G METHOD IN ACTION: FAST EXACT SAMPLING FROM SET OF PERMUTATIONS OF ORDER n ACCORDING TO MALLOWS MODEL THROUGH KENDALL METRIC

UDREA PĂUN

Communicated by Marius Iosifescu

Using G method, we construct a Gibbs sampler in a generalized sense on $\mathbb{S}_n$, the set of permutations of order n, which attains its stationarity at time 1, its stationary probability distribution being the Mallows model through Kendall metric (a model for ranked data). This chain leads to a fast exact (not approximate) Markovian method for sampling from $\mathbb{S}_n$ according to the Mallows model through Kendall metric. The method has $n-1$ steps because the transition matrix of chain is a product of $n-1$ (stochastic) matrices. On the other hand, for the Mallows model through Kendall metric, this chain can be used for other things — we compute the normalizing constant and, by Uniqueness Theorem, certain important probabilities.

AMS 2010 Subject Classification: 60J10, 62Dxx, 62F07, 68U20.

Key words: G method, exact sampling, Gibbs sampler in a generalized sense, Kendall metric, Mallows model, normalizing constant, important probabilities.

1. AUXILIARY RESULTS

In this section, we present the basic result from [9] we need. We then give three results on $\mathbb{S}_n$, the set of permutations of order n — the first one refers to certain components of certain permutations, the second one refers to certain systems of generators for $\mathbb{S}_n$, the last one refers to Kendall metric.

Set

$$\operatorname{Par}(E) = \{\Delta \mid \Delta \text{ is a partition of } E\},$$

where E is a nonempty set. We shall agree that the partitions do not contain the empty set.

Definition 1.1. Let $\Delta_1, \Delta_2 \in \operatorname{Par}(E)$. We say that Δ_1 *is finer than* Δ_2 if $\forall V \in \Delta_1$, $\exists W \in \Delta_2$ such that $V \subseteq W$.

Write $\Delta_1 \preceq \Delta_2$ when Δ_1 is finer than Δ_2.

In this article, a vector is a row vector and a stochastic matrix is a row stochastic matrix.

The entry (i, j) of a matrix Z will be denoted by Z_{ij} or, if confusion can arise, $Z_{i \to j}$.

Set

$$\langle m \rangle = \{1, 2, ..., m\} \ \ (m \geq 1),$$
$$N_{m,n} = \{P \mid P \text{ is a nonnegative } m \times n \text{ matrix}\},$$
$$S_{m,n} = \{P \mid P \text{ is a stochastic } m \times n \text{ matrix}\},$$
$$N_n = N_{n,n},$$
$$S_n = S_{n,n}.$$

Let $P = (P_{ij}) \in N_{m,n}$. Let $\emptyset \neq U \subseteq \langle m \rangle$ and $\emptyset \neq V \subseteq \langle n \rangle$. Set the matrices

$$P_U = (P_{ij})_{i \in U, j \in \langle n \rangle}, \ P^V = (P_{ij})_{i \in \langle m \rangle, j \in V}, \text{ and } P_U^V = (P_{ij})_{i \in U, j \in V}.$$

Set

$$(\{i\})_{i \in \{s_1, s_2, ..., s_t\}} = (\{s_1\}, \{s_2\}, ..., \{s_t\});$$
$$(\{i\})_{i \in \{s_1, s_2, ..., s_t\}} \in \operatorname{Par}(\{s_1, s_2, ..., s_t\}).$$

Definition 1.2. Let $P \in N_{m,n}$. We say that P is a *generalized stochastic matrix* if $\exists a \geq 0$, $\exists Q \in S_{m,n}$ such that $P = aQ$.

Definition 1.3 ([9]). Let $P \in N_{m,n}$. Let $\Delta \in \operatorname{Par}(\langle m \rangle)$ and $\Sigma \in \operatorname{Par}(\langle n \rangle)$. We say that P is a $[\Delta]$-*stable matrix on* Σ if P_K^L is a generalized stochastic matrix, $\forall K \in \Delta, \forall L \in \Sigma$. In particular, a $[\Delta]$-stable matrix on $(\{i\})_{i \in \langle n \rangle}$ is called $[\Delta]$-*stable* for short.

Definition 1.4 ([9]). Let $P \in N_{m,n}$. Let $\Delta \in \operatorname{Par}(\langle m \rangle)$ and $\Sigma \in \operatorname{Par}(\langle n \rangle)$. We say that P is a Δ-*stable matrix on* Σ if Δ is the least fine partition for which P is a $[\Delta]$-stable matrix on Σ. In particular, a Δ-stable matrix on $(\{i\})_{i \in \langle n \rangle}$ is called Δ-*stable* while a $(\langle m \rangle)$-stable matrix on Σ is called *stable on* Σ for short. A stable matrix on $(\{i\})_{i \in \langle n \rangle}$ is called *stable* for short.

Let $\Delta_1 \in \operatorname{Par}(\langle m \rangle)$ and $\Delta_2 \in \operatorname{Par}(\langle n \rangle)$. Set (see [9] for G_{Δ_1, Δ_2} and [10] for $\overline{G}_{\Delta_1, \Delta_2}$)

$$G_{\Delta_1, \Delta_2} = \{P \mid P \in S_{m,n} \text{ and } P \text{ is a } [\Delta_1]\text{-stable matrix on } \Delta_2\}$$

and

$$\overline{G}_{\Delta_1, \Delta_2} = \{P \mid P \in N_{m,n} \text{ and } P \text{ is a } [\Delta_1]\text{-stable matrix on } \Delta_2\}.$$

When we study or even when we construct products of nonnegative matrices (in particular, products of stochastic matrices) using G_{Δ_1,Δ_2} or $\overline{G}_{\Delta_1,\Delta_2}$, we shall refer this as the *G method.*

We now give the basic result from [9] we need.

THEOREM 1.5 ([9]). *Let* $P_1 \in G_{(\langle m_1 \rangle),\Delta_2} \subseteq S_{m_1,m_2}$, $P_2 \in G_{\Delta_2,\Delta_3} \subseteq S_{m_2,m_3}$, ..., $P_{n-1} \in G_{\Delta_{n-1},\Delta_n} \subseteq S_{m_{n-1},m_n}$, $P_n \in G_{\Delta_n,(\{i\})_{i\in\langle m_{n+1}\rangle}} \subseteq S_{m_n,m_{n+1}}$. *Then*

$$P_1P_2...P_n$$

is a stable matrix (i.e., a matrix with identical rows, see Definition 1.4).

Proof. See [9]. $\square$

Consider the group $(\mathbb{S}_n, \circ)$, where $\mathbb{S}_n$ is the set of permutations of order n $(n \geq 1)$ and $\circ$ is the usual composition of functions. $(u_1, u_2, ..., u_k)$ is a cycle of length k, where k, u_1, u_2, ..., $u_k \in \langle n \rangle$, $u_s \neq u_t, \forall s,t \in \langle k \rangle$, $s \neq t$; (u_1) is a degenerate (improper) cycle and (u_1, u_2) is a transposition. Set $(u) =\text{Id}$, $\forall u \in \langle n \rangle$, where (u) is a degenerate cycle, $\forall u \in \langle n \rangle$, and Id is the identity permutation. Set

$$C_n = \{(1)\,,\ (2,1)\,,\ ...,\ (n, n-1, ..., 1)\}\,,\ \forall n \geq 1.$$

$(C_n \subseteq \mathbb{S}_n, \forall n \geq 1.)$ C_n is a set of cycles of $\mathbb{S}_n$: (1) is a degenerate cycle ($(1) =\text{Id}$), $(2,1)$ is a cycle of length 2 (a transposition), ..., $(n, n-1, ..., 1)$ is a cycle of length n.

Let $n \geq 3$. Let $1 \leq s \leq n-2$. Let $n_1, n_2, ..., n_s \in \langle n \rangle$, $n_i \neq n_j, \forall i, j \in \langle s \rangle$, $i \neq j$. The set $\langle n \rangle - \{n_1, n_2, ..., n_s\}$ has $n-s$ elements. Denoting these elements by w_1, w_2, ..., w_{n-s}, we have

$$\langle n \rangle - \{n_1, n_2, ..., n_s\} = \{w_1, w_2, ..., w_{n-s}\}\,.$$

Suppose that $w_1 < w_2 < ... < w_{n-s}$. Set

$$C_{n,n_1,n_2,...,n_s} = \{(w_1)\,,\ (w_2, w_1)\,,\ ...,\ (w_{n-s}, w_{n-s-1}, ..., w_1)\}\,.$$

$(C_{n,n_1,n_2,...,n_s} \subseteq \mathbb{S}_n$ and $w_1 = \min\{w_1, w_2, ..., w_{n-s}\}\,.)$ $C_{n,n_1,n_2,...,n_s}$ is also a set of cycles of $\mathbb{S}_n$ ((w_1) is a degenerate cycle ($(w_1) =\text{Id}$)).

Let $\psi = (u_1, u_2, ..., u_k)$ be a (degenerate or not) cycle ($\psi \in \mathbb{S}_n$, $1 \leq k \leq n$). Set

$$g(\psi) = \max\{u_1, u_2, ..., u_k\}\,.$$

We call $g(\psi)$ the *greatest element of* ψ.

THEOREM 1.6. *Let* $n \geq 2$. *Let* $l \in \langle n-1 \rangle$. *Let* $\tau_1 \in C_n$, $\tau_2 \in C_{n,g(\tau_1)}$, ..., $\tau_l \in C_{n,g(\tau_1),g(\tau_2),...,g(\tau_{l-1})}$ ($g(\tau_1)\,,\ g(\tau_2)\,,\ ...,\ g(\tau_{l-1})$ *vanish when* $l = 1$). *Then*

$$(\tau_l \circ \tau_{l-1} \circ ... \circ \tau_1)(k) = g(\tau_k)\,,\ \forall k \in \langle l \rangle$$

(i.e., the first l *entries of* $\tau_l \circ \tau_{l-1} \circ ... \circ \tau_1$ *are* $g(\tau_1)\,,\ g(\tau_2)\,,\ ...,\ g(\tau_l)$*, respectively).*

Proof. Let $k \in \langle l \rangle$.

First, we show that

$$(\tau_k \circ \tau_{k-1} \circ ... \circ \tau_1)(k) = g(\tau_k).$$

Case 1. $k = 1$. Obvious.

