Annals of the University of Craiova, Mathematics and Computer Science Series
Volume 45(1), 2018, Pages 103–121
ISSN: 1223-6934

A Gibbs sampler in a generalized sense, II

Udrea Păun

Abstract. We consider two new conditions for finite sequences of (finite) stochastic matrices. The Gibbs samplers in a generalized sense which satisfy these conditions have important properties, and thus became among the first our favorite chains — our interest is to design very fast Markov chains and having, if possible, other important properties. We show, in the finite case, that the probability distribution of a random vector with independent components is a wavy probability distribution with respect to the lexicographic order and $n+1$ partitions which will be specified, where n is the dimension of random vector. We define the wavy probability distributions in a generalized sense. When these probability distributions have normalization constant, we give, under certain conditions, a formula to compute this constant. To give other examples of wavy probability distributions and of wavy probability distributions in a generalized sense, we consider the Potts model (a model used in statistical physics and other fields). Moreover, the normalization constant for the Ising model on $\mathcal{C}_n$, the cycle graph with n vertices, is computed.

2010 Mathematics Subject Classification. 60J10, 15B51, 60E05, 82B20.
Key words and phrases. Gibbs sampler in a generalized sense, wavy probability distribution, wavy probability distribution in a generalized sense, normalization constant, Potts model, Ising model.

1. A short introduction

The reader is assumed to be acquainted with [16].

The Metropolis chain (1953) [9], its generalization, the Metropolis-Hastings chain (1970) [3], and the Gibbs sampler (1984) [2] are Markov chains which were created to simulate random variables. As far as we know, no application in which the exact sampling holds was obtained for the first two chains. Moreover, as far as we know too, for the speeds of convergence of these three chains — these speeds are the most important things on these chains —, only in some special cases were obtained some results, not very good results. However, some people are very enthusiastic, such as, the authors of [18]. (For the theory of Markov chains, see, *e.g.*, [4], and for the Metropolis-Hastings chain and Gibbs sampler, besides [2]-[3] and [9], see, *e.g.*, also [6].)

Our hybrid Metropolis-Hastings chain (2011) [11] is also a Markov chain which was created to simulate random variables. The construction of this chain was suggested by some recent enough results from the theory of Markov chains from [10] (see also [12]). The Gibbs sampler (the cyclic Gibbs sampler, see [16]) and its generalization from [16], the Gibbs sampler in a generalized sense, are two important cases of our hybrid Metropolis-Hastings chain(s). Four applications in which the exact sampling

Received February 23, 2018.

holds were obtained, one for the Gibbs sampler in [12, Application 3.5] and three for the Gibbs sampler in a generalized sense in [13]−[15]. Moreover, besides the exact sampling, other important things were obtained in [13]−[15]: closed-form expressions for normalization constants (in [13]−[15]), bounds for normalization constants (in [15]), etc.

So, we have sufficient reasons to continue our work on the hybrid Metropolis-Hastings chain, Gibbs sampler in a generalized sense, wavy probability distributions, ...

2. New conditions for finite sequences of stochastic matrices

In this section, we present two new conditions for finite sequences of stochastic matrices, which can be used as special conditions for the hybrid Metropolis-Hastings chain(s). These conditions lead to some important results, the most important being one on the Gibbs sampler(s) in a generalized sense — we find an important subcollection of the collection of Gibbs samplers in a generalized sense. This subcollection is (the two conditions are also) in conjunction with a main problem, a problem of interest to us: finding the fastest Gibbs samplers in a generalized sense.

Let $Q_1, Q_2, ..., Q_t \in S_r$ ($r, t \geq 1$). In [11] (see also [12]) we considered four special conditions for our hybrid Metropolis-Hastings chain denoted (c1), (c2), (c3), (c4) (the conditions (c1), (c2), and (c4) are also presented in [16]). Due this fact, the conditions presented below are denoted (c5), (c6).

(c5) $(Q_l)_{ii} > 0, \forall l \in \langle t \rangle, \forall i \in \langle r \rangle$.

(c6) $\forall l \in \langle t \rangle, \forall i, j, k \in \langle r \rangle, j \neq k$, if $(Q_l)_{ij} > 0$ and $(Q_l)_{ik} > 0$, then $(Q_l)_{jk} > 0$.

These new conditions imply the basic condition (C1) of hybrid Metropolis-Hastings chain from [11] ((C1) is also presented in [16]).

Theorem 2.1. *Let $Q_1, Q_2, ..., Q_t \in S_r$. If* (c5) *and* (c6) *hold, then* (C1) *holds.*

Proof. We must show that (see [16]) $\overline{Q}_l$, the incidence matrix of Q_l, is a symmetric matrix, $\forall l \in \langle t \rangle$ — this is equivalent to

$$\forall l \in \langle t \rangle, \ \forall i, j \in \langle r \rangle, \ i \neq j, \ [(Q_l)_{ij} > 0 \Longleftrightarrow (Q_l)_{ji} > 0].$$

Let $l \in \langle t \rangle$. Let $i, j \in \langle r \rangle, i \neq j$.

"$\Longrightarrow$" By (c5), $(Q_l)_{ii} > 0$. By (c6), $(Q_l)_{ij} > 0$ and $(Q_l)_{ii} > 0$ imply

$$(Q_l)_{ji} > 0.$$

"$\Longleftarrow$" By (c5), $(Q_l)_{jj} > 0$. By (c6), $(Q_l)_{ji} > 0$ and $(Q_l)_{jj} > 0$ imply

$$(Q_l)_{ij} > 0.$$

□

Let $A \in N_{m,n}$ (see [16] for $N_{m,n}$). Let $i \in \langle m \rangle$. Set

$$\mathcal{N}_{A,i} = \{j \mid j \in \langle n \rangle \text{ and } A_{ij} > 0\}.$$

Below we give another basic result on the structure of matrices $Q_1, Q_2, ..., Q_t$ when the conditions (c5) and (c6) hold.

Theorem 2.2. *Let $Q_1, Q_2, ..., Q_t \in S_r$. Suppose that* (c5) *and* (c6) *hold. Let $l \in \langle t \rangle$. Let $i \in \langle r \rangle$. Then*

(i)

$$(Q_l)_{\mathcal{N}_{Q_l,i}}^{\mathcal{N}_{Q_l,i}} > 0$$

$((Q_l)_{\mathcal{N}_{Q_l,i}}^{\mathcal{N}_{Q_l,i}}$ *is a matrix, see* [16]; *by* (c5), $i \in \mathcal{N}_{Q_l,i}$, *so* $\mathcal{N}_{Q_l,i} \neq \emptyset$);

(ii)

$$(Q_l)_{\mathcal{N}_{Q_l,i}}^{\mathcal{N}_{Q_l,i}^c} = 0 \ \ \textit{if} \ \mathcal{N}_{Q_l,i}^c \neq \emptyset$$

$((Q_l)_{\mathcal{N}_{Q_l,i}}^{\mathcal{N}_{Q_l,i}^c}$ *is also a matrix*; $\mathcal{N}_{Q_l,i}^c$ *is the complement of* $\mathcal{N}_{Q_l,i}$);

(iii)

$$(Q_l)_{\mathcal{N}_{Q_l,i}^c}^{\mathcal{N}_{Q_l,i}} = 0 \ \ \textit{if} \ \mathcal{N}_{Q_l,i}^c \neq \emptyset;$$

(iv)

$$\mathcal{N}_{Q_l,j} = \mathcal{N}_{Q_l,i}, \ \forall j \in \mathcal{N}_{Q_l,i}.$$

Proof. (i) Obviously,

$$(Q_l)_{\mathcal{N}_{Q_l,i}}^{\mathcal{N}_{Q_l,i}} > 0 \Longleftrightarrow (Q_l)_{jk} > 0, \ \forall j, k \in \mathcal{N}_{Q_l,i}.$$

Let j, $k \in \mathcal{N}_{Q_l,i}$. We show that

$$(Q_l)_{jk} > 0.$$

Case 1. $j = k$. By (c5).

