mathcal{M}_{x,x'} \neq \emptyset \right\}.$$

- **Multi-row merging blocks:**

$$\mathcal{M}_{>1} \triangleq \left\{ (x,x') \in \mathcal{M}, |\mathcal{S}_x| > 1 \right\}.$$

Examples. $\mathcal{W}_\kappa(\mathcal{Y},\mathcal{E}) \sim$  , $\mathcal{W}_\kappa(\mathcal{Y},\mathcal{E}) \sim$  ,

Lumpable cone

Definition.

$$\mathcal{G}_\kappa(\mathcal{Y}, \mathcal{E}) \triangleq \left\{ \log F \colon F \in \mathcal{F}_\kappa^+(\mathcal{Y}, \mathcal{E}) \right\}$$

Theorem. Characterization of log-affinity of the cone $\mathcal{F}_\kappa^+(\mathcal{Y}, \mathcal{E})$.

The two following statements are equivalent.

(i) $(\mathcal{Y}, \mathcal{E})$ has **no multi-row merging block** with respect to κ.

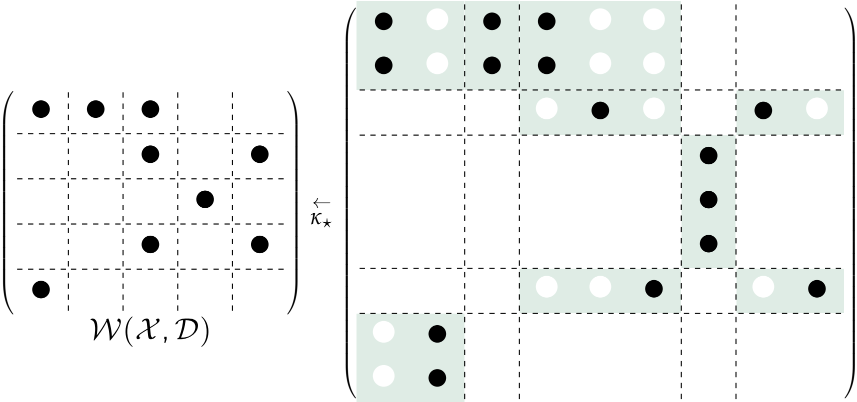
(ii) $\mathcal{G}_\kappa(\mathcal{Y}, \mathcal{E})$ is an **affine space**.

Corollary. No multi-row merging block criterion.

If $(\mathcal{Y}, \mathcal{E})$ has **no multi-row merging block**, then $\mathcal{W}_\kappa(\mathcal{Y}, \mathcal{E})$ **is e-family**.

Example: Hudson expansion (Kemeny and Snell, 1983)

$$\{X_t\}_t \sim P \in \mathcal{W}(\mathcal{X}, \mathcal{D}) \qquad \{(X_t, X_{t+1})\}_t \sim ?$$



$\mathcal{W}(\mathcal{X}, \mathcal{D})$

$H_\star \mathcal{W}(\mathcal{X}, \mathcal{D}) = \mathcal{W}_\kappa(\mathcal{Y} = \mathcal{D}, \mathcal{E})$

$$\mathcal{E} = \left\{(e = (x_1, x_2), e' = (x_1', x_2')) \in \mathcal{D}^2 \colon x_2 = x_1'\right\}, \ \kappa \colon \mathcal{Y} \to \mathcal{X}, (x_1, x_2) \mapsto x_2$$

No merging row $\implies$ $\mathcal{W}_\kappa(\mathcal{Y}, \mathcal{E})$ forms an e-family.

Monotonicity & Lazy cycle criterion

Theorem. Monotonicity.

Let $\mathcal{E} \subset \mathcal{E}'$. If $\mathcal{W}_\kappa(\mathcal{Y}, \mathcal{E}')$ forms an e-family, then $\mathcal{W}_\kappa(\mathcal{Y}, \mathcal{E})$ forms an e-family.