Case 2. $k > 1$ (when $l > 1$). By hypothesis, $\tau_k \in C_{n,g(\tau_1),g(\tau_2),...,g(\tau_{k-1})} = \{(w_1), (w_2, w_1), ..., (w_{n-k+1}, w_{n-k}, ..., w_1)\}$, where $w_1 < w_2 < ... < w_{n-k+1}$ and $\{w_1, w_2, ..., w_{n-k+1}\} = \langle n \rangle - \{g(\tau_1), g(\tau_2), ..., g(\tau_{k-1})\}$. Notice that any cycle τ_k contains w_1, the smallest element of set $\{w_1, w_2, ..., w_{n-k+1}\}$. Now, we consider the sequence k, $k-1$, ..., 1. If $k \leq g(\tau_1)$, we have

$$\tau_1(k) = k - 1$$

(because $k > 1$); since $\tau_1(k) = k - 1$, we delete k from this sequence and, consequently, we obtain the sequence $k-1$, $k-2$, ..., 1. If $k > g(\tau_1)$, we have

$$\tau_1(k) = k;$$

since $\tau_2 \in C_{n,g(\tau_1)}$, we delete $g(\tau_1)$ from the sequence k, $k-1$, ..., 1 and, consequently, we obtain a sequence with $k-1$ elements too. Since the two obtained sequences have the same length, $k-1$, we use a common notation for them, $a_1^{(1)}, a_2^{(1)}, ..., a_{k-1}^{(1)}$, and, when $k > 2$, suppose that $a_1^{(1)} > a_2^{(1)} > ... > a_{k-1}^{(1)}$ ($a_1^{(1)} = k-1$ in the first sequence with $k-1$ elements and $a_1^{(1)} = k$ in the second sequence with $k-1$ elements, ...). Obviously,

$$\tau_1(k) = a_1^{(1)}.$$

Further, we proceed in this way for τ_2, using the sequence left $a_1^{(1)}$, $a_2^{(1)}$, ..., $a_{k-1}^{(1)}$ — we do this when $k > 2$ (k is fixed; if $k = 2$, we do not proceed in this way for τ_2 anymore, see below). In this case, if $a_1^{(1)} \leq g(\tau_2)$, we have

$$\tau_2\left(a_1^{(1)}\right) = a_2^{(1)}$$

(because $k > 2$); since $\tau_2\left(a_1^{(1)}\right) = a_2^{(1)}$, we delete $a_1^{(1)}$ from the sequence $a_1^{(1)}$, $a_2^{(1)}$, ..., $a_{k-1}^{(1)}$ and, consequently, we obtain the sequence $a_2^{(1)}$, $a_3^{(1)}$, ..., $a_{k-1}^{(1)}$. If $a_1^{(1)} > g(\tau_2)$, we have

$$\tau_2\left(a_1^{(1)}\right) = a_1^{(1)};$$

since $\tau_3 \in C_{n,g(\tau_1),g(\tau_2)}$, we delete $g(\tau_2)$ from the sequence $a_1^{(1)}$, $a_2^{(1)}$, ..., $a_{k-1}^{(1)}$ and, consequently, we obtain a sequence with $k-2$ elements too. Since the two obtained sequences have the same length, $k-2$, we use a common notation for them, $a_1^{(2)}$, $a_2^{(2)}$, ..., $a_{k-2}^{(2)}$, and, when $k > 3$, suppose that $a_1^{(2)} > a_2^{(2)} > ... > a_{k-2}^{(2)}$. Obviously,

$$\tau_2\left(a_1^{(1)}\right) = a_1^{(2)}.$$

Proceeding in this way for τ_3, τ_4, ..., τ_{k-1}, we obtain a sequence having only one element, $a_1^{(k-1)}$. Obviously,

$$\tau_{k-1}\left(a_1^{(k-2)}\right) = a_1^{(k-1)}.$$

$a_1^{(k-1)}$ is the smallest element from the sequence k, $k-1$, ..., 1, different from $g(\tau_1), g(\tau_2), ..., g(\tau_{k-1})$ (see the way it was obtained again). So,

$$a_1^{(k-1)} = w_1.$$

Finally, we have

$$(\tau_k \circ \tau_{k-1} \circ ... \circ \tau_1)(k) = \tau_k((\tau_{k-1} \circ ... \circ \tau_1)(k)) =$$
$$= \tau_k\left(a_1^{(k-1)}\right) = \tau_k(w_1) = g(\tau_k).$$

Second, we show that

$$(\tau_l \circ \tau_{l-1} \circ ... \circ \tau_1)(k) = g(\tau_k)$$

(recall that $k \in \langle l \rangle$).

Case 1. $l = 1$. Obvious ($l = 1 \Longrightarrow k = 1$).

Case 2. $l > 1$.

Subcase 2.1. $l = k$. Obvious (by the first part of proof).

Subcase 2.2. $l > k$. Since $(\tau_k \circ \tau_{k-1} \circ ... \circ \tau_1)(k) = g(\tau_k)$ and $\tau_{k+1} \in C_{n,g(\tau_1),g(\tau_2),...,g(\tau_k)}$, $\tau_{k+2} \in C_{n,g(\tau_1),g(\tau_2),...,g(\tau_{k+1})}$, ..., $\tau_l \in C_{n,g(\tau_1),g(\tau_2),...,g(\tau_{l-1})}$, we have

$$(\tau_l \circ \tau_{l-1} \circ ... \circ \tau_1)(k) = (\tau_l \circ \tau_{l-1} \circ ... \circ \tau_{k+1})((\tau_k \circ \tau_{k-1} \circ ... \circ \tau_1)(k)) =$$
$$= (\tau_l \circ \tau_{l-1} \circ ... \circ \tau_{k+1})(g(\tau_k)) = g(\tau_k). \quad \square$$

THEOREM 1.7. *Let $n \geq 2$. Let $\sigma_0 \in \mathbb{S}_n$. Let*

$$\mathbb{A}_{n,l} = \left\{\tau_l \circ \tau_{l-1} \circ ... \circ \tau_1 \circ \sigma_l \circ \sigma_0 \,\middle|\, \tau_1 \in C_n,\ \tau_2 \in C_{n,g(\tau_1)},\ ...\right.$$
$$\left. ...,\ \tau_l \in C_{n,g(\tau_1),g(\tau_2),...,g(\tau_{l-1})},\ \sigma_l \in \mathbb{S}_n,\ \sigma_l(v) = v,\ \forall v \in \langle l \rangle \right\},\ \forall l \in \langle n-1 \rangle$$

($g(\tau_1), g(\tau_2), ..., g(\tau_{l-1})$ vanish when $l = 1$). Then

$$\mathbb{A}_{n,l} = \mathbb{S}_n,\ \forall l \in \langle n-1 \rangle.$$

Proof. Let $l \in \langle n-1 \rangle$. Since $(\mathbb{S}_n, \circ)$ is a group, we have $\mathbb{A}_{n,l} \subseteq \mathbb{S}_n$. Therefore, $|\mathbb{A}_{n,l}| \leq |\mathbb{S}_n| = n!$ ($|\cdot|$ is the cardinal). To finish the proof, we show that $|\mathbb{A}_{n,l}| = n!$.

The number of permutations $\sigma_l \in \mathbb{S}_n$ with $\sigma_l(v) = v$, $\forall v \in \langle l \rangle$, is equal to $(n-l)!$. Since $\tau_1 \in C_n$ and $|C_n| = n$, $\tau_2 \in C_{n,g(\tau_1)}$ and $\left|C_{n,g(\tau_1)}\right| = n-1$, ..., $\tau_l \in C_{n,g(\tau_1),g(\tau_2),\ldots,g(\tau_{l-1})}$ and $\left|C_{n,g(\tau_1),g(\tau_2),\ldots,g(\tau_{l-1})}\right| = n-l+1$, it follows that $|\mathbb{A}_{n,l}|$ is at most equal to

$$n(n-1)\ldots(n-l+1)[(n-l)!] = n!.$$

We show that

$$\tau_l \circ \tau_{l-1} \circ \ldots \circ \tau_1 \circ \sigma_l \circ \sigma_0 = \mu_l \circ \mu_{l-1} \circ \ldots \circ \mu_1 \circ \nu_l \circ \sigma_0$$

if and only if $\tau_q = \mu_q$, $\forall q \in \langle l \rangle$, and $\sigma_l = \nu_l$, where $\tau_1, \mu_1 \in C_n$, $\tau_2 \in C_{n,g(\tau_1)}$, $\mu_2 \in C_{n,g(\mu_1)}$, ..., $\tau_l \in C_{n,g(\tau_1),g(\tau_2),\ldots,g(\tau_{l-1})}$, $\mu_l \in C_{n,g(\mu_1),g(\mu_2),\ldots,g(\mu_{l-1})}$, $\sigma_l, \nu_l \in \mathbb{S}_n$, $\sigma_l(v) = \nu_l(v) = v$, $\forall v \in \langle l \rangle$.

" $\Longleftarrow$ " Obvious.

" $\Longrightarrow$ " Since σ_0 can be eliminated, we eliminate it. So, we suppose that

$$\tau_l \circ \tau_{l-1} \circ \ldots \circ \tau_1 \circ \sigma_l = \mu_l \circ \mu_{l-1} \circ \ldots \circ \mu_1 \circ \nu_l$$

It follows that

$$(\tau_l \circ \tau_{l-1} \circ \ldots \circ \tau_1 \circ \sigma_l)(k) = (\mu_l \circ \mu_{l-1} \circ \ldots \circ \mu_1 \circ \nu_l)(k), \ \forall k \in \langle l \rangle.$$

Therefore, by Theorem 1.6 we have

$$g(\tau_1) = g(\mu_1), \ g(\tau_2) = g(\mu_2), \ \ldots, \ g(\tau_l) = g(\mu_l),$$

so,

$$\tau_1 = \mu_1, \ \tau_2 = \mu_2, \ \ldots, \ \tau_l = \mu_l,$$

and, as a result of these equations,

$$\sigma_l = \nu_l.$$

We conclude that

$$|\mathbb{A}_{n,l}| = n!. \quad \square$$

Theorem 1.7 says that we can work with $\mathbb{A}_{n,l}$ instead of $\mathbb{S}_n$, $\forall l \in \langle n-1 \rangle$ (this fact will be used in Theorem 2.1).

Let $\sigma, \tau \in \mathbb{S}_n$. Set

$$K(\sigma,\tau) = \sum_{1 \le i < j \le n} \mathbf{1}[(\sigma(i) - \sigma(j))(\tau(i) - \tau(j)) < 0],$$

where

$$\mathbf{1}[(\sigma(i)-\sigma(j))(\tau(i)-\tau(j))<0] = \begin{cases} 1 & \text{if } (\sigma(i)-\sigma(j))(\tau(i)-\tau(j))<0, \\ 0 & \text{if } (\sigma(i)-\sigma(j))(\tau(i)-\tau(j))\ge 0, \end{cases}$$

$\forall i, j \in \langle n \rangle$. K is a metric on $\mathbb{S}_n$, called the *Kendall metric* (see, *e.g.*, [1] or [8]).