Case 2. $j \neq k$. Recall that, by (c5), $i \in \mathcal{N}_{Q_l,i}$.

Subcase 2.1. $j = i$. By $k \in \mathcal{N}_{Q_l,i}$ we have $(Q_l)_{ik} > 0$. Since $j = i$, we have

$$(Q_l)_{jk} = (Q_l)_{ik} > 0.$$

Subcase 2.2. $k = i$. By $j \in \mathcal{N}_{Q_l,i}$ we have $(Q_l)_{ij} > 0$. By (c5), $(Q_l)_{ii} > 0$. By $j \neq k$ and $k = i$ we have $j \neq i$. By (c6), since $(Q_l)_{ij} > 0$, $(Q_l)_{ii} > 0$, and $j \neq i$, we have $(Q_l)_{ji} > 0$. Finally,

$$(Q_l)_{jk} = (Q_l)_{ji} > 0.$$

Subcase 2.3. j, $k \neq i$ (when $r \geq 3$). By j, $k \in \mathcal{N}_{Q_l,i}$ we have $(Q_l)_{ij} > 0$, $(Q_l)_{ik} > 0$. So, using (c6), we have

$$(Q_l)_{jk} > 0.$$

(ii) Obviously, for $\mathcal{N}_{Q_l,i}^c \neq \emptyset$,

$$(Q_l)_{\mathcal{N}_{Q_l,i}}^{\mathcal{N}_{Q_l,i}^c} = 0 \Longleftrightarrow (Q_l)_{jk} = 0, \ \forall j \in \mathcal{N}_{Q_l,i}, \ \forall k \in \mathcal{N}_{Q_l,i}^c.$$

Let $j \in \mathcal{N}_{Q_l,i}$ and $k \in \mathcal{N}_{Q_l,i}^c \neq \emptyset$. We show that

$$(Q_l)_{jk} = 0.$$

Case 1. $j = i$. By $k \in \mathcal{N}_{Q_l,i}^c$ we have $(Q_l)_{ik} = 0$. So,

$$(Q_l)_{jk} = (Q_l)_{ik} = 0.$$

Case 2. $j \neq i$ (when $r \geq 3$). Since $i \in \mathcal{N}_{Q_l,i}$ (by (c5)) and $k \in \mathcal{N}_{Q_l,i}^c$, we have $i \neq k$. By (i) we have $(Q_l)_{ji} > 0$. Suppose that $(Q_l)_{jk} > 0$. From $(Q_l)_{ji} > 0$, $(Q_l)_{jk} > 0$, and $i \neq k$, using (c6), we have $(Q_l)_{ik} > 0$. It follows that $k \in \mathcal{N}_{Q_l,i}$. Contradiction.

(iii) Since, by Theorem 2.1, $\overline{Q}_l$ (the incidence matrix of Q_l) is a symmetric matrix, we have, for $\mathcal{N}^c_{Q_l,i} \neq \emptyset$,

$$(Q_l)^{\mathcal{N}_{Q_l,i}}_{\mathcal{N}^c_{Q_l,i}} = 0 \Longleftrightarrow (Q_l)^{\mathcal{N}^c_{Q_l,i}}_{\mathcal{N}_{Q_l,i}} = 0.$$

As, by (ii),

$$(Q_l)^{\mathcal{N}^c_{Q_l,i}}_{\mathcal{N}_{Q_l,i}} = 0 \text{ if } \mathcal{N}^c_{Q_l,i} \neq \emptyset,$$

it follows that

$$(Q_l)^{\mathcal{N}_{Q_l,i}}_{\mathcal{N}^c_{Q_l,i}} = 0 \text{ if } \mathcal{N}^c_{Q_l,i} \neq \emptyset.$$

(iv) Let $j \in \mathcal{N}_{Q_l,i}$. By (i) we have

$$(Q_l)^{\mathcal{N}_{Q_l,i}}_{\{j\}} > 0.$$

By (ii) we have

$$(Q_l)^{\mathcal{N}^c_{Q_l,i}}_{\{j\}} = 0 \text{ if } \mathcal{N}^c_{Q_l,i} \neq \emptyset.$$

Consequently,

$$\mathcal{N}_{Q_l,j} = \mathcal{N}_{Q_l,i}.$$

□

Below we give a corrected version of Theorem 2.1 from [16]. Although Theorem 2.1 from [16] is wrong, however, it contains the important case $P_l = Q_l$, $\forall l \in \langle t \rangle$, which was suggested by the applications from [12]-[15]. On the other hand, the mistake from this theorem was, fortunately, fruitful because, due to it, due to the important case $P_l = Q_l$, $\forall l \in \langle t \rangle$, we found the conditions (c5) and (c6).

Theorem 2.3. *Consider a hybrid Metropolis-Hastings chain with state space $S = \langle r \rangle$ and transition matrix $P = P_1 P_2 ... P_t$, P_1, P_2, ..., P_t corresponding to Q_1, Q_2, ..., Q_t, respectively. Suppose that $\forall l \in \langle t \rangle$, $\forall i, j \in S$,*

$$(Q_l)_{ij} = \frac{\pi_j}{\sum\limits_{k \in S,\ (Q_l)_{ik} > 0} \pi_k} \quad \textit{if } (Q_l)_{ij} > 0$$

(see Section 1 in [16] for Q_l, $l \in \langle t \rangle$, $\pi = (\pi_i)_{i \in S}$, ...). Then

$$(P_l)_{ij} = \begin{cases} 0 & \textit{if } j \neq i \textit{ and } (Q_l)_{ij} = 0, \\ (Q_l)_{ij} & \textit{if } j \neq i \textit{ and } \pi_j (Q_l)_{ji} \geq \pi_i (Q_l)_{ij} > 0, \\ \frac{\pi_j}{\sum\limits_{k \in S,\ (Q_l)_{jk} > 0} \pi_k} & \textit{if } j \neq i \textit{ and } \pi_j (Q_l)_{ji} < \pi_i (Q_l)_{ij}, \\ 1 - \sum\limits_{k \neq i} (P_l)_{ik} & \textit{if } j = i, \end{cases}$$

$\forall l \in \langle t \rangle$, $\forall i, j \in S$. If, moreover,

$$\pi_i (Q_l)_{ij} = \pi_j (Q_l)_{ji}, \ \forall l \in \langle t \rangle, \forall i, j \in S,$$

then

$$P_l = Q_l, \ \forall l \in \langle t \rangle.$$

Proof. Note that "$j \neq i$" is superfluous in "$j \neq i$ and $\pi_j (Q_l)_{ji} < \pi_i (Q_l)_{ij}$" because if $\pi_j (Q_l)_{ji} < \pi_i (Q_l)_{ij}$, then $j \neq i$. Moreover, $\pi_j (Q_l)_{ji} < \pi_i (Q_l)_{ij}$ implies $0 < \pi_j (Q_l)_{ji}$. Indeed, if $\pi_j (Q_l)_{ji} < \pi_i (Q_l)_{ij}$, then $(Q_l)_{ij} > 0$. By (C1) (in [16]), $(Q_l)_{ij} > 0$ implies $(Q_l)_{ji} > 0$. Finally, $(Q_l)_{ji} > 0$ implies $0 < \pi_j (Q_l)_{ji}$.