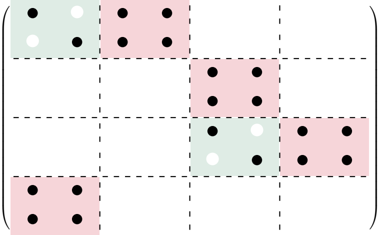
What happens if there is a multi-row merging block?

Theorem. Lazy cycle criterion.

If for any $P \in \mathcal{W}_\kappa(\mathcal{Y}, \mathcal{E})$,

$$P = D + \Pi,$$

$\kappa_\star \Pi$ is a permutation matrix over $\mathcal{X}$ and D is diagonal, then $\mathcal{W}_\kappa(\mathcal{Y}, \mathcal{E})$ forms an e-family.



- There can be **e-families with multi-row merging blocks**.
- There can be e-families with an **arbitrarily large number** of multi-row merging blocks.

Redundant merging block criterion

Theorem. Redundant merging block criterion.

If $(\mathcal{Y}, \mathcal{E})$ has a **redundant multi-row merging block**, then $\mathcal{W}_\kappa(\mathcal{Y}, \mathcal{E})$ is **not e-family**.

$$\mathcal{Y} = \{0, 1, 2, 3, 4, 5\}, \mathcal{X} = \{a, b, c\}, \mathcal{S}_a = \{0, 1\}, \mathcal{S}_b = \{2, 3\}, \mathcal{S}_c = \{4, 5\},$$

$$\mathcal{W}_\kappa(\mathcal{Y}, \mathcal{E}) \sim \left(\begin{array}{cc|cc|cc} \bullet & \circ & \bullet & \bullet & \bullet & \circ \\ \circ & \bullet & \bullet & \bullet & \circ & \bullet \\ \hline \bullet & \circ & \bullet & \circ & \bullet & \bullet \\ \circ & \bullet & \circ & \bullet & \bullet & \bullet \\ \hline \bullet & \bullet & \bullet & \circ & \bullet & \circ \\ \bullet & \bullet & \circ & \bullet & \circ & \bullet \end{array}\right).$$

Remove merging block (b, c). Closed path exhausting $\mathcal{Y}$:

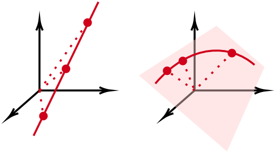
$$1 \to 2 \to 0 \to 4 \to 2 \to 0 \to 3 \to 1 \to 5 \to 3 \to 1,$$

$$\implies \mathcal{W}_\kappa(\mathcal{Y}, \mathcal{E}) \text{ \underline{not} e-family}.$$

Dimensional criterion

▷ **Recall class:**

$$\mathcal{G}_\kappa(\mathcal{Y},\mathcal{E}) \triangleq \left\{\log F\colon F \in \mathcal{F}_\kappa^+(\mathcal{Y},\mathcal{E})\right\}.$$



▷ **Affine hull in $\mathcal{F}(\mathcal{Y},\mathcal{E})$:**

$$\operatorname{aff}(\mathcal{G}_\kappa(\mathcal{Y},\mathcal{E})) \triangleq \left\{\sum_{i=1}^{k} \alpha_i G_i\colon k \in \mathbb{N}, \alpha \in \mathbb{R}^k, \sum_{i=1}^{k} \alpha_i = 1, G_1,\ldots,G_k \in \mathcal{G}_\kappa(\mathcal{Y},\mathcal{E})\right\}.$$

▷ **Recall:** $\operatorname{aff}(\mathcal{G}_\kappa(\mathcal{Y},\mathcal{E})) = \mathcal{G}_\kappa(\mathcal{Y},\mathcal{E})$ iff no multi-row merging block.

Lemma. Linear space, basis and dimension.

$$0_{\mathcal{F}(\mathcal{Y},\mathcal{E})} \in \operatorname{aff}(\mathcal{G}_\kappa(\mathcal{Y},\mathcal{E})).$$

$$\dim \operatorname{aff}(\mathcal{G}_\kappa(\mathcal{Y},\mathcal{E})) = \text{ Can compute.}$$

Can construct a basis for $\operatorname{aff}(\mathcal{G}_\kappa(\mathcal{Y},\mathcal{E}))$.