THEOREM 1.8. *Let $n \geq 2$. Let $\sigma_0 \in \mathbb{S}_n$. Consider on $\mathbb{S}_n$ the Kendall metric. Then*

$$K\left(\tau_l \circ \tau_{l-1} \circ ... \circ \tau_1 \circ \sigma_l \circ \sigma_0, \sigma_0\right) = \lfloor \tau_l \rceil - 1 + K\left(\tau_{l-1} \circ \tau_{l-2} \circ ... \circ \tau_1 \circ \sigma_l \circ \sigma_0, \sigma_0\right),$$

$\forall l \in \langle n-1 \rangle$ *(warning! σ_l appears in two terms of the above equation), where $\tau_1 \in C_n$, $\tau_2 \in C_{n,g(\tau_1)}$, ..., $\tau_l \in C_{n,g(\tau_1),g(\tau_2),...,g(\tau_{l-1})}$, $\sigma_l \in \mathbb{S}_n$, $\sigma_l(v) = v$, $\forall v \in \langle l \rangle$, and $\lfloor \tau_l \rceil$ is the length of τ_l (a degenerate cycle has length 1) — $\tau_{l-1} \circ \tau_{l-2} \circ ... \circ \tau_1$ and $g(\tau_1), g(\tau_2), ..., g(\tau_{l-1})$ vanish when $l = 1$.*

Proof. Since K is right-invariant (see, *e.g.*, [1]), the equation we must show is equivalent to

$$K\left(\tau_l \circ \tau_{l-1} \circ ... \circ \tau_1 \circ \sigma_l, \mathrm{Id}\right) = \lfloor \tau_l \rceil - 1 + K\left(\tau_{l-1} \circ \tau_{l-2} \circ ... \circ \tau_1 \circ \sigma_l, \mathrm{Id}\right),$$

$\forall l \in \langle n-1 \rangle$. We shall show the latter equation. More precisely, we shall show that

$$K\left(\tau_l \circ \tau_{l-1} \circ ... \circ \tau_1 \circ \sigma_l, \mathrm{Id}\right) - K\left(\tau_{l-1} \circ \tau_{l-2} \circ ... \circ \tau_1 \circ \sigma_l, \mathrm{Id}\right) = \lfloor \tau_l \rceil - 1,$$

$\forall l \in \langle n-1 \rangle$.

Let $l \in \langle n-1 \rangle$. Since the above equation is obvious for τ_l =Id, further, we suppose that

$$\tau_l = (w_t, w_{t-1}, ..., w_1),$$

where $t \geq 2$ $(w_1 < w_2 < ... < w_t)$. It follows that $\lfloor \tau_l \rceil = t$. Let $i_1, i_2, ..., i_t \in \langle n \rangle$, $i_a \neq i_b$, $\forall a, b \in \langle t \rangle$, $a \neq b$. Suppose that

$$\left(\tau_{l-1} \circ \tau_{l-2} \circ ... \circ \tau_1 \circ \sigma_l\right)(i_a) = w_a, \forall a \in \langle t \rangle.$$

(If $l = 1$, since $\tau_{l-1} \circ \tau_{l-2} \circ ... \circ \tau_1$ vanishes when $l = 1$, we have $\sigma_1(i_a) = w_a$, $\forall a \in \langle t \rangle$.)

Let i, $j \in \langle n \rangle$, $i \neq j$. Below we consider four cases.

Case 1. i, $j \in \{i_1, i_2, ..., i_t\}$. We have

$$\left(\tau_{l-1} \circ \tau_{l-2} \circ ... \circ \tau_1 \circ \sigma_l\right)(i_1) = w_1 < w_b = \left(\tau_{l-1} \circ \tau_{l-2} \circ ... \circ \tau_1 \circ \sigma_l\right)(i_b)$$

and

$$\left(\tau_l \circ \tau_{l-1} \circ ... \circ \tau_1 \circ \sigma_l\right)(i_1) = w_t > w_{b-1} = \left(\tau_l \circ \tau_{l-1} \circ ... \circ \tau_1 \circ \sigma_l\right)(i_b),$$

$\forall b \in \langle t \rangle - \{1\}$. We also have

$$\left(\tau_{l-1} \circ \tau_{l-2} \circ ... \circ \tau_1 \circ \sigma_l\right)(i_a) = w_a < w_b = \left(\tau_{l-1} \circ \tau_{l-2} \circ ... \circ \tau_1 \circ \sigma_l\right)(i_b)$$

and

$$\left(\tau_l \circ \tau_{l-1} \circ ... \circ \tau_1 \circ \sigma_l\right)(i_a) = w_{a-1} < w_{b-1} = \left(\tau_l \circ \tau_{l-1} \circ ... \circ \tau_1 \circ \sigma_l\right)(i_b),$$

$\forall a, b \in \langle t \rangle - \{1\}$, $a < b$. Therefore,

$$\left(\tau_{l-1} \circ \tau_{l-2} \circ ... \circ \tau_1 \circ \sigma_l\right)(i) < \left(\tau_{l-1} \circ \tau_{l-2} \circ ... \circ \tau_1 \circ \sigma_l\right)(j)$$

and

$$(\tau_l \circ \tau_{l-1} \circ ... \circ \tau_1 \circ \sigma_l)(i) > (\tau_l \circ \tau_{l-1} \circ ... \circ \tau_1 \circ \sigma_l)(j)$$

when $i \in \{i_1\}$ and $j \in \{i_2, i_3, ..., i_t\}$ only. Further, we show that

$$i < j \text{ if } i \in \{i_1\} \text{ and } j \in \{i_2, i_3, ..., i_t\}$$

($i < j$ is of interest to us, see the definition of K). This follows from $i_1 < i_2$, i_3, ..., i_t. We show that

$$i_1 < i_2, \ i_3, \ ..., \ i_t.$$

By Theorem 1.6, $(\tau_l \circ \tau_{l-1} \circ ... \circ \tau_1 \circ \sigma_l)(l) = g(\tau_l) = w_t$. It follows that

$$(\tau_{l-1} \circ \tau_{l-2} \circ ... \circ \tau_1 \circ \sigma_l)(l) = w_1$$

and, consequently,

$$i_1 = l.$$

Consider that $l > 1$. By Theorem 1.6, the first $l-1$ components of $\tau_{l-1} \circ \tau_{l-2} \circ ... \circ \tau_1 \circ \sigma_l$ are $g(\tau_1), g(\tau_2), ..., g(\tau_{l-1})$, respectively. As $\tau_l \in C_{n, g(\tau_1), g(\tau_2), ..., g(\tau_{l-1})}$ and $\tau_l = (w_t, w_{t-1}, ..., w_1)$, it follows that

$$(\tau_{l-1} \circ \tau_{l-2} \circ ... \circ \tau_1 \circ \sigma_l)(i_a) = w_a \notin \{g(\tau_1), g(\tau_2), ..., g(\tau_{l-1})\}, \forall a \in \langle t \rangle.$$

Therefore, we have

$$i_1, i_2, ..., i_t \geq l.$$

Since $i_1 = l$, we have $i_1 < i_2$, i_3, ..., i_t. Now consider that $l = 1$. It follows that $i_1 = 1$. Therefore, we also have $i_1 < i_2$, i_3, ..., i_t. So, for $l \geq 1$, we have $i_1 < i_2$, i_3, ..., i_t. Finally, we have (see the result we must show and definition of K again)

$$\begin{aligned}&\mathbf{1}[[(\tau_l \circ \tau_{l-1} \circ ... \circ \tau_1 \circ \sigma_l)(i) - (\tau_l \circ \tau_{l-1} \circ ... \circ \tau_1 \circ \sigma_l)(j)][\mathrm{Id}(i) - \mathrm{Id}(j)] < 0] - \\ &-\mathbf{1}[[(\tau_{l-1} \circ \tau_{l-2} \circ ... \circ \tau_1 \circ \sigma_l)(i) - (\tau_{l-1} \circ \tau_{l-2} \circ ... \circ \tau_1 \circ \sigma_l)(j)][\mathrm{Id}(i) - \mathrm{Id}(j)] < 0] = \\ &= \begin{cases} 1 & \text{if } i \in \{i_1\}, \ j \in \{i_2, i_3, ..., i_t\} \ (i < j), \\ 0 & \text{if } i, \ j \in \{i_2, i_3, ..., i_t\} \ (i < j \text{ or } i > j). \end{cases}\end{aligned}$$

Case 2. $i, \ j \notin \{i_1, i_2, ..., i_t\}$. This case holds when $t \leq n - 2$. Since i, $j \notin \{i_1, i_2, ..., i_t\}$, we have

$$(\tau_{l-1} \circ \tau_{l-2} \circ ... \circ \tau_1 \circ \sigma_l)(i), \ \ (\tau_{l-1} \circ \tau_{l-2} \circ ... \circ \tau_1 \circ \sigma_l)(j) \notin \{w_1, w_2, ..., w_t\}.$$

Suppose that (the first subcase)

$$(\tau_{l-1} \circ \tau_{l-2} \circ ... \circ \tau_1 \circ \sigma_l)(i) < (\tau_{l-1} \circ \tau_{l-2} \circ ... \circ \tau_1 \circ \sigma_l)(j).$$

It follows that

$$(\tau_l \circ \tau_{l-1} \circ ... \circ \tau_1 \circ \sigma_l)(i) < (\tau_l \circ \tau_{l-1} \circ ... \circ \tau_1 \circ \sigma_l)(j)$$

because

$$(\tau_l \circ \tau_{l-1} \circ ... \circ \tau_1 \circ \sigma_l)(i) = (\tau_{l-1} \circ \tau_{l-2} \circ ... \circ \tau_1 \circ \sigma_l)(i)$$

and

$$(\tau_l \circ \tau_{l-1} \circ ... \circ \tau_1 \circ \sigma_l)(j) = (\tau_{l-1} \circ \tau_{l-2} \circ ... \circ \tau_1 \circ \sigma_l)(j).$$

Further, we have

$$\mathbf{1}[[(\tau_l \circ \tau_{l-1} \circ ... \circ \tau_1 \circ \sigma_l)(i) - (\tau_l \circ \tau_{l-1} \circ ... \circ \tau_1 \circ \sigma_l)(j)][\mathrm{Id}(i) - \mathrm{Id}(j)] < 0] - \\ -\mathbf{1}[[(\tau_{l-1} \circ \tau_{l-2} \circ ... \circ \tau_1 \circ \sigma_l)(i) - (\tau_{l-1} \circ \tau_{l-2} \circ ... \circ \tau_1 \circ \sigma_l)(j)][\mathrm{Id}(i) - \mathrm{Id}(j)] < 0] = 0$$

both when $i < j$ and when $i > j$ ($i < j$ is of interest to us, see the definition of K). The subcase when

$$(\tau_{l-1} \circ \tau_{l-2} \circ ... \circ \tau_1 \circ \sigma_l)(i) > (\tau_{l-1} \circ \tau_{l-2} \circ ... \circ \tau_1 \circ \sigma_l)(j)$$

is similar to the above subcase.