If $j \neq i$ and $\pi_j (Q_l)_{ji} < \pi_i (Q_l)_{ij}$ ($i, j \in S$, $l \in \langle t \rangle$), we have (see Section 1 in [16])

$$(P_l)_{ij} = (Q_l)_{ij} \min \left(1, \frac{\pi_j (Q_l)_{ji}}{\pi_i (Q_l)_{ij}} \right) = (Q_l)_{ij} \cdot \frac{\pi_j (Q_l)_{ji}}{\pi_i (Q_l)_{ij}} =$$

$$= \frac{\pi_j}{\pi_i} \cdot (Q_l)_{ji} = \frac{\pi_j}{\pi_i} \cdot \frac{\pi_i}{\sum\limits_{k \in S,\ (Q_l)_{jk} > 0} \pi_k} = \frac{\pi_j}{\sum\limits_{k \in S,\ (Q_l)_{jk} > 0} \pi_k}.$$

The others are obvious. □

The hybrid Metropolis-Hastings chain from Theorem 2.1 in [16] (here, this chain is in Theorem 2.3) was called the *cyclic Gibbs sampler in a generalized sense* — the *Gibbs sampler in a generalized sense* for short. Therefore, the Gibbs sampler in a generalized sense is a hybrid Metropolis-Hastings chain having the property: $\forall l \in \langle t \rangle$, $\forall i, j \in S$,

$$(Q_l)_{ij} = \frac{\pi_j}{\sum\limits_{k \in S,\ (Q_l)_{ik} > 0} \pi_k} \text{ if } (Q_l)_{ij} > 0.$$

Using the sets $\mathcal{N}_{Q_l, i}$, $l \in \langle t \rangle$, $i \in S$, this property can be written differently: $\forall l \in \langle t \rangle$, $\forall i, j \in S$,

$$(Q_l)_{ij} = \frac{\pi_j}{\sum\limits_{k \in \mathcal{N}_{Q_l, i}} \pi_k} \text{ if } j \in \mathcal{N}_{Q_l, i}.$$

Remark 2.1. (a) For our hybrid Metropolis-Hastings chain (in particular, for our Gibbs sampler in a generalized sense) with state space $S = \langle r \rangle$ and transition matrix $P = P_1 P_2 ... P_t$, P_1, P_2, ..., P_t corresponding to Q_1, Q_2, ..., Q_t, respectively, we have

$$|\mathcal{N}_{P_l, i}| \geq |\mathcal{N}_{Q_l, i}|, \ \forall l \in \langle t \rangle, \ \forall i \in S$$

($|\cdot|$ is the cardinal), more precisely, we have

$$\text{either } |\mathcal{N}_{P_l, i}| = |\mathcal{N}_{Q_l, i}| \text{ or } |\mathcal{N}_{P_l, i}| = |\mathcal{N}_{Q_l, i}| + 1$$

(the latter equation holds when $(Q_l)_{ii} = 0$ while $(P_l)_{ii} > 0$), $\forall l \in \langle t \rangle$, $\forall i \in S$. The nearer the value $|\mathcal{N}_{P_l, i}|$ is to the value $|\mathcal{N}_{Q_l, i}|$, $\forall l \in \langle t \rangle$, $\forall i \in S$, the faster the sampling is. So, the best possible case for sampling is when

$$|\mathcal{N}_{P_l, i}| = |\mathcal{N}_{Q_l, i}|, \ \forall l \in \langle t \rangle, \ \forall i \in S.$$

(b) The case when the Gibbs sampler in a generalized sense has the property that $P_l = Q_l$, $\forall l \in \langle t \rangle$, is an important one due to the following things.

(b1) When we run the chain, some of the values $(P_l)_{ij}$, $l \in \langle t \rangle$, $i, j \in S$, are computed — these values will be computed using the simple formula

$$(P_l)_{ij} = \begin{cases} 0 & \text{if } (Q_l)_{ij} = 0, \\ \frac{\pi_j}{\sum\limits_{k \in S,\ (Q_l)_{ik} > 0} \pi_k} & \text{if } (Q_l)_{ij} > 0, \end{cases}$$

$\forall l \in \langle t \rangle$, $\forall i, j \in S$.

(b2)

$$|\mathcal{N}_{P_l,i}| = |\mathcal{N}_{Q_l,i}|, \ \forall l \in \langle t \rangle, \ \forall i \in S.$$

See (a) again now.

(b3)

$$(P_l)_{ij} = \frac{\pi_j}{\sum\limits_{k \in S, \ (P_l)_{ik} > 0} \pi_k} \text{ if } (P_l)_{ij} > 0$$

$(l \in \langle t \rangle, \ i, j \in S)$, so, $(P_l)_{ij}$ and π_j are directly proportional when $(P_l)_{ij} > 0$.

Due to Remark 2.1 and [12]−[16], the subcollection of Gibbs samplers in a generalized sense with $P_l = Q_l$, $\forall l \in \langle t \rangle$, is among the greatest subcollections of our collection of hybrid Metropolis-Hastings chains. We have a Gibbs sampler in a generalized sense belonging to this subcollection in Theorem 3.1 from [16], which has four applications in [12]−[15] (in [12, Application 3.5], we even have a Gibbs sampler). (In Theorem 3.1 from [16], the matrices Q_l, $l \in \langle t \rangle$, of Gibbs sampler in a generalized sense from there do not appear — take $Q_l = P_l$, $\forall l \in \langle t \rangle$.)

Related to the above considerations, below we give a result, the main result of this section.

Theorem 2.4. *Consider a Gibbs sampler in a generalized sense with state space* $S = \langle r \rangle$ *and transition matrix* $P = P_1 P_2 ... P_t$, P_1, P_2, ..., P_t *corresponding to* Q_1, Q_2, ..., Q_t, *respectively. If* Q_1, Q_2, ..., Q_t *satisfy the conditions* (c5) *and* (c6)*, then*

$$P_l = Q_l, \ \forall l \in \langle t \rangle.$$

Proof. Let $l \in \langle t \rangle$. Since P_l and Q_l are stochastic matrices, it is sufficient to show that

$$(P_l)_{ij} = (Q_l)_{ij}, \ \forall i, j \in S \text{ with } i \neq j \text{ and } (Q_l)_{ij} > 0.$$

Let i, $j \in S$ with $i \neq j$ and $(Q_l)_{ij} > 0$. We have

$$(Q_l)_{ij} = \frac{\pi_j}{\sum\limits_{k \in S, \ (Q_l)_{ik} > 0} \pi_k}.$$

Since $\overline{Q}_l$ (the incidence matrix of Q_l) is symmetric, we have $(Q_l)_{ji} > 0$, so,

$$(Q_l)_{ji} = \frac{\pi_i}{\sum\limits_{k \in S, \ (Q_l)_{jk} > 0} \pi_k}.$$

Since $(Q_l)_{ij} > 0$, we have $j \in \mathcal{N}_{Q_l,i}$. By Theorem 2.2(iv) we have $\mathcal{N}_{Q_l,i} = \mathcal{N}_{Q_l,j}$. It follows that

$$\sum_{k \in S, \ (Q_l)_{ik} > 0} \pi_k = \sum_{k \in \mathcal{N}_{Q_l,i}} \pi_k = \sum_{k \in \mathcal{N}_{Q_l,j}} \pi_k = \sum_{k \in S, \ (Q_l)_{jk} > 0} \pi_k.$$

Further, we have

$$\pi_i (Q_l)_{ij} = \pi_j (Q_l)_{ji},$$

so, by Theorem 2.3,

$$(P_l)_{ij} = (Q_l)_{ij}.$$

□

Remark 2.2. (a) It is easy to see that $P_t = Q_t$ for any Gibbs sampler in a generalized sense with transition matrix $P = P_1P_2...P_t$, P_1, P_2, ..., P_t corresponding to Q_1, Q_2, ..., Q_t, respectively. (This result follows from the condition (C3) in [16], $\Delta_{t+1} = (\{i\})_{i\in S}$, and definition of Gibbs sampler in a generalized sense.)