Dimensional criterion

Theorem. Dimensional criterion.

$\mathcal{W}_\kappa(\mathcal{Y},\mathcal{E})$ forms an **e-family** if and only if

$$\dim\left(\operatorname{aff}\left(\mathcal{G}_\kappa(\mathcal{Y},\mathcal{E})\right)\oplus\mathcal{N}(\mathcal{Y},\mathcal{E})\right)=|\mathcal{Y}|+\underbrace{|\mathcal{E}|+|\mathcal{D}|-|\mathcal{X}|-\sum_{(x,x')\in\mathcal{D}}|\mathcal{S}_x|}_{=\dim\mathcal{W}_\kappa(\mathcal{Y},\mathcal{E})}.$$

Algorithm.

- 1. Construct a basis for $\operatorname{aff}(\mathcal{G}_\kappa(\mathcal{Y},\mathcal{E}))$ (skipped in this talk).
- 2. Construct a basis for $\mathcal{N}(\mathcal{Y},\mathcal{E})=\{N\colon N(y,y')=f(y')-f(y)+c\}$.
- 3. Concatenate the two bases for $\operatorname{aff}(\mathcal{G}_\kappa(\mathcal{Y},\mathcal{E}))$ and $\mathcal{N}(\mathcal{Y},\mathcal{E})$.
- 4. Compute the rank of the system (flatten into vectors of dimension $|\mathcal{E}|$).

Dimensional criterion — Example

$$\mathcal{W}_\kappa(\mathcal{Y},\mathcal{E}) \sim \left(\begin{array}{c|ccc} \bullet & \bullet & \bullet & \\ \hline \bullet & \bullet & \bullet & \bullet \\ \bullet & & \bullet & \\ \bullet & & & \bullet \end{array}\right),$$

- ▷ The only multi-row merging block is **not redundant**.
- ▷ Basis for $\operatorname{aff}(\mathcal{G}_\kappa(\mathcal{Y},\mathcal{E}))$.

$$\left(\begin{array}{c|ccc} 1 & & & \\ \hline & & & \\ & & & \\ & & & \end{array}\right), \left(\begin{array}{c|ccc} & 1 & & \\ \hline & & & \\ & & & \\ & & & \end{array}\right), \left(\begin{array}{c|ccc} & & 1 & \\ \hline & & & \\ & & & \\ & & & \end{array}\right), \left(\begin{array}{c|ccc} & & & \\ \hline 1 & & & \\ 1 & & & \\ 1 & & & \end{array}\right),$$

$$\left(\begin{array}{c|ccc} & & & \\ \hline & 1 & & \\ & & 1 & \\ & & & 1 \end{array}\right), \left(\begin{array}{c|ccc} & & & \\ \hline & 1 & & \\ & & & \\ & & & \end{array}\right), \left(\begin{array}{c|ccc} & & & \\ \hline & & 1 & \\ & & & \\ & & & \end{array}\right), \left(\begin{array}{c|ccc} & & & \\ \hline & & & 1 \\ & & & \\ & & & \end{array}\right).$$

$$\mathcal{W}_\kappa(\mathcal{Y},\mathcal{E}) \sim \left(\begin{array}{c:ccc} \bullet & \bullet & \bullet & \\ \hdashline \bullet & \bullet & \bullet & \bullet \\ \bullet & & \bullet & \\ \bullet & & & \bullet \end{array}\right),$$

▷ Basis for $\mathcal{N}(\mathcal{Y},\mathcal{E})$

$$\left(\begin{array}{c:ccc} 1 & 1 & 1 & \\ \hdashline 1 & 1 & 1 & 1 \\ 1 & & 1 & \\ 1 & & & 1 \end{array}\right), \left(\begin{array}{c:ccc} & 1 & & \\ \hdashline -1 & & -1 & -1 \\ & & & \\ & & & \end{array}\right), \left(\begin{array}{c:ccc} & & 1 & \\ \hdashline & & 1 & \\ -1 & & & \\ & & & \end{array}\right), \left(\begin{array}{c:ccc} & & & \\ \hdashline & & & 1 \\ & & & \\ -1 & & & \end{array}\right).$$