Case 3. $i \in \{i_1, i_2, ..., i_t\}$, $j \notin \{i_1, i_2, ..., i_t\}$. (Recall that i_1, i_2, ..., $i_t \geq l$, see Case 1 again.) This case holds when $t \leq n-1$.

Subcase 3.1. $i \in \{i_1, i_2, ..., i_t\}$, $j \in \langle l-1 \rangle$. This subcase holds when $l > 1$. Since i_1, i_2, ..., $i_t \geq l$, we have $i > j$, so, this subcase is not of interest to us.

Subcase 3.2. $i \in \{i_1, i_2, ..., i_t\}$, $j \in \langle n \rangle - (\langle l-1 \rangle \cup \{i_1, i_2, ..., i_t\})$ ($\langle l-1 \rangle$ vanishes when $l = 1$). Since $(\tau_{l-1} \circ \tau_{l-2} \circ ... \circ \tau_1 \circ \sigma_l)(i_a) = w_a$, $\forall a \in \langle t \rangle$, and $g(\tau_l) = w_t$, we have

$$(\tau_{l-1} \circ \tau_{l-2} \circ ... \circ \tau_1 \circ \sigma_l)(i) \in \langle g(\tau_l) \rangle$$

($\langle g(\tau_l) \rangle = \{1, 2, ..., g(\tau_l)\}$). Further, we show that

$$(\tau_{l-1} \circ \tau_{l-2} \circ ... \circ \tau_1 \circ \sigma_l)(j) \notin \langle g(\tau_l) \rangle.$$

Consider that $l > 1$. Since $j \notin \langle l-1 \rangle$ and, by Theorem 1.6, the first $l-1$ entries of $\tau_{l-1} \circ \tau_{l-2} \circ ... \circ \tau_1 \circ \sigma_l$ are $g(\tau_1)$, $g(\tau_2)$, ..., $g(\tau_{l-1})$, respectively, we have

$$(\tau_{l-1} \circ \tau_{l-2} \circ ... \circ \tau_1 \circ \sigma_l)(j) \notin \{g(\tau_1), g(\tau_2), ..., g(\tau_{l-1})\}.$$

Since $j \notin \{i_1, i_2, ..., i_t\}$ and $(\tau_{l-1} \circ \tau_{l-2} \circ ... \circ \tau_1 \circ \sigma_l)(i_a) = w_a$, $\forall a \in \langle t \rangle$, we have

$$(\tau_{l-1} \circ \tau_{l-2} \circ ... \circ \tau_1 \circ \sigma_l)(j) \notin \{w_1, w_2, ..., w_t\}.$$

As $\tau_l \in C_{n, g(\tau_1), g(\tau_2), ..., g(\tau_{l-1})}$, $\tau_l = (w_t,\ w_{t-1},\ ...,\ w_1)$, and $g(\tau_l) = w_t$, we have

$$\langle g(\tau_l) \rangle \subseteq \{w_1, w_2, ..., w_t\} \cup \{g(\tau_1), g(\tau_2), ..., g(\tau_{l-1})\}.$$

Since (see above)

$$(\tau_{l-1} \circ \tau_{l-2} \circ ... \circ \tau_1 \circ \sigma_l)(j) \notin \\ \notin \{w_1, w_2, ..., w_t\} \cup \{g(\tau_1), g(\tau_2), ..., g(\tau_{l-1})\} \supseteq \langle g(\tau_l) \rangle,$$

we have

$$\left(\tau_{l-1}\circ\tau_{l-2}\circ...\circ\tau_1\circ\sigma_l\right)(j)\notin\langle g(\tau_l)\rangle.$$

Now consider that $l=1$. As $\tau_1\in C_n$ ($w_a=a$, $\forall a\in\langle t\rangle$; therefore, $\tau_1=(t,t-1,...,1)$) and $j\in\langle n\rangle-\{i_1,i_2,...,i_t\}$ ($i_1=1$ when $l=1$, see Case 1; therefore, $j\neq1$), we have

$$\sigma_1(j)\notin\langle g(\tau_1)\rangle=\langle t\rangle$$

(recall that $\tau_{l-1}\circ\tau_{l-2}\circ...\circ\tau_1$ vanishes from $\left(\tau_{l-1}\circ\tau_{l-2}\circ...\circ\tau_1\circ\sigma_l\right)(j)$ when $l=1$; recall that $\sigma_1(i_a)=w_a$, $\forall a\in\langle t\rangle$; recall that $\sigma_1(1)=1$). So, for $l\geq1$, $\left(\tau_{l-1}\circ\tau_{l-2}\circ...\circ\tau_1\circ\sigma_l\right)(j)\notin\langle g(\tau_l)\rangle$. Further, since $\left(\tau_{l-1}\circ\tau_{l-2}\circ...\circ\tau_1\circ\sigma_l\right)(i)\in\langle g(\tau_l)\rangle$ and $\left(\tau_{l-1}\circ\tau_{l-2}\circ...\circ\tau_1\circ\sigma_l\right)(j)\notin\langle g(\tau_l)\rangle$, we have

$$\left(\tau_{l-1}\circ\tau_{l-2}\circ...\circ\tau_1\circ\sigma_l\right)(i)<\left(\tau_{l-1}\circ\tau_{l-2}\circ...\circ\tau_1\circ\sigma_l\right)(j).$$

Since $\left(\tau_l\circ\tau_{l-1}\circ...\circ\tau_1\circ\sigma_l\right)(i)\in\langle g(\tau_l)\rangle$ and $\left(\tau_l\circ\tau_{l-1}\circ...\circ\tau_1\circ\sigma_l\right)(j)=\left(\tau_{l-1}\circ\tau_{l-2}\circ...\circ\tau_1\circ\sigma_l\right)(j)\notin\langle g(\tau_l)\rangle$, we have

$$\left(\tau_l\circ\tau_{l-1}\circ...\circ\tau_1\circ\sigma_l\right)(i)<\left(\tau_l\circ\tau_{l-1}\circ...\circ\tau_1\circ\sigma_l\right)(j).$$

Finally, we have

$$\mathbf{1}\left[\left[\left(\tau_l\circ\tau_{l-1}\circ...\circ\tau_1\circ\sigma_l\right)(i)-\left(\tau_l\circ\tau_{l-1}\circ...\circ\tau_1\circ\sigma_l\right)(j)\right]\left[\mathrm{Id}(i)-\mathrm{Id}(j)\right]<0\right]-$$
$$-\mathbf{1}\left[\left[\left(\tau_{l-1}\circ\tau_{l-2}\circ...\circ\tau_1\circ\sigma_l\right)(i)-\left(\tau_{l-1}\circ\tau_{l-2}\circ...\circ\tau_1\circ\sigma_l\right)(j)\right]\left[\mathrm{Id}(i)-\mathrm{Id}(j)\right]<0\right]=0$$

both when $i<j$ and when $i>j$.

Case 4. $i\notin\{i_1,i_2,...,i_t\}$, $j\in\{i_1,i_2,...,i_t\}$. This case holds when $t\leq n-1$.

Subcase 4.1. $i\in\langle l-1\rangle$, $j\in\{i_1,i_2,...,i_t\}$. This subcase holds when $l>1$. Since i_1, i_2, ..., $i_t\geq l$ (see Case 1 again), we have $i<j$. Suppose that (the first sub-subcase)

$$\left(\tau_{l-1}\circ\tau_{l-2}\circ...\circ\tau_1\circ\sigma_l\right)(i)<\left(\tau_{l-1}\circ\tau_{l-2}\circ...\circ\tau_1\circ\sigma_l\right)(j).$$

As $j\in\{i_1,i_2,...,i_t\}$, it follows that $\exists a\in\langle t\rangle$ such that $j=i_a$. Therefore,

$$\left(\tau_{l-1}\circ\tau_{l-2}\circ...\circ\tau_1\circ\sigma_l\right)(j)=\left(\tau_{l-1}\circ\tau_{l-2}\circ...\circ\tau_1\circ\sigma_l\right)(i_a)=w_a.$$

As τ_l is a cycle, $\exists b\in\langle t\rangle$ such that

$$\left(\tau_l\circ\tau_{l-1}\circ...\circ\tau_1\circ\sigma_l\right)(i_b)=w_a$$

(recall that $\left(\tau_{l-1}\circ\tau_{l-2}\circ...\circ\tau_1\circ\sigma_l\right)(i_u)=w_u$, $\forall u\in\langle t\rangle$, and $\tau_l=(w_t,w_{t-1},...,w_1)$). Therefore, since $\left(\tau_l\circ\tau_{l-1}\circ...\circ\tau_1\circ\sigma_l\right)(i)=\left(\tau_{l-1}\circ\tau_{l-2}\circ...\circ\tau_1\circ\sigma_l\right)(i)$, we have (using what we supposed in this sub-subcase)

$$\left(\tau_l\circ\tau_{l-1}\circ...\circ\tau_1\circ\sigma_l\right)(i)<\left(\tau_l\circ\tau_{l-1}\circ...\circ\tau_1\circ\sigma_l\right)(i_b)=w_a.$$

Obviously, $i < i_b$. We associate the pair of inequalities

$$i < j = i_a,$$
$$(\tau_{l-1} \circ \tau_{l-2} \circ ... \circ \tau_1 \circ \sigma_l)(i) < (\tau_{l-1} \circ \tau_{l-2} \circ ... \circ \tau_1 \circ \sigma_l)(j) = w_a$$

with the pair of inequalities

$$i < i_b,$$
$$(\tau_l \circ \tau_{l-1} \circ ... \circ \tau_1 \circ \sigma_l)(i) < (\tau_l \circ \tau_{l-1} \circ ... \circ \tau_1 \circ \sigma_l)(i_b) = w_a.$$

By this association (trick) we have (see the definition of K and result we must show again)

$$\mathbf{1}[[(\tau_l \circ \tau_{l-1} \circ ... \circ \tau_1 \circ \sigma_l)(i) - (\tau_l \circ \tau_{l-1} \circ ... \circ \tau_1 \circ \sigma_l)(i_b)][\mathrm{Id}(i) - \mathrm{Id}(i_b)] < 0] -$$
$$-\mathbf{1}[[(\tau_{l-1} \circ \tau_{l-2} \circ ... \circ \tau_1 \circ \sigma_l)(i) - (\tau_{l-1} \circ \tau_{l-2} \circ ... \circ \tau_1 \circ \sigma_l)(j)][\mathrm{Id}(i) - \mathrm{Id}(j)] < 0] = 0.$$

The sub-subcase when

$$(\tau_{l-1} \circ \tau_{l-2} \circ ... \circ \tau_1 \circ \sigma_l)(i) > (\tau_{l-1} \circ \tau_{l-2} \circ ... \circ \tau_1 \circ \sigma_l)(j)$$

is similar to the above sub-subcase.