(b) It is easy to see that the conditions (c5) and (c6) hold for any Gibbs sampler in a generalized sense with transition matrix $P = P_1$, P_1 corresponding to Q_1 ($t = 1$ in this case; (c5) and (c6) refer to Q_1, not to P_1; $Q_1 > 0$ (by the condition (C3) in [16]), $Q_1 = e'\pi$ (by $Q_1 > 0$ and the definition of Gibbs sampler in a generalized sense), $P_1 = Q_1$ (by (a)); $Q_1 > 0$ implies (c5) and (c6)).

(c) If the conditions (c5) and (c6) do not hold for a Gibbs sampler in a generalized sense with transition matrix $P = P_1P_2...P_t$, P_1, P_2, ..., P_t corresponding to Q_1, Q_2, ..., Q_t, respectively ((c5) and (c6) refer to Q_1, Q_2, ..., Q_t, not to P_1, P_2, ..., P_t), we can have either $P_l \neq Q_l$ for some $l \in \langle t-1\rangle$ ($t > 1$, see (b); $l \in \langle t-1\rangle$ because, see (a), $P_t = Q_t$) or $P_l = Q_l$, $\forall l \in \langle t\rangle$. Indeed, this follows from the next two examples.

Example 2.1. Let $S = \langle 4\rangle$. Let

$$\pi = \left(\frac{\theta}{Z}, \frac{\theta^2}{Z}, \frac{\theta^2}{Z}, \frac{\theta^3}{Z}\right),$$

a probability distribution on S, where $\theta \in \mathbb{R}^+$ ($0 < \theta \leq 1$ or $\theta > 1$) and $Z = \theta + 2\theta^2 + \theta^3$ (the normalization constant). Let

$$\Delta_1 = (S)\,,\ \Delta_2 = (\{1,2\}, \{3,4\})\,,$$
$$\Delta_3 = (\{1\}, \{2\}, \{3\}, \{4\}) = (\{i\})_{i\in S}\,.$$

Note that π is a wavy probability distribution (with respect to Δ_1, Δ_2, Δ_3), $\forall\theta \in \mathbb{R}^+$. Further, we consider the case $\theta > 1$ only. For S, π, and Δ_1, Δ_2, Δ_3, we consider the Gibbs sampler in a generalized sense with transition matrix $P = P_1P_2$, P_1, P_2 corresponding to Q_1, Q_2, respectively, where

$$Q_1 = \begin{pmatrix} 0 & \frac{1}{1+\theta} & 0 & \frac{\theta}{1+\theta} \\ \frac{1}{1+\theta} & 0 & \frac{\theta}{1+\theta} & 0 \\ 0 & \frac{1}{1+\theta} & 0 & \frac{\theta}{1+\theta} \\ \frac{1}{1+\theta} & 0 & \frac{\theta}{1+\theta} & 0 \end{pmatrix},\ Q_2 = \begin{pmatrix} \frac{1}{1+\theta} & \frac{\theta}{1+\theta} & & \\ \frac{1}{1+\theta} & \frac{\theta}{1+\theta} & & \\ & & \frac{1}{1+\theta} & \frac{\theta}{1+\theta} \\ & & \frac{1}{1+\theta} & \frac{\theta}{1+\theta} \end{pmatrix}$$

($(Q_1)_{12} = \frac{\pi_2}{\pi_2+\pi_4} = \frac{1}{1+\theta}$, $(Q_1)_{14} = \frac{\pi_4}{\pi_2+\pi_4} = \frac{\theta}{1+\theta}$ (we considered $(Q_1)_{12}$, $(Q_1)_{14} > 0$), etc.; $\overline{Q}_1$ (the incidence matrix of Q_1) and $\overline{Q}_2$ are symmetric, Q_1 and Q_2 are not). Since $(Q_1)_{11} = 0$, the condition (c5) does not hold (for the sequence Q_1, Q_2). Since $2 \neq 4$, $(Q_1)_{12} > 0$, and $(Q_1)_{14} > 0$, but $(Q_1)_{24} = 0$, the condition (c6) does not hold. Moreover, we have

$$\pi_1 (Q_1)_{12} = \frac{\theta}{Z(1+\theta)} \neq \pi_2 (Q_1)_{21} = \frac{\theta^2}{Z(1+\theta)},\ \text{etc.}$$

By Theorem 2.3 (or by the definition of matrices P_l, $l \in \langle t\rangle$, from [16]) we have

$$P_1 = \begin{pmatrix} 0 & \frac{1}{1+\theta} & 0 & \frac{\theta}{1+\theta} \\ \frac{1}{\theta(1+\theta)} & \frac{\theta-1}{\theta} & \frac{1}{1+\theta} & 0 \\ 0 & \frac{1}{1+\theta} & 0 & \frac{\theta}{1+\theta} \\ \frac{1}{\theta(1+\theta)} & 0 & \frac{1}{1+\theta} & \frac{\theta-1}{\theta} \end{pmatrix} \text{ and (see Remark 2.2(a)) } P_2 = Q_2.$$

Consequently, $P_1 \neq Q_1$ — now, see Remark 2.2(c) again. Moreover, the second and fourth row of P_1 have 3 entries different from 0 while the second and fourth row of Q_1 have 2 entries different from 0 — now, it is interesting to go to Remark 2.1.

Example 2.2. Let $S = \langle 4 \rangle$. Let

$$\pi = \left(\frac{1}{10}, \frac{3}{10}, \frac{4}{10}, \frac{2}{10} \right),$$

a probability distribution on S. Let

$$\Delta_1 = (S)\,,\ \Delta_2 = (\{1,2\}\,,\{3,4\})\,,\ \Delta_3 = (\{i\})_{i \in S}\,.$$

Note that $\pi_1 + \pi_3 = \pi_2 + \pi_4$. For S, π, and Δ_1, Δ_2, Δ_3, we consider the Gibbs sampler in a generalized sense with transition matrix $P = P_1 P_2$, P_1, P_2 corresponding to Q_1, Q_2, respectively, where

$$Q_1 = \begin{pmatrix} 0 & \frac{3}{5} & 0 & \frac{2}{5} \\ \frac{1}{5} & 0 & \frac{4}{5} & 0 \\ 0 & \frac{3}{5} & 0 & \frac{2}{5} \\ \frac{1}{5} & 0 & \frac{4}{5} & 0 \end{pmatrix},\ Q_2 = \begin{pmatrix} \frac{1}{4} & \frac{3}{4} & & \\ \frac{1}{4} & \frac{3}{4} & & \\ & & \frac{4}{6} & \frac{2}{6} \\ & & \frac{4}{6} & \frac{2}{6} \end{pmatrix}$$

($\overline{Q}_1$ and $\overline{Q}_2$ are symmetric, Q_1 and Q_2 are not). The conditions (c5) and (c6) do not hold. Since $\pi_i \left(Q_l\right)_{ij} = \pi_j \left(Q_l\right)_{ji}$, $\forall l \in \langle 2 \rangle$, $\forall i,\ j \in S$ ($\pi_1 \left(Q_1\right)_{12} = \frac{3}{50} = \pi_2 \left(Q_1\right)_{21}$, etc.), by Theorem 2.3 we have $P_l = Q_l$, $\forall l \in \langle 2 \rangle$ — now, see Remark 2.2(c) again.

Remark 2.3. For the Gibbs samplers in a generalized sense, by the proof of Theorem 2.4, the conditions (c5) and (c6) imply

$$\pi_i \left(Q_l\right)_{ij} = \pi_j \left(Q_l\right)_{ji}\,,\ \forall l \in \langle t \rangle\,,\ \forall i, j \in S$$

— conversely, it is not true, see Example 2.2.