$$\dim \mathcal{W}_\kappa(\mathcal{Y},\mathcal{E}) = |\mathcal{E}| + |\mathcal{D}| - |\mathcal{X}| - \sum_{(x,x')\in\mathcal{D}} |\mathcal{S}_x| = 11 + 4 - 2 - 8 = 5$$

$$\dim\left(\mathcal{N}(\mathcal{Y},\mathcal{E})\right) = |\mathcal{Y}| = 4$$

$$\dim\left(\mathrm{aff}\left(\mathcal{G}_\kappa(\mathcal{Y},\mathcal{E})\right) \oplus \mathcal{N}(\mathcal{Y},\mathcal{E})\right) \overset{\text{comp.}}{=} 10 > 4 + 5.$$

$$\Longrightarrow \mathcal{W}_\kappa(\mathcal{Y},\mathcal{E}) \text{ \underline{not} e-family}.$$

Conclusion

Find Embedding

Lazy Cycle Criterion

$\mathcal{O}(|\mathcal{X}|^2 + |\mathcal{E}|)$

No Merging Block Criterion

$\mathcal{O}(|\mathcal{X}|\,|\mathcal{Y}| + |\mathcal{E}|)$

Dimension Criterion

$\mathcal{O}(|\mathcal{E}|^\omega)$

Witness Verification

$\mathcal{O}(|\mathcal{Y}|^\omega)$

$\Longrightarrow$ Corollary

$\mathcal{O}(|\mathcal{X}|\,|\mathcal{Y}| + |\mathcal{E}|)$

Redundant Block Criterion

$\mathcal{O}(|\mathcal{D}|\,(|\mathcal{Y}| + |\mathcal{E}|))$

Monotonicity

Thank you for listening.

arXiv:2412.08400

- **Characterization of Exponential Families of Lumpable Stochastic Matrices**
  S. Watanabe, G. Wolfer
  arXiv:2412.08400

References

CJ Burke and M Rosenblatt. A Markovian function of a Markov chain. *The Annals of Mathematical Statistics*, 29(4):1112–1122, 1958.

Imre Csiszár, Thomas M. Cover, and Byoung-Seon Choi. Conditional limit theorems under Markov conditioning. *IEEE Transactions on Information Theory*, 33(6):788–801, 1987.

Masahito Hayashi and Shun Watanabe. Information geometry approach to parameter estimation in Markov chains. *The Annals of Statistics*, 44(4):1495 – 1535, 2016. doi: 10.1214/15-AOS1420.

Hisashi Ito and Shun-ichi Amari. Geometry of information sources. In *Proceedings of the 11th Symposium on Information Theory and Its Applications (SITA '88)*, pages 57–60, 1988.

John G Kemeny and J Laurie Snell. *Finite Markov chains: with a new appendix" Generalization of a fundamental matrix"*. Springer, New York, 1983.

HD Miller. A convexity property in the theory of random variables defined on a finite Markov chain. *The Annals of mathematical statistics*, pages 1260–1270, 1961.

Hiroshi Nagaoka. The exponential family of Markov chains and its information geometry. In *The proceedings of the Symposium on Information Theory and Its Applications*, volume 28(2), pages 601–604, 2005.

Leonard CG Rogers and JW Pitman. Markov functions. *The Annals of Probability*, pages 573–582, 1981.

Geoffrey W. and Shun Watanabe. Information geometry of reversible Markov chains. *Information Geometry*, 4(2):393–433, 12 2021. ISSN 2511-2481. doi: 10.1007/s41884-021-00061-7.

Geoffrey W. and Shun Watanabe. Geometric aspects of data-processing of Markov chains. *Transactions of Mathematics and Its Applications*, 8(1), 05 2024. ISSN 2398-4945. doi: 10.1093/imatrm/tnae001.

Shun Watanabe and G. W. Combinatorial characterization of exponential families of lumpable stochastic matrices. *arXiv:2412.08400*, 2024.