Subcase 4.2. $i \in \langle n \rangle - (\langle l-1 \rangle \cup \{i_1, i_2, ..., i_t\})$, $j \in \{i_1, i_2, ..., i_t\}$. Similar to Subcase 3.2.

From Cases 1–4, we have

$$K(\tau_l \circ \tau_{l-1} \circ ... \circ \tau_1 \circ \sigma_l, \mathrm{Id}) - K(\tau_{l-1} \circ \tau_{l-2} \circ ... \circ \tau_1 \circ \sigma_l, \mathrm{Id}) = \lfloor \tau_l \rceil - 1$$

($\lfloor \tau_l \rceil - 1$ is due to Case 1). $\square$

2. OUR MARKOVIAN METHOD

In this section, we present Mallows model and our fast Markovian method for sampling exactly (not approximately) from $\mathbb{S}_n$, the set of permutations of order n, according to the Mallows model through Kendall metric. In addition to sampling, for this special Mallows model, the Markov chain we construct can do other things — we compute the normalizing constant and, by Uniqueness Theorem, certain important probabilities.

Recall that $\mathbb{R}^+ = \{x \mid x \in \mathbb{R} \text{ and } x > 0\}$.

Let

$$\pi_\sigma = \frac{\theta^{d(\sigma,\sigma_0)}}{Z}, \ \forall \sigma \in \mathbb{S}_n,$$

where $\theta \in \mathbb{R}^+$ $(0 < \theta \leq 1$ or $\theta > 1)$, $\sigma_0 \in \mathbb{S}_n$ $(n \geq 1)$, d is a metric on $\mathbb{S}_n$, and

$$Z = \sum_{\sigma \in \mathbb{S}_n} \theta^{d(\sigma,\sigma_0)}.$$

The probability distribution $\pi = (\pi_\sigma)_{\sigma\in\mathbb{S}_n}$ (on $\mathbb{S}_n$) is called the *Mallows model through metric* d (see [7]; see, *e.g.*, also [1, 4, 8]). This is a model — an exponential model when $\theta \neq 1$ — for ranked data (see the above references). Here we considered a generalization of the classical Mallows model (through metric d) — $0 < \theta \leq 1$ in the classical case.

In this article, the transpose of a vector x is denoted by x'. Set $e = e(n) = (1, 1, ..., 1) \in \mathbb{R}^n$, $\forall n \geq 1$.

Below we give the main result of this article.

THEOREM 2.1. *Let $n \geq 2$. Let $\pi = (\pi_\sigma)_{\sigma\in\mathbb{S}_n}$ be the Mallows model through Kendall metric. Consider a Markov chain with state space $\mathbb{S}_n$ and transition matrix $P = P_1P_2...P_{n-1}$, where P_l, $l \in \langle n-1\rangle$, are stochastic matrices on $\mathbb{S}_n$,*

$$(P_l)_{\tau_l\circ\tau_{l-1}\circ...\circ\tau_1\circ\sigma_l\circ\sigma_0\to\xi} =$$

$$= \begin{cases} \dfrac{\pi_{\varphi\circ\tau_{l-1}\circ\tau_{l-2}\circ...\circ\tau_1\circ\sigma_l\circ\sigma_0}}{\sum\limits_{\nu\in C_{n,g(\tau_1),g(\tau_2),...,g(\tau_{l-1})}} \pi_{\nu\circ\tau_{l-1}\circ\tau_{l-2}\circ...\circ\tau_1\circ\sigma_l\circ\sigma_0}} & \begin{array}{l}\text{if } \xi = \varphi\circ\tau_{l-1}\circ\tau_{l-2}\circ...\circ\\ \circ\tau_1\circ\sigma_l\circ\sigma_0 \text{ for some}\\ \varphi\in C_{n,g(\tau_1),g(\tau_2),...,g(\tau_{l-1})},\end{array} \\ 0 & \begin{array}{l}\text{if } \xi \neq \varphi\circ\tau_{l-1}\circ\tau_{l-2}\circ...\circ\\ \circ\tau_1\circ\sigma_l\circ\sigma_0,\ \forall\varphi\in\\ \in C_{n,g(\tau_1),g(\tau_2),...,g(\tau_{l-1})},\end{array} \end{cases}$$

$\forall l \in \langle n-1\rangle$, $\forall\tau_1 \in C_n$, $\forall\tau_2 \in C_{n,g(\tau_1)}$, ..., $\forall\tau_l \in C_{n,g(\tau_1),g(\tau_2),...,g(\tau_{l-1})}$, $\forall\sigma_l \in \mathbb{S}_n$, $\sigma_l(v) = v$, $\forall v \in \langle l\rangle$, $\forall\xi \in \mathbb{S}_n$ ($\tau_{l-1}\circ\tau_{l-2}\circ...\circ\tau_1$ and $g(\tau_1)$, $g(\tau_2)$, ..., $g(\tau_{l-1})$ vanish when $l = 1$), where σ_0 is the parameter from $\mathbb{S}_n$ of Mallows model through Kendall metric. Then

$$P = e'\pi$$

(therefore, the chain attains its stationarity at time 1 and, obviously, π is its stationary probability distribution (limit probability distribution)).

Proof. Set

$$L_{\tau_1,\tau_2,...,\tau_l} = \{\tau_l\circ\tau_{l-1}\circ...\circ\tau_1\circ\sigma_l\circ\sigma_0 \mid \sigma_l \in \mathbb{S}_n, \sigma_l(v) = v, \forall v \in \langle l\rangle\},$$

$\forall l \in \langle n-1\rangle$, $\forall\tau_1 \in C_n$, $\forall\tau_2 \in C_{n,g(\tau_1)}$, ..., $\forall\tau_l \in C_{n,g(\tau_1),g(\tau_2),...,g(\tau_{l-1})}$. We have

$$\bigcup_{\substack{\tau_1\in C_n\\ \tau_2\in C_{n,g(\tau_1)}\\ \vdots\\ \tau_l\in C_{n,g(\tau_1),g(\tau_2),...,g(\tau_{l-1})}}} L_{\tau_1,\tau_2,...,\tau_l} = \mathbb{A}_{n,l} = \mathbb{S}_n, \forall l \in \langle n-1\rangle$$

(see Theorem 1.7). We show that

$$L_{\tau_1,\tau_2,...,\tau_l} \cap L_{\mu_1,\mu_2,...,\mu_l} = \emptyset$$

if $\exists u \in \langle l \rangle$ such that $\tau_u \neq \mu_u$, where $l \in \langle n-1 \rangle$, $\tau_1, \mu_1 \in C_n$, $\tau_2 \in C_{n,g(\tau_1)}$, $\mu_2 \in C_{n,g(\mu_1)}$, ..., $\tau_l \in C_{n,g(\tau_1),g(\tau_2),...,g(\tau_{l-1})}$, $\mu_l \in C_{n,g(\mu_1),g(\mu_2),...,g(\mu_{l-1})}$. For this we suppose that $\exists u \in \langle l \rangle$ with $\tau_u \neq \mu_u$ such that

$$L_{\tau_1,\tau_2,...,\tau_l} \cap L_{\mu_1,\mu_2,...,\mu_l} \neq \emptyset.$$

Let $\omega \in L_{\tau_1,\tau_2,...,\tau_l} \cap L_{\mu_1,\mu_2,...,\mu_l}$. We have

$$\omega = \tau_l \circ \tau_{l-1} \circ ... \circ \tau_1 \circ \sigma_l \circ \sigma_0 = \mu_l \circ \mu_{l-1} \circ ... \circ \mu_1 \circ \phi_l \circ \sigma_0,$$

where $\sigma_l, \phi_l \in \mathbb{S}_n$, $\sigma_l(v) = \phi_l(v) = v$, $\forall v \in \langle l \rangle$. Proceeding as in the proof of Theorem 1.7 (σ_0 is eliminated, ...), we obtain

$$\tau_1 = \mu_1,\ \tau_2 = \mu_2,\ ...,\ \tau_l = \mu_l,\ \sigma_l = \phi_l.$$

Therefore, we obtained a contradiction.

The above results lead to the fact that

$$\left(L_{\tau_1,\tau_2,...,\tau_l}\right)_{\substack{\tau_1 \in C_n \\ \tau_2 \in C_{n,g(\tau_1)} \\ \vdots \\ \tau_l \in C_{n,g(\tau_1),g(\tau_2),...,g(\tau_{l-1})}}}$$

is a partition of $\mathbb{A}_{n,l}$ ($\mathbb{A}_{n,l} = \mathbb{S}_n$), $\forall l \in \langle n-1 \rangle$. Set the partitions (this can now be done)

$$\Delta_1 = (\mathbb{S}_n),$$

$$\Delta_{l+1} = \left(L_{\tau_1,\tau_2,...,\tau_l}\right)_{\substack{\tau_1 \in C_n \\ \tau_2 \in C_{n,g(\tau_1)} \\ \vdots \\ \tau_l \in C_{n,g(\tau_1),g(\tau_2),...,g(\tau_{l-1})}}},$$

$\forall l \in \langle n-1 \rangle$. Obviously, we have $\Delta_n = (\{\sigma\})_{\sigma \in \mathbb{S}_n}$.