We conclude this section saying that to design good Gibbs samplers in a generalized sense we can use the conditions (c5) and (c6) or, more generally, the equations from Remark 2.3 — a case, a happy case, in which these equations hold is in Theorem 3.1 from [16]. (For more information on Theorem 3.1 from [16], see the first paragraph after Remark 2.1.)

3. Probability distribution of a random vector with independent components is a wavy probability distribution

In this section, we show that the probability distribution of a random vector with independent components is a wavy probability distribution with respect to the lexicographic order and certain partitions which will be specified. We work with random vectors in the finite case only — the number of components is finite and each component has a finite number of values, at least two values.

Let $X = (X_1, X_2, ..., X_n)$ be a random vector with the set of values $\langle\langle h_1 \rangle\rangle \times \langle\langle h_2 \rangle\rangle \times ... \times \langle\langle h_n \rangle\rangle$, where n, h_1, h_2, ..., $h_n \geq 1$. For simplification, suppose that $h_1 = h_2 = ... = h_n := h$, so, the set of values is $\langle\langle h \rangle\rangle^n$. Suppose that X_1, X_2, ..., X_n are independent random variables. Let π be the probability distribution of X (π is positive; $\pi = (\pi_x)_{x \in \langle\langle h \rangle\rangle^n}$). Let

$$U_{(x_1, x_2, ..., x_l)} = \{(y_1, y_2, ..., y_n) \mid\ (y_1, y_2, ..., y_n) \in \langle\langle h \rangle\rangle^n \text{ and } y_s = x_s, \forall s \in \langle l \rangle\}\,,$$

$\forall l \in \langle n \rangle$, $\forall x_1, x_2, ..., x_l \in \langle\langle h \rangle\rangle$. (The set

$$\left\{ (y_1, y_2, ..., y_n) \mid (y_1, y_2, ..., y_n) \in \langle\langle h \rangle\rangle^n \text{ and } y_s = x_s, \forall s \in \langle l \rangle \right\}$$

also appears in [16, Section 2], but in a different context and differently denoted, $K_{(x_1,x_2,...,x_l)}$ instead of $U_{(x_1,x_2,...,x_l)}$.) Let

$$\Delta_1 = \left(\langle\langle h \rangle\rangle^n \right),$$

$$\Delta_{l+1} = \left(U_{(x_1,x_2,...,x_l)} \right)_{x_1,x_2,...,x_l \in \langle\langle h \rangle\rangle}, \forall l \in \langle n \rangle$$

(obviously,

$$\Delta_{n+1} = \left(U_{(x_1,x_2,...,x_n)} \right)_{x_1,x_2,...,x_n \in \langle\langle h \rangle\rangle} = \left(\{ (x_1, x_2, ..., x_n) \} \right)_{(x_1,x_2,...,x_n) \in \langle\langle h \rangle\rangle^n}$$

and

$$\Delta_1 \succ \Delta_2 \succ ... \succ \Delta_{n+1}).$$

Suppose that $\langle\langle h \rangle\rangle^n$ is equipped with the lexicographic order. The sets of each partition Δ_{l+1}, where $l \in \langle n \rangle$, can be considered, if this is of interest to the reader, in the order induced by the lexicographic order on $\langle\langle h \rangle\rangle^l$: $\forall (a_1, a_2, ..., a_l), (b_1, b_2, ..., b_l) \in \langle\langle h \rangle\rangle^l$,

$$U_{(a_1,a_2,...,a_l)} \overset{ilex}{\leq} U_{(b_1,b_2,...,b_l)} \text{ if } (a_1, a_2, ..., a_l) \overset{lex}{\leq} (b_1, b_2, ..., b_l),$$

where $\overset{lex}{\leq}$ is the lexicographic order on $\langle\langle h \rangle\rangle^l$ and $\overset{ilex}{\leq}$ is the induced order by this on Δ_{l+1}.

Theorem 3.1. *Under the above conditions π is a wavy probability distribution on $\langle\langle h \rangle\rangle^n$ (with respect to the lexicographic order and partitions Δ_1, Δ_2, ..., Δ_{n+1}).*

Proof. Let $p_u^{(v)} = P(X_v = u)$, $\forall u \in \langle\langle h \rangle\rangle$, $\forall v \in \langle n \rangle$. Let $l \in \langle n \rangle$. Let $x_1, x_2, ..., x_l \in \langle\langle h \rangle\rangle$. Let $(y_1, y_2, ..., y_n) \in U_{(x_1,x_2,...,x_l)}$. It follows that $y_s = x_s$, $\forall s \in \langle l \rangle$. We have

$$\pi_{(y_1,y_2,...,y_n)} = \pi_{(x_1,x_2,...,x_l,y_{l+1},y_{l+2},...,y_n)} =$$
$$= P(X_1 = x_1) P(X_2 = x_2) ...$$
$$...P(X_l = x_l) P(X_{l+1} = y_{l+1}) P(X_{l+2} = y_{l+2}) ... P(X_n = y_n) =$$
$$= p_{x_1}^{(1)} p_{x_2}^{(2)} ... p_{x_l}^{(l)} p_{y_{l+1}}^{(l+1)} p_{y_{l+2}}^{(l+2)} ... p_{y_n}^{(n)}$$

($y_{l+1}, y_{l+2}, ..., y_n$, etc. vanish when $l = n$).

Set

$$K = \begin{cases} S & \text{if } l = 1, \\ U_{(x_1,x_2,...,x_{l-1})} & \text{if } l \in \langle n \rangle - \{1\}. \end{cases}$$

The sets from Δ_{l+1} which are included in K ($K \in \Delta_l$) are

$$U_{(x_1,x_2,...,x_{l-1},0)}, U_{(x_1,x_2,...,x_{l-1},1)}, ..., U_{(x_1,x_2,...,x_{l-1},h)}$$

($x_1, x_2, ..., x_{l-1}$ vanish when $l = 1$). $U_{(x_1,x_2,...,x_{l-1},0)}$ contains the first $\left| U_{(x_1,x_2,...,x_{l-1},0)} \right|$ elements of K (see the definition of wavy probability distribution in [16]). Suppose, further, that $x_l \neq 0$. Consequently, $(x_1, x_2, ..., x_l, y_{l+1}, y_{l+2}, ..., y_n) \notin U_{(x_1,x_2,...,x_{l-1},0)}$ ($(x_1, x_2, ..., x_l, y_{l+1}, y_{l+2}, ..., y_n) \in U_{(x_1,x_2,...,x_{l-1},1)} \subseteq K$, or $\in U_{(x_1,x_2,...,x_{l-1},2)} \subseteq K$, ..., or $\in U_{(x_1,x_2,...,x_{l-1},h)} \subseteq K$). We have

$$\pi_{(x_1,x_2,...,x_l,y_{l+1},y_{l+2},...,y_n)} = p_{x_1}^{(1)} p_{x_2}^{(2)} ... p_{x_l}^{(l)} p_{y_{l+1}}^{(l+1)} p_{y_{l+2}}^{(l+2)} ... p_{y_n}^{(n)}$$

$$= \frac{p^{(l)}_{x_l}}{p^{(l)}_0} p^{(1)}_{x_1} p^{(2)}_{x_2} ... p^{(l-1)}_{x_{l-1}} p^{(l)}_0 p^{(l+1)}_{y_{l+1}} p^{(l+2)}_{y_{l+2}} ... p^{(n)}_{y_n} = \frac{p^{(l)}_{x_l}}{p^{(l)}_0} \pi_{(x_1, x_2, ..., x_{l-1}, 0, y_{l+1}, y_{l+2}, ..., y_n)}$$

($\frac{p^{(l)}_{x_l}}{p^{(l)}_0}$ is the proportionality factor). Therefore, π is a wavy probability distribution on $\langle\langle h \rangle\rangle^n$. □

Remark 3.1. The Gibbs sampler with the state space $S = \langle\langle h \rangle\rangle^n$ (h, $n \geq 1$) and transition matrix $P = P_1 P_2 ... P_n$ from [12, Theorem 3.2] attains its stationarity at time 1 (1 step due to P or n steps due to P_1, P_2, ..., P_n). Using Theorem 3.1 from [16], we can construct the Gibbs sampler in a generalized sense for the wavy probability distribution π of random vector X. (For more information on Theorem 3.1 from [16], see the first paragraph after Remark 2.1 and last paragraph from Section 2.) This chain attains its stationarity at time 1 and, moreover, is even a Gibbs sampler. So, we have another case in which the Gibbs sampler attains its stationarity at time 1.

Remark 3.2. (a) If the random vector $X = (X_1, X_2, ..., X_n)$ has the set of values $\langle\langle 1 \rangle\rangle^n$ and independent and identically distributed components, then

$$X_1 + X_2 + ... + X_n \sim Bi\,(n, p)\,,$$

where $p = P\,(X_1 = 1)\,, 0 < p < 1$. It is interesting to connect the binomial distribution to the wavy probability distributions, *i.e.*, to obtain the results on the binomial distribution using the fact that π (the probability distribution of X) is a wavy probability distribution.

(b) The geometric distribution is a probability distribution related to the wavy probability distributions (see [12, Application 3.5]).

In the finite case, in the collection of probability distributions of random vectors with dependent components, we found three important wavy probability distributions: the Mallows model through Cayley metric and that through Kendall metric (see [13]−[14]; see also [16]) and, when $\theta \neq 1$, the Potts model on the tree (see [15]; see also [16]; see also Remark 5.1; for θ, see [15] or Section 5). Our interest is to find other, as many as possible, important wavy probability distributions. It is also of interest to us to find, besides the geometric distribution (see Remark 3.2(b)), other important probability distributions having connections with the wavy probability distributions.

4. Wavy probability distributions in a generalized sense

In this section, we generalize the notion of wavy probability distribution and Theorem 3.2 from [16].

Definition 4.1. Let $S = \langle r \rangle$. Let π be a positive probability distribution on S. Let $t \geq 1$. Let Δ_1, Δ_2, ..., Δ_{t+1} $\in$Par(S) with $\Delta_1 = (S) \succ \Delta_2 \succ ... \succ \Delta_{t+1} = (\{i\})_{i \in S}$. ($\Delta_1 \succ \Delta_2$ implies $r \geq 2$.) We say that π is a *wavy probability distribution in a generalized sense* (*with respect to* Δ_1, Δ_2, ..., Δ_{t+1}) if there exists a hybrid Metropolis-Hastings chain (in particular, a Gibbs sampler in a generalized sense) with state space S and transition matrix $P = P_1 P_2 ... P_t$, P_1, P_2, ..., P_t corresponding to Q_1, Q_2, ..., Q_t, respectively, the latter matrices being defined by means of Δ_1, Δ_2, ..., Δ_{t+1} (see [16]), such that

$$P = e'\pi.$$

The notion of wavy probability distribution in a generalized sense makes sense if the wavy probability distributions are wavy probability distributions in a generalized sense and if there exists at least one wavy probability distribution in a generalized sense which is not a wavy probability distribution. By Theorem 3.1 from [16], the wavy probability distributions are wavy probability distributions in a generalized sense. (For more information on Theorem 3.1 from [16], see the first paragraph after Remark 2.1 and last paragraph from Section 2.) An example of wavy probability distribution in a generalized sense (with respect to 3 partitions) which is not a wavy probability distribution (with respect to the same 3 partitions) is presented in Example 5.1 (from Section 5).

Although an order relation on S is not required for a wavy probability distribution in a generalized sense on S, such a relation will be considered when we will need it.

It is interesting to find the structure, if any, of wavy probability distributions in a generalized sense, at least in the case of Gibbs samplers in a generalized sense.

One way to obtain wavy probability distributions in a generalized sense is presented in the next result.

Theorem 4.1. *Let $S = \langle r \rangle$. Let π be a positive probability distribution on S. Let $t \geq 1$. Let Δ_1, Δ_2, ..., $\Delta_{t+1} \in \mathrm{Par}(S)$ with $\Delta_1 = (S) \succ \Delta_2 \succ ... \succ \Delta_{t+1} = (\{i\})_{i \in S}$. If there exists a hybrid Metropolis-Hastings chain (in particular, a Gibbs sampler in a generalized sense) with state space S and transition matrix $P = P_1 P_2 ... P_t$, P_1, P_2, ..., P_t corresponding to Q_1, Q_2, ..., Q_t, respectively, the latter matrices being defined by means of Δ_1, Δ_2, ..., Δ_{t+1}, such that*

$$P_1 \in G_{\Delta_1, \Delta_2}, \ P_2 \in G_{\Delta_2, \Delta_3}, \ ..., \ P_t \in G_{\Delta_t, \Delta_{t+1}},$$

then π is a wavy probability distribution in a generalized sense with respect to Δ_1, Δ_2, ..., Δ_{t+1}.

Proof. By Theorem 1.2 from [16] we have

$$\pi P = \pi.$$

By Theorem 1.1 from [16], P is a stable matrix. So, $\exists \psi$, ψ is a probability distribution on S, such that

$$P = e' \psi.$$

Finally, we have

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

so,

$$P = e' \pi.$$

□

Let $\Delta_1 \in \mathrm{Par}(\langle m \rangle)$ and $\Delta_2 \in \mathrm{Par}(\langle n \rangle)$. Let $P \in G_{\Delta_1, \Delta_2} \subseteq S_{m,n}$ (see, *e.g.*, [16] for G_{Δ_1, Δ_2}). Let $K \in \Delta_1$ and $L \in \Delta_2$. Then $\exists a_{K,L} \geq 0, \exists Q_{K,L} \in S_{|K|,|L|}$ such that $P_K^L = a_{K,L} Q_{K,L}$. Set the matrix

$$P^{-+} = \left(P_{KL}^{-+}\right)_{K \in \Delta_1, L \in \Delta_2}, \ P_{KL}^{-+} = a_{K,L}, \ \forall K \in \Delta_1, \ \forall L \in \Delta_2$$

(P_{KL}^{-+}, $K \in \Delta_1$, $L \in \Delta_2$, are the entries of matrix P^{-+}). If confusion can arise we write $P^{-+(\Delta_1, \Delta_2)}$ instead of P^{-+}. (For an example, see, *e.g.*, [12].)

Below we give a generalization of Theorem 3.2 from [16].

Theorem 4.2. *Let* $S = \langle r \rangle$. *Let* $\pi = (\pi_i)_{i \in S}$ *be a wavy probability distribution in a generalized sense* (*on* S) *with respect to the partitions* $\Delta_1, \Delta_2, ..., \Delta_{t+1}$. *Suppose that*

$$P_1 \in G_{\Delta_1,\Delta_2},\ P_2 \in G_{\Delta_2,\Delta_3},\ ...,\ P_t \in G_{\Delta_t,\Delta_{t+1}}$$

(*we use the notation from Definition* 4.1, ...). *Suppose that*

$$\pi_i = \frac{\nu_i}{Z},\ \forall i \in S,$$

where

$$Z = \sum_{i \in S} \nu_i,$$

Z *is the normalization constant* ($\nu_i \in \mathbb{R}^+$, $\forall i \in S$, *so,* $Z \in \mathbb{R}^+$). *Then* $\exists! U_2 \in \Delta_2$, $U_3 \in \Delta_3$, ..., $U_t \in \Delta_t$ *such that*

$$S \supseteq U_2 \supseteq U_3 \supseteq ... \supseteq U_t \supseteq \{1\};$$

further, we have

$$Z = \frac{\nu_1}{(P_1)^{-+}_{SU_2} (P_2)^{-+}_{U_2U_3} ... (P_t)^{-+}_{U_t\{1\}}}$$

$((P_1)^{-+} = (P_1)^{-+(\Delta_1,\Delta_2)}, (P_2)^{-+} = (P_2)^{-+(\Delta_2,\Delta_3)}, ..., (P_t)^{-+} = (P_t)^{-+(\Delta_t,\Delta_{t+1})})$.