By hypothesis and Theorem 1.8 we have

$$(P_l)_{\tau_l \circ \tau_{l-1} \circ ... \circ \tau_1 \circ \sigma_l \circ \sigma_0 \to \xi} =$$

$$= \begin{cases} \dfrac{\theta^{K(\varphi \circ \tau_{l-1} \circ \tau_{l-2} \circ ... \circ \tau_1 \circ \sigma_l \circ \sigma_0, \sigma_0)}}{\sum\limits_{\nu \in C_{n,g(\tau_1),g(\tau_2),...,g(\tau_{l-1})}} \theta^{K(\nu \circ \tau_{l-1} \circ \tau_{l-2} \circ ... \circ \tau_1 \circ \sigma_l \circ \sigma_0, \sigma_0)}} & \text{if } \xi = \varphi \circ \tau_{l-1} \circ \tau_{l-2} \circ ... \circ \\ & \circ \tau_1 \circ \sigma_l \circ \sigma_0 \text{ for some} \\ & \varphi \in C_{n,g(\tau_1),g(\tau_2),...,g(\tau_{l-1})}, \\ 0 & \text{if } \xi \neq \varphi \circ \tau_{l-1} \circ \tau_{l-2} \circ ... \circ \\ & \circ \tau_1 \circ \sigma_l \circ \sigma_0,\ \forall \varphi \in \\ & \in C_{n,g(\tau_1),g(\tau_2),...,g(\tau_{l-1})}, \end{cases}$$

$$= \begin{cases} \frac{\theta^{\lfloor\varphi\rceil-1}}{1+\theta+\theta^2+...+\theta^{n-l}} & \text{if } \xi = \varphi \circ \tau_{l-1} \circ \tau_{l-2} \circ ... \circ \\ & \circ \tau_1 \circ \sigma_l \circ \sigma_0 \text{ for some} \\ & \varphi \in C_{n,g(\tau_1),g(\tau_2),...,g(\tau_{l-1})}, \\ & \text{if } \xi \neq \varphi \circ \tau_{l-1} \circ \tau_{l-2} \circ ... \circ \\ 0 & \circ \tau_1 \circ \sigma_l \circ \sigma_0, \ \forall \varphi \in \\ & \in C_{n,g(\tau_1),g(\tau_2),...,g(\tau_{l-1})}, \end{cases}$$

$\forall l \in \langle n-1 \rangle$, $\forall \tau_1 \in C_n$, $\forall \tau_2 \in C_{n,g(\tau_1)}, ...,$ $\forall \tau_l \in C_{n,g(\tau_1),g(\tau_2),...,g(\tau_{l-1})}$, $\forall \sigma_l \in \mathbb{S}_n$, $\sigma_l(v) = v$, $\forall v \in \langle l \rangle$, $\forall \xi \in \mathbb{S}_n$, where $\lfloor\varphi\rceil$ is the length of (cycle) φ. (Recall that a degenerate cycle has the length 1 and $1 + \theta + \theta^2 + ... + \theta^{n-l} = \frac{1-\theta^{n-l+1}}{1-\theta}$ if $\theta \neq 1$.) It follows that

$$P_l \in G_{\Delta_l,\Delta_{l+1}}, \ \forall l \in \langle n-1 \rangle .$$

Since $P = P_1P_2...P_{n-1}$, by Theorem 1.5, P is a stable matrix. Consequently, $\exists\psi$, ψ is a probability distribution on $\mathbb{S}_n$, such that

$$P = e'\psi.$$

It is easy to see that

$$\pi_\sigma (P_l)_{\sigma\tau} = \pi_\tau (P_l)_{\tau\sigma}, \forall l \in \langle n-1 \rangle, \forall \sigma, \tau \in \mathbb{S}_n$$

($\mathbb{S}_n = \mathbb{A}_{n,l}$, $\forall l \in \langle n-1 \rangle$). This thing implies

$$\pi P_l = \pi, \forall l \in \langle n-1 \rangle,$$

and, further,

$$\pi P = \pi.$$

Finally, we have

$$\pi = \pi P = \pi e'\psi = \psi,$$

so,

$$P = e'\pi. \quad \square$$

We comment on Theorem 2.1 and its proof.

1. Any 1-step of the chain with transition matrix $P = P_1P_2...P_{n-1}$ is performed via P_1, P_2, ..., P_{n-1}, *i.e.*, doing $n-1$ transitions: one using P_1, one using P_2, ..., one using P_{n-1}. This chain: a) attains its stationarity at time 1 — one step due to P or $n-1$ steps due to P_1, P_2, ..., P_{n-1}; b) belongs to our collection of hybrid Metropolis-Hastings chains from [10] (this follows from $L_{\tau_1,\tau_2,...,\tau_{l+1}} \subset L_{\tau_1,\tau_2,...,\tau_l}$, $\forall l \in \langle n-2 \rangle$, $\forall \tau_1 \in C_n$, $\forall \tau_2 \in C_{n,g(\tau_1)}, ...,$ $\forall \tau_{l+1} \in C_{n,g(\tau_1),g(\tau_2),...,g(\tau_l)}$, where $n \geq 3$ (hint: use Theorem 1.6), etc.; for our collection, see also [11–12]); c) is a cyclic Gibbs sampler in a generalized sense because the state space is, here, $\mathbb{S}_n$, ratios used to define the transition probabilities of matrices P_l, $l \in \langle n-1 \rangle$, are similar to those of (usual) cyclic Gibbs sampler with — not leaving the finite framework — finite state space (this chain also belongs to our collection of hybrid Metropolis-Hastings chains from [10], see [11]), and matrices P_1, P_2, ..., P_{n-1} are used cyclically. (For finite Markov chain theory, see, *e.g.*, [5], and for the Gibbs sampler, see, *e.g.*, [6].)

2. To define transition probabilities of P_l, l is fixed ($l \in \langle n-1 \rangle$), we used states from $\mathbb{A}_{n,l}$. So, using P_l, the chain passes from a state, say, γ of $\mathbb{A}_{n,l}$ to a state, say, δ of $\mathbb{A}_{n,l}$ also. For the next matrix, P_{l+1}, when $l+1 \leq n-1$, we need states from $\mathbb{A}_{n,l+1}$, so, when we run the chain, we must rewrite δ using the generators of $\mathbb{A}_{n,l+1}$.

3. There exists a case, a happy case, for which rewriting the states from Comment 2 is not, practically speaking (see below), necessary, namely, when $\sigma_l = \mathrm{Id}$. So, to avoid rewriting the states, we consider the chain with initial state σ_0 (warning! σ_0 is here and σ_l is in the previous sentence). In this case, $(p_0)_{\sigma_0} = 1$, p_0 is the initial probability distribution of chain. Obviously, $p_0 P = \pi$ because $P = e'\pi$. (Moreover,

$$p_0 P^m = \pi, \ \forall m \geq 1, \forall p_0, \ p_0 = \text{ initial probability distribution.)}$$

From

$$\sigma_0 = (1) \circ \mathrm{Id} \circ \sigma_0 \in \mathbb{A}_{n,1}$$

($\sigma_1 = \mathrm{Id}$), the chain passes in one of the states — the rule is very simple —

$$\begin{gathered}\sigma_0 = (1) \circ \sigma_0 = (1) \circ \mathrm{Id} \circ \sigma_0 \in \mathbb{A}_{n,1}, \\ (2,1) \circ \sigma_0 = (2,1) \circ \mathrm{Id} \circ \sigma_0 \in \mathbb{A}_{n,1}, \\ \vdots \\ (n, n-1, ..., 1) \circ \sigma_0 = (n, n-1, ..., 1) \circ \mathrm{Id} \circ \sigma_0 \in \mathbb{A}_{n,1}.\end{gathered}$$

Suppose that it passed in the state $(3,2,1) \circ \sigma_0$. From

$$(3,2,1) \circ \sigma_0 = (1) \circ (3,2,1) \circ \mathrm{Id} \circ \sigma_0 \in \mathbb{A}_{n,2}$$

($\sigma_2 = \mathrm{Id}$), the chain passes in one of the states

$$\begin{gathered}(3,2,1) \circ \sigma_0 = (1) \circ (3,2,1) \circ \sigma_0 = (1) \circ (3,2,1) \circ \mathrm{Id} \circ \sigma_0 \in \mathbb{A}_{n,2}, \\ (2,1) \circ (3,2,1) \circ \sigma_0 = (2,1) \circ (3,2,1) \circ \mathrm{Id} \circ \sigma_0 \in \mathbb{A}_{n,2}, \\ (4,2,1) \circ (3,2,1) \circ \sigma_0 = (4,2,1) \circ (3,2,1) \circ \mathrm{Id} \circ \sigma_0 \in \mathbb{A}_{n,2}, \\ \vdots \\ (n, n-1, ..., 4, 2, 1) \circ (3,2,1) \circ \sigma_0 = (n, n-1, ..., 4, 2, 1) \circ (3,2,1) \circ \mathrm{Id} \circ \sigma_0 \in \mathbb{A}_{n,2}.\end{gathered}$$

Suppose that it passed in the state $(5,4,2,1) \circ (3,2,1) \circ \sigma_0$, etc. For these transitions of the chain, we use the probability distributions

$$\left(\frac{1}{1+\theta+\theta^2+...+\theta^{n-1}}, \frac{\theta}{1+\theta+\theta^2+...+\theta^{n-1}}, \frac{\theta^2}{1+\theta+\theta^2+...+\theta^{n-1}}, ..., \frac{\theta^{n-1}}{1+\theta+\theta^2+...+\theta^{n-1}}\right)$$

(for P_1),

$$\left(\frac{1}{1+\theta+\theta^2+...+\theta^{n-2}}, \frac{\theta}{1+\theta+\theta^2+...+\theta^{n-2}}, \frac{\theta^2}{1+\theta+\theta^2+...+\theta^{n-2}}, ..., \frac{\theta^{n-2}}{1+\theta+\theta^2+...+\theta^{n-2}}\right)$$

(for P_2), ...,

$$\left(\frac{1}{1+\theta}, \frac{\theta}{1+\theta}\right)$$

(for P_{n-1}). These probability distributions have something in common with the geometric distribution. (For the geometric distribution, see, *e.g.*, [2, pp. 498–500].)

4. By $P = e'\pi$ we can compute the normalizing constant Z. Indeed, since $\mathbb{S}_n \supset L_{(1)} \supset L_{(1),(2)} \supset ... \supset L_{(1),(2),...,(n-1)} = \{\sigma_0\}$ (recall that, by Theorem 1.6, $L_{\tau_1,\tau_2,...,\tau_{l+1}} \subset L_{\tau_1,\tau_2,...,\tau_l}$, $\forall l \in \langle n-2 \rangle$, $\forall \tau_1 \in C_n$, $\forall \tau_2 \in C_{n,g(\tau_1)}$, ..., $\forall \tau_{l+1} \in C_{n,g(\tau_1),g(\tau_2),...,g(\tau_l)}$, where $n \geq 3$), P_l is a block diagonal matrix (eventually by permutation of rows and columns), $\forall l \in \langle n-1 \rangle - \{1\}$, and $P_l \in G_{\Delta_l,\Delta_{l+1}}$, $\forall l \in \langle n-1 \rangle$ (moreover, P_l is a Δ_l-stable matrix on Δ_{l+1}, $\forall l \in \langle n-1 \rangle$), we have (using $P = e'\pi$)

$$\pi_{\sigma_0} = \frac{1}{1+\theta+\theta^2+...+\theta^{n-1}} \cdot \frac{1}{1+\theta+\theta^2+...+\theta^{n-2}} \cdot ... \cdot \frac{1}{1+\theta}.$$

On the other hand,

$$\pi_{\sigma_0} = \frac{\theta^0}{Z} = \frac{1}{Z}.$$

So,

$$Z = (1+\theta)\left(1+\theta+\theta^2\right)...\left(1+\theta+\theta^2+...+\theta^{n-1}\right).$$

If $\theta \neq 1$, we have

$$Z = \frac{1-\theta^2}{1-\theta}\frac{1-\theta^3}{1-\theta}...\frac{1-\theta^n}{1-\theta}.$$

This result is known, see, *e.g.*, [3] (there, $\sigma_0 =$Id and $\theta \geq 1$ (using our notation)), but our computation method is new, probabilistic (Markovian), and interesting.