Proof. The first part of conclusion is obvious because $\Delta_2, \Delta_3, ..., \Delta_t$ are partitions and $\Delta_1 = (S) \succ \Delta_2 \succ ... \succ \Delta_{t+1} = (\{i\})_{i \in S}$.

By Definition 4.1 we have $P = e'\pi$. It follows that

$$P^{-+(\Delta_1,\Delta_{t+1})} = (e'\pi)^{-+(\Delta_1,\Delta_{t+1})}$$

and, consequently,

$$\left(P^{-+(\Delta_1,\Delta_{t+1})}\right)_{S\{1\}} = \pi_1.$$

By Theorem 2.3(ii) from [10] (or Theorem 1.5(ii) from [11]) we have

$$P^{-+(\Delta_1,\Delta_{t+1})} = (P_1)^{-+(\Delta_1,\Delta_2)} (P_2)^{-+(\Delta_2,\Delta_3)} ... (P_t)^{-+(\Delta_t,\Delta_{t+1})} =$$

$$= (P_1)^{-+} (P_2)^{-+} ... (P_t)^{-+}.$$

We now have

$$\pi_1 = \left(P^{-+(\Delta_1,\Delta_{t+1})}\right)_{S\{1\}} = \left((P_1)^{-+} (P_2)^{-+} ... (P_t)^{-+}\right)_{S\{1\}} =$$

$$= (P_1)^{-+}_{SU_2} (P_2)^{-+}_{U_2U_3} ... (P_t)^{-+}_{U_t\{1\}}.$$

On the other hand,

$$\pi_1 = \frac{\nu_1}{Z}.$$

So,

$$Z = \frac{\nu_1}{(P_1)^{-+}_{SU_2} (P_2)^{-+}_{U_2U_3} ... (P_t)^{-+}_{U_t\{1\}}}.$$

□

5. Potts model

In this section, we show that the Potts model on a (finite) graph is a wavy probability distribution with respect to an order relation and three partitions which will be specified. For the Ising model on $\mathcal{C}_n$, the cycle graph with n vertices ($n \geq 3$), we compute the normalization constant and give an example for $n = 3$.

Let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ be a (nondirected) graph with vertex set $\mathcal{V} = \{V_1, V_2, ..., V_n\}$ and edge set $\mathcal{E}$. Suppose that $n \geq 2$ and $|\mathcal{E}| \geq 1$ ($|\cdot|$ is the cardinal). $[V_i, V_j]$ is the edge whose ends are vertices V_i and V_j, where $i,\ j \in \langle n \rangle$ ($i \neq j$). Consider the set of functions

$$\langle\langle h \rangle\rangle^{\mathcal{V}} = \{ f \mid f : \mathcal{V} \to \langle\langle h \rangle\rangle \},$$

where $h \geq 1$ ($h \in \mathbb{N}$). Represent the functions from $\langle\langle h \rangle\rangle^{\mathcal{V}}$ by vectors: if $f \in \langle\langle h \rangle\rangle^{\mathcal{V}}$, $V_i \longmapsto f(V_i) := x_i$, $\forall i \in \langle n \rangle$, then its vectorial representation is $(x_1, x_2, ..., x_n)$. $(x_1, x_2, ..., x_n)$, $x_1, x_2, ..., x_n \in \langle\langle h \rangle\rangle$, are called configurations. $\langle\langle h \rangle\rangle$ can be seen as a set of colors; in this case, if $(x_1, x_2, ..., x_n)$ is a configuration, then x_1 is the color of V_1, x_2 is the color of V_2, ..., x_n is the color of V_n.

Set (see, *e.g.*, [6, Chapter 6])

$$H(x) = \sum_{[V_i, V_j] \in \mathcal{E}} \mathbf{1}[x_i \neq x_j],\ \forall x \in \langle\langle h \rangle\rangle^n\ \ (x = (x_1, x_2, ..., x_n)),$$

where

$$\mathbf{1}[x_i \neq x_j] = \begin{cases} 1 & \text{if } x_i \neq x_j, \\ 0 & \text{if } x_i = x_j, \end{cases}$$

$\forall x \in \langle\langle h \rangle\rangle^n$, $\forall i,\ j \in \langle n \rangle$. The function H is called the *Hamiltonian* or *energy.* $H(x)$ represents the energy of configuration x.

Set

$$\pi_x = \frac{\theta^{H(x)}}{Z},\ \forall x \in \langle\langle h \rangle\rangle^n,$$

where $\theta \in \mathbb{R}^+$ and

$$Z = \sum_{x \in \langle\langle h \rangle\rangle^n} \theta^{H(x)}.$$

The probability distribution $\pi = (\pi_x)_{x \in \langle\langle h \rangle\rangle^n}$ (on $\langle\langle h \rangle\rangle^n$) is called, when $0 < \theta < 1$, the *Potts model on the graph* $\mathcal{G}$ (see [17]; see, *e.g.*, also [6, Chapter 6], [7], and [19]) — we extend this notion considering $\theta \in \mathbb{R}^+$. In particular, if $h = 1$ and $0 < \theta < 1$, π is called the *Ising model on the graph* $\mathcal{G}$ (see [5]; see, *e.g.*, also [6, Chapter 6] and [8]; no external field is allowed in our article) — we also extend this notion considering $\theta \in \mathbb{R}^+$. Z is called the *normalization constant.*

In this section, $\oplus$ is the addition modulo $h + 1$.

Consider the subsets $U_{(k)}$, $k \in \langle\langle h \rangle\rangle$, of $\langle\langle h \rangle\rangle^n$,

$$U_{(k)} = \{(y_1, y_2, ..., y_n) \mid (y_1, y_2, ..., y_n) \in \langle\langle h \rangle\rangle^n \text{ and } y_1 = k\},\ \forall k \in \langle\langle h \rangle\rangle.$$

Theorem 5.1. *We have*

$$U_{(k)} = U_{(0)} \oplus (k, k, ..., k),\ \forall k \in \langle\langle h \rangle\rangle,$$

where

$$U_{(0)} \oplus (k, k, ..., k) = \left\{(x_1, x_2, ..., x_n) \oplus (k, k, ..., k) \mid (x_1, x_2, ..., x_n) \in U_{(0)}\right\} =$$

$$= \left\{ (x_1 \oplus k, x_2 \oplus k, ..., x_n \oplus k) \,\middle|\, (x_1, x_2, ..., x_n) \in U_{(0)} \right\}, \ \forall k \in \langle\langle h \rangle\rangle .$$

Proof. Let $k \in \langle\langle h \rangle\rangle$. Consider the function $f_{0,k} : U_{(0)} \longrightarrow U_{(k)}$,

$$f_{0,k}(x_1, x_2, ..., x_n) = (x_1 \oplus k, x_2 \oplus k, ..., x_n \oplus k)$$