5. Using Uniqueness Theorem from [11] (the presentation of this result is too long, so, we omit to give it here), we can compute certain important probabilities of the Mallows model through Kendall metric. Indeed, by Uniqueness Theorem we have

$$P\left(L_{\tau_1}\right)=\sum_{\sigma\in L_{\tau_1}}\pi_\sigma=\frac{\theta^{\lfloor\tau_1\rceil-1}}{1+\theta+\theta^2+...+\theta^{n-1}},\ \forall\tau_1\in C_n$$

($\lfloor\tau_1\rceil$ is the length of τ_1). Note that

$$L_{\tau_1}=\left\{\sigma\,\middle|\,\sigma\in\mathbb{S}_n,\ \sigma\left(\sigma_0^{-1}(1)\right)=g(\tau_1)\right\},\ \forall\tau_1\in C_n$$

(L_{τ_1} is the set of permutations from $\mathbb{S}_n$, each permutation having the component i equal to $g(\tau_1)$, where $i=\sigma_0^{-1}(1)$ (therefore, $\sigma_0(i)=1$)). In particular, if $\sigma_0=$Id, we have

$$L_{\tau_1}=\{\sigma\,|\,\sigma\in\mathbb{S}_n,\ \sigma(1)=g(\tau_1)\},\ \forall\tau_1\in C_n$$

(L_{τ_1} is, here, the set of permutations from $\mathbb{S}_n$, each permutation having the component 1 equal to $g(\tau_1)$). Further, by Uniqueness Theorem we have

$$\frac{P\left(L_{\tau_1,\tau_2}\right)}{P\left(L_{\tau_1}\right)}=\frac{\sum\limits_{\sigma\in L_{\tau_1,\tau_2}}\pi_\sigma}{\sum\limits_{\sigma\in L_{\tau_1}}\pi_\sigma}=\frac{\theta^{\lfloor\tau_2\rceil-1}}{1+\theta+\theta^2+...+\theta^{n-2}},$$

$\forall\tau_1\in C_n,\forall\tau_2\in C_{n,g(\tau_1)}$, so,

$$P\left(L_{\tau_1,\tau_2}\right)=\begin{cases}\frac{1}{(1+\theta+\theta^2+...+\theta^{n-1})(1+\theta+\theta^2+...+\theta^{n-2})} & \text{if } \tau_1=(1),\ \tau_2=(2),\\ \frac{\theta^{\lfloor\tau_2\rceil-1}}{(1+\theta+\theta^2+...+\theta^{n-1})(1+\theta+\theta^2+...+\theta^{n-2})} & \text{if } \tau_1=(1),\ \tau_2\neq(2),\\ \frac{\theta^{\lfloor\tau_1\rceil-1}}{(1+\theta+\theta^2+...+\theta^{n-1})(1+\theta+\theta^2+...+\theta^{n-2})} & \text{if } \tau_1\neq(1),\ \tau_2=(1),\\ \frac{\theta^{\lfloor\tau_1\rceil+\lfloor\tau_2\rceil-2}}{(1+\theta+\theta^2+...+\theta^{n-1})(1+\theta+\theta^2+...+\theta^{n-2})} & \text{if } \tau_1\neq(1),\ \tau_2\neq(1).\end{cases}$$

Note that

$$L_{\tau_1,\tau_2}=\left\{\sigma\,\middle|\,\sigma\in\mathbb{S}_n,\ \sigma\left(\sigma_0^{-1}(1)\right)=g(\tau_1),\ \sigma\left(\sigma_0^{-1}(2)\right)=g(\tau_2)\right\},$$

$\forall\tau_1\in C_n$, $\forall\tau_2\in C_{n,g(\tau_1)}$. In particular, if $\sigma_0=$Id, we have

$$L_{\tau_1,\tau_2}=\{\sigma\,|\,\sigma\in\mathbb{S}_n,\ \sigma(1)=g(\tau_1),\ \sigma(2)=g(\tau_2)\},$$

$\forall\tau_1\in C_n$, $\forall\tau_2\in C_{n,g(\tau_1)}$. To compute $P\left(L_{\tau_1,\tau_2,\tau_3}\right)$, etc., we use (see Uniqueness Theorem)

$$\frac{P\left(L_{\tau_1,\tau_2,...,\tau_u}\right)}{P\left(L_{\tau_1,\tau_2,...,\tau_{u-1}}\right)}=\frac{\sum\limits_{\sigma\in L_{\tau_1,\tau_2,...,\tau_u}}\pi_\sigma}{\sum\limits_{\sigma\in L_{\tau_1,\tau_2,...,\tau_{u-1}}}\pi_\sigma}=\frac{\theta^{\lfloor\tau_u\rceil-1}}{1+\theta+\theta^2+...+\theta^{n-u}},$$

$\forall \tau_1 \in C_n$, $\forall \tau_2 \in C_{n,g(\tau_1)}$, ..., $\forall \tau_u \in C_{n,g(\tau_1),g(\tau_2),\ldots,g(\tau_{u-1})}$ $(3 \leq u \leq n-1)$.

6. $\sigma_0 =$Id is the best case of our sampling method. To illustrate this fact, we consider $\sigma_0 = (12345)$. Consider that the initial state of chain is σ_0. From (12345), since

$$C_5 = \{(1), \ (2,1), \ (3,2,1), \ (4,3,2,1), \ (5,4,3,2,1)\},$$

the chain passes in one of the states

$$\sigma_0 = (12345), \ (21345), \ (31245), \ (41235), \ (51234)$$

(see Comment 3 or, direct, Theorem 2.1; see also Theorem 1.6), the transition probabilities being

$$\frac{1}{1+\theta+\theta^2+\theta^3+\theta^4}, \ \frac{\theta}{1+\theta+\theta^2+\theta^3+\theta^4},$$
$$\frac{\theta^2}{1+\theta+\theta^2+\theta^3+\theta^4}, \frac{\theta^3}{1+\theta+\theta^2+\theta^3+\theta^4}, \frac{\theta^4}{1+\theta+\theta^2+\theta^3+\theta^4},$$

respectively (see Theorem 2.1). This sequence of probabilities is decreasing when $0 < \theta \leq 1$ (recall that $0 < \theta \leq 1$ is the classical case) and increasing when $\theta > 1$. Note that (21345) is obtained from (12345) modifying the first two components of (12345), (31245) is obtained from (12345) modifying the first three components of (12345), etc. Suppose that the chain passed in the state (31245). $(31245) = (3,2,1) \circ (12345)$, so, $\tau_1 = (3,2,1)$ and $g(\tau_1) = 3$. In this case, to pass in the next state, we use

$$\tau_2 \in C_{5,3} = \{(1), \ (2,1), \ (4,2,1), \ (5,4,2,1)\}.$$

$g(\tau_2) \in \{1,2,4,5\}$ — the elements from this set are placed after 3 in (31245) (this fact is important for the implementation of our method in the special case $\sigma_0 =$Id). From (31245), the chain passes in one of the states

$$(31245), \ (32145), \ (34125), \ (35124),$$

the transition probabilities being

$$\frac{1}{1+\theta+\theta^2+\theta^3}, \ \frac{\theta}{1+\theta+\theta^2+\theta^3}, \ \frac{\theta^2}{1+\theta+\theta^2+\theta^3}, \ \frac{\theta^3}{1+\theta+\theta^2+\theta^3},$$

respectively. This sequence of probabilities is also decreasing when $0 < \theta \leq 1$ and increasing when $\theta > 1$. Note that (32145) is obtained from (31245) modifying the first two components after 3 of (31245), (34125) is obtained from (31245) modifying the first three components after 3 of (31245), etc. Suppose that the chain passed in the state (34125). $(34125) = (4,2,1) \circ (31245)$, so, $\tau_2 = (4,2,1)$ and $g(\tau_2) = 4$. In this case, to pass in the next state, we use

$$\tau_3 \in C_{5,3,4} = \{(1), \ (2,1), \ (5,2,1)\}.$$

$g(\tau_3) \in \{1,2,5\}$ — the elements from this set are placed after 3 and 4 in (34125). Etc. Now, we consider the general case $\sigma_0 \in \mathbb{S}_n$. Since $(\mathbb{S}_n, \circ)$ is a group and K is right-invariant, we have

$$K(\sigma, \mathrm{Id}) = K(\sigma \circ \sigma_0, \sigma_0), \ \forall \sigma \in \mathbb{S}_n,$$

so,

$$\sum_{\sigma \in \mathbb{S}_n} \theta^{K(\sigma,\mathrm{Id})} = \sum_{\sigma \in \mathbb{S}_n} \theta^{K(\sigma\circ\sigma_0,\sigma_0)}$$

(the terms of this equation are normalizing constants) and

$$\frac{\theta^{K(\sigma,\mathrm{Id})}}{\sum\limits_{\sigma \in \mathbb{S}_n} \theta^{K(\sigma,\mathrm{Id})}} = \frac{\theta^{K(\sigma\circ\sigma_0,\sigma_0)}}{\sum\limits_{\sigma \in \mathbb{S}_n} \theta^{K(\sigma\circ\sigma_0,\sigma_0)}}$$

(the terms of this equation are probabilities), $\forall \sigma \in \mathbb{S}_n$. It follows that the general case $\sigma_0 \in \mathbb{S}_n$ can easily be obtained from the simple case $\sigma_0 =$Id using the bijective transformation $\sigma \longmapsto \sigma \circ \sigma_0$ — if, using the simple case $\sigma_0 =$Id, we generate a permutation, say, χ, then $\chi \circ \sigma_0$ is the result of generation for the general case $\sigma_0 \in \mathbb{S}_n$. So, for any $\sigma_0 \in \mathbb{S}_n$, our exact sampling method is simple and fast — faster when $\sigma_0 =$Id than when $\sigma_0 \neq$Id, faster when $0 < \theta \leq 1$ than when $\theta > 1$.

In [3], it is presented a method for sampling from $\mathbb{S}_n$ according to the Mallows model through Kendall metric when $\sigma_0 =$Id and $\theta \geq 1$ (using our notation).

Finally, we give an example to illustrate Theorem 2.1, its proof, and the above comments, excepting Comment 6 (this contains an example for itself).

Example 2.2. Consider the Mallows model through Kendall metric on $\mathbb{S}_3$ with $\sigma_0 = (312)$. By Theorem 2.1 (and its proof) we have

$$P_1 = \begin{array}{c} \\ (123) \\ (321) \\ (132) \\ (231) \\ (213) \\ (312) \end{array} \begin{array}{c} \begin{array}{cccccc} (123) & (321) & (132) & (231) & (213) & (312) \end{array} \\ \left(\begin{array}{cccccc} \frac{\theta}{1+\theta+\theta^2} & 0 & \frac{\theta^2}{1+\theta+\theta^2} & 0 & \frac{1}{1+\theta+\theta^2} & 0 \\ 0 & \frac{\theta}{1+\theta+\theta^2} & 0 & \frac{\theta^2}{1+\theta+\theta^2} & 0 & \frac{1}{1+\theta+\theta^2} \\ \frac{\theta}{1+\theta+\theta^2} & 0 & \frac{\theta^2}{1+\theta+\theta^2} & 0 & \frac{1}{1+\theta+\theta^2} & 0 \\ 0 & \frac{\theta}{1+\theta+\theta^2} & 0 & \frac{\theta^2}{1+\theta+\theta^2} & 0 & \frac{1}{1+\theta+\theta^2} \\ \frac{\theta}{1+\theta+\theta^2} & 0 & \frac{\theta^2}{1+\theta+\theta^2} & 0 & \frac{1}{1+\theta+\theta^2} & 0 \\ 0 & \frac{\theta}{1+\theta+\theta^2} & 0 & \frac{\theta^2}{1+\theta+\theta^2} & 0 & \frac{1}{1+\theta+\theta^2} \end{array} \right) \end{array}$$

and

$$P_2 = \begin{array}{c} \\ (123) \\ (321) \\ (132) \\ (231) \\ (213) \\ (312) \end{array} \begin{array}{c} \begin{array}{cccccc} (123) & (321) & (132) & (231) & (213) & (312) \end{array} \\ \left(\begin{array}{cccccc} \frac{\theta}{1+\theta} & \frac{1}{1+\theta} & & & & \\ \frac{\theta}{1+\theta} & \frac{1}{1+\theta} & & & & \\ & & \frac{\theta}{1+\theta} & \frac{1}{1+\theta} & & \\ & & \frac{\theta}{1+\theta} & \frac{1}{1+\theta} & & \\ & & & & \frac{\theta}{1+\theta} & \frac{1}{1+\theta} \\ & & & & \frac{\theta}{1+\theta} & \frac{1}{1+\theta} \end{array} \right) \end{array}$$

because, for P_1,

$$(123) = (2,1) \circ (2,3) \circ (312) \in \mathbb{A}_{3,1},$$
$$(321) = (2,1) \circ \mathrm{Id} \circ (312) \in \mathbb{A}_{3,1},$$
$$(132) = (3,2,1) \circ (2,3) \circ (312) \in \mathbb{A}_{3,1},$$
$$(231) = (3,2,1) \circ \mathrm{Id} \circ (312) \in \mathbb{A}_{3,1},$$
$$(213) = (1) \circ (2,3) \circ (312) \in \mathbb{A}_{3,1},$$
$$(312) = (1) \circ \mathrm{Id} \circ (312) \in \mathbb{A}_{3,1}$$

and, for P_2,

$$(123) = (3,1) \circ (2,1) \circ \mathrm{Id} \circ (312) \in \mathbb{A}_{3,2},$$
$$(321) = (1) \circ (2,1) \circ \mathrm{Id} \circ (312) \in \mathbb{A}_{3,2},$$
$$(132) = (2,1) \circ (3,2,1) \circ \mathrm{Id} \circ (312) \in \mathbb{A}_{3,2},$$
$$(231) = (1) \circ (3,2,1) \circ \mathrm{Id} \circ (312) \in \mathbb{A}_{3,2},$$
$$(213) = (3,2) \circ (1) \circ \mathrm{Id} \circ (312) \in \mathbb{A}_{3,2},$$
$$(312) = (2) \circ (1) \circ \mathrm{Id} \circ (312) \in \mathbb{A}_{3,2}.$$

Further, we have

$$L_{(1)} = \{(213), (312)\}, \ L_{(2,1)} = \{(123), (321)\}, \ L_{(3,2,1)} = \{(132), (231)\}$$

(the permutations from $L_{(1)}$, $L_{(2,1)}$, $L_{(3,2,1)}$ have the component 2 equal to 1, 2, 3, respectively),

$$L_{(1),(2)} = \{(312)\}, \ L_{(1),(3,2)} = \{(213)\},$$
$$L_{(2,1),(1)} = \{(321)\}, \ L_{(2,1),(3,1)} = \{(123)\},$$
$$L_{(3,2,1),(1)} = \{(231)\}, \ L_{(3,2,1),(2,1)} = \{(132)\},$$
$$\Delta_1 = (\mathbb{S}_3),$$
$$\Delta_2 = \left(L_{(1)}, \ L_{(2,1)}, \ L_{(3,2,1)}\right),$$

$$\Delta_3 = \left(L_{(1),(2)},\ L_{(1),(3,2)}, L_{(2,1),(1)},\ L_{(2,1),(3,1)},\ L_{(3,2,1),(1)},\ L_{(3,2,1),(2,1)}\right).$$

Obviously, $\Delta_3 = (\{\sigma\})_{\sigma\in\mathbb{S}_3}$ and, moreover, $\Delta_3 \preceq \Delta_2 \preceq \Delta_1$, $\Delta_3 \neq \Delta_2 \neq \Delta_1$. It is easy to see that $P_1 \in G_{\Delta_1,\Delta_2}$, $P_2 \in G_{\Delta_2,\Delta_3}$, and $\pi_\sigma (P_l)_{\sigma\tau} = \pi_\tau (P_l)_{\tau\sigma}$, $\forall l \in \langle 2\rangle$, $\forall \sigma,\tau \in \mathbb{S}_3$. By Theorem 2.1 or direct computation, $P = e'\pi$. Since $\pi_{\sigma_0} = \frac{1}{Z}$, it is easy to see, using $P = e'\pi$, that $Z = (1+\theta)\left(1+\theta+\theta^2\right)$. Obviously, P_2 is a block diagonal matrix and Δ_2-stable matrix on Δ_2. Moreover, P_2 is a Δ_2-stable matrix, see Definition 1.4. P_1 is a stable matrix both on Δ_1 and on Δ_2. By Uniqueness Theorem from [11] or direct computation we have

$$P\left(L_{(1)}\right) = \frac{1}{1+\theta+\theta^2},\ P\left(L_{(2,1)}\right) = \frac{\theta}{1+\theta+\theta^2},\ P\left(L_{(3,2,1)}\right) = \frac{\theta^2}{1+\theta+\theta^2},$$
$$P\left(L_{(1),(2)}\right) = \frac{1}{(1+\theta+\theta^2)(1+\theta)},\ P\left(L_{(1),(3,2)}\right) = \frac{\theta}{(1+\theta+\theta^2)(1+\theta)},$$
$$P\left(L_{(2,1),(1)}\right) = \frac{\theta}{(1+\theta+\theta^2)(1+\theta)},\ P\left(L_{(2,1),(3,1)}\right) = \frac{\theta^2}{(1+\theta+\theta^2)(1+\theta)},$$
$$P\left(L_{(3,2,1),(1)}\right) = \frac{\theta^2}{(1+\theta+\theta^2)(1+\theta)},\ P\left(L_{(3,2,1),(2,1)}\right) = \frac{\theta^3}{(1+\theta+\theta^2)(1+\theta)}.$$

Obviously, our exact sampling Markovian method has, here, 2 steps (due to P_1 and P_2).

"We conclude saying that we believe that the homogeneous Markov chain framework is too narrow to design fast algorithms of Metropolis-Hastings type. Moreover, we believe that our hybrid Metropolis-Hastings chain works better than the Metropolis-Hastings chain, at least on $\mathbb{S}_n$, the set of permutations of order n, and on $\{0,1,...,h\}^n$." We said this in [10, p. 227]. We say this again because, being guided by [9, 10] and, especially, [11], the achievements from [12] and this article of our hybrid Metropolis-Hastings chain are impressive: fast exact samplings, computing the normalizing constants (exactly), computing certain important probabilities. The Metropolis-Hastings chain cannot do these things – we think so.

REFERENCES

[1] D.E. Critchlow, *Metric Methods for Analyzing Partially Ranked Data.* Lecture Notes in Statistics, Vol. **34**, Springer-Verlag, Berlin, 1985.

[2] L. Devroye, *Non-Uniform Random Variate Generation.* Springer-Verlag, New York, 1986; available at http://cg.scs.carleton.ca/~luc/rnbookindex.html.

[3] P. Diaconis and A. Ram, *Analysis of systematic scan Metropolis algorithms using Iwahori-Hecke algebra techniques.* Michigan Math. J. **48** (2000), 157–190.

[4] M.A. Fligner and J.S. Verducci (Eds.), *Probability Models and Statistical Analyses for Ranking Data.* Lecture Notes in Statistics, Vol. **80**, Springer-Verlag, New York, 1993.

[5] M. Iosifescu, *Finite Markov Processes and Their Applications.* Wiley, Chichester & Edit. Tehnică, Bucharest, 1980; corrected republication by Dover, Mineola, N.Y., 2007.

[6] N. Madras, *Lectures on Monte Carlo Methods.* AMS, Providence, Rhode Island, 2002.

[7] C.L. Mallows, *Non-null ranking models*, I. Biometrika **44** (1957), 114–130.

[8] J.I. Marden, *Analyzing and Modeling Rank Data.* Chapman & Hall, London, 1995.

[9] U. Păun, G_{Δ_1,Δ_2} *in action.* Rev. Roumaine Math. Pures Appl. **55** (2010), 387–406.

[10] U. Păun, *A hybrid Metropolis-Hastings chain.* Rev. Roumaine Math. Pures Appl. **56** (2011), 207–228.

[11] U. Păun, *G method in action*: *from exact sampling to approximate one.* Rev. Roumaine Math. Pures Appl. **62** (2017), 413–452.

[12] U. Păun, *G method in action*: *fast exact sampling from set of permutations of order n according to Mallows model through Cayley metric.* Braz. J. Probab. Stat. **31** (2017), 338–352.

Received 23 May 2017

Romanian Academy,
"Gheorghe Mihoc – Caius Iacob" Institute
of Mathematical Statistics and Applied
Mathematics,
Calea 13 Septembrie nr. 13,
050711 Bucharest 5, Romania
paun@csm.ro