$((x_1 \oplus k, x_2 \oplus k, ..., x_n \oplus k) \in U_{(k)}$ because $x_1 \oplus k = 0 \oplus k = k$; $f_{0,k}$ is a special case of the function $f_{j,j\oplus k}$ $(j, k \in \langle\langle h \rangle\rangle)$ from [15]). This function is bijective because it is injective (it is easy to show this fact) and $\left|U_{(k)}\right| = \left|U_{(0)}\right|$. Consequently,

$$U_{(k)} = \left\{ (x_1 \oplus k, x_2 \oplus k, ..., x_n \oplus k) \,\middle|\, (x_1, x_2, ..., x_n) \in U_{(0)} \right\}.$$

□

Let $\stackrel{0}{\leq}$ be an order relation on $U_{(0)}$. The case when $\stackrel{0}{\leq} = \stackrel{lex}{\leq}$ is an interesting one, see Remark 5.1, where $\stackrel{lex}{\leq}$ is, as in Section 3, the lexicographic order (here, on $U_{(0)}$). Let $k \in \langle h \rangle$ $(\langle h \rangle = \langle\langle h \rangle\rangle - \{0\})$. Consider $U_{(k)}$ equipped with the order relation $\stackrel{k}{\leq}$ defined as follows (see the formula for $U_{(k)}$ from Theorem 5.1):

$$(x_1 \oplus k, x_2 \oplus k, ..., x_n \oplus k) \stackrel{k}{\leq} (z_1 \oplus k, z_2 \oplus k, ..., z_n \oplus k)$$

if

$$(x_1, x_2, ..., x_n) \stackrel{0}{\leq} (z_1, z_2, ..., z_n),$$

where $(x_1, x_2, ..., x_n), (z_1, z_2, ..., z_n) \in U_{(0)}$.

Consider $\langle\langle h \rangle\rangle^n$ equipped with the order relation $\preccurlyeq$ defined as follows ($\langle\langle h \rangle\rangle^n = \bigcup_{k \in \langle\langle h \rangle\rangle} U_{(k)}$):

$$(u_1, u_2, ..., u_n) \preccurlyeq (v_1, v_2, ..., v_n)$$

if

$$(u_1, u_2, ..., u_n) \in U_{(k_1)} \text{ and } (v_1, v_2, ..., v_n) \in U_{(k_2)} \text{ for some } k_1, k_2 \in \langle\langle h \rangle\rangle, \ k_1 < k_2,$$

or if

$$(u_1, u_2, ..., u_n), (v_1, v_2, ..., v_n) \in U_{(k)} \text{ and}$$
$$(u_1, u_2, ..., u_n) \stackrel{k}{\leq} (v_1, v_2, ..., v_n) \text{ for some } k \in \langle\langle h \rangle\rangle,$$

where $(u_1, u_2, ..., u_n), (v_1, v_2, ..., v_n) \in \langle\langle h \rangle\rangle^n$.

Theorem 5.2. *The Potts model on the graph $\mathcal{G}$ is a wavy probability distribution with respect to the order relation $\preccurlyeq$ and partitions*

$$\Delta_1 = \left(\langle\langle h \rangle\rangle^n\right), \ \Delta_2 = \left(U_{(0)}, U_{(1)}, ..., U_{(h)}\right), \ \Delta_3 = \left(\{x\}\right)_{x \in \langle\langle h \rangle\rangle^n}.$$

Proof. First, we consider Δ_1 and Δ_2 (see the definition of wavy probability distribution in [16]). $U_{(0)}$ contains the first $\left|U_{(0)}\right|$ elements of $\langle\langle h \rangle\rangle^n$ ($\langle\langle h \rangle\rangle^n \in \Delta_1$; $\langle\langle h \rangle\rangle^n$ is equipped with the order relation $\preccurlyeq$; $U_{(0)} \in \Delta_2$). Fix $U_{(k)}$ ($U_{(k)} \in \Delta_2$), where $k \in \langle h \rangle$ (not $k \in \langle\langle h \rangle\rangle$). (It follows that $U_{(k)} \neq U_{(0)}$.) Since $H(u_1, u_2, ..., u_n) = H(u_1 \oplus k, u_2 \oplus k, ..., u_n \oplus k)$, $\forall (u_1, u_2, ..., u_n) \in \langle\langle h \rangle\rangle^n$ (it is easy to show this), it follows that

$$\pi_{(x_1 \oplus k, x_2 \oplus k, ..., x_n \oplus k)} = \pi_{(x_1, x_2, ..., x_n)}$$

(the proportionality factor is 1), $\forall (x_1, x_2, ..., x_n) \in U_{(0)}$ $((x_1 \oplus k, x_2 \oplus k, ..., x_n \oplus k) \in U_{(k)}$ if $(x_1, x_2, ..., x_n) \in U_{(0)}$, see Theorem 5.1).

Second, we consider Δ_2 and Δ_3. Fix $\{x\}$ ($\{x\} \in \Delta_3$; $x = (x_1, x_2, ..., x_n)$). Then $\exists s \in \langle\langle h \rangle\rangle$ such that $\{x\} \subseteq U_{(s)}$ ($U_{(s)} \in \Delta_2$). Let y be the smallest element of $U_{(s)}$ ($y = (y_1, y_2, ..., y_n)$; $y_1 = s$). Suppose, moreover, that $x \neq y$. Since

$$\pi_y = \frac{\theta^{H(y)}}{Z},$$

we have

$$\pi_x = \frac{\theta^{H(x)}}{Z} = \frac{\theta^{H(x)}}{\theta^{H(y)}} \cdot \frac{\theta^{H(y)}}{Z} = \theta^{H(x)-H(y)} \pi_y$$

(the proportionality factor is $\theta^{H(x)-H(y)}$). □

Remark 5.1. Since $\pi_{(0,0,...,0)} = \frac{1}{Z}$, by Theorem 3.2 from [16] and Theorem 5.2 we obtain

$$Z = (1+h)\left(1 + \sum_{x \in U_{(0)},\, x \neq 0} \theta^{H(x)}\right) = (h+1) \sum_{x \in U_{(0)}} \theta^{H(x)}$$

for the Potts model on the graph $\mathcal{G}$ (to compute Z, the normalization constant, we also used the fact that $(0,0,...,0) \in U_{(0)}$ and $\theta^{H(0,0,...,0)} = \theta^0 = 1$). For the Potts model on the tree, we obtained more, namely,

$$Z_T = (h+1)(h\theta+1)^{n-1}$$

(see [15]), where Z_T is the normalization constant for this special model and n is the number of nodes (vertices) of tree ($n \geq 2$). This good formula is due to the fact that the Potts model on the tree with n nodes is a wavy probability distribution with respect to the $n+1$ partitions from [15] and order relation $\overset{f_c}{\leq}$, the induced order by (the function) f_c, defined as follows. By [15] it is easy to see that the Potts model on the star graph with n vertices ($n \geq 2$) is a wavy probability distribution with respect to the $n+1$ partitions from [15] and order relation $\leqslant$ when (do not forget this!) $\overset{0}{\leq} = \overset{lex}{\leq}$. Set

$$(x_1, x_2, ..., x_n) \overset{f_c}{\leq} (y_1, y_2, ..., y_n) \text{ if } f_c(x_1, x_2, ..., x_n) \leqslant f_c(y_1, y_2, ..., y_n),$$

where $(x_1, x_2, ..., x_n), (y_1, y_2, ..., y_n) \in \langle\langle h \rangle\rangle^n$, $(x_1, x_2, ..., x_n), (y_1, y_2, ..., y_n)$ are configurations of the tree with n nodes, $f_c : \langle\langle h \rangle\rangle^n \longrightarrow \langle\langle h \rangle\rangle^n$, the function for configurations from [15] (see this work for the complete definition of f_c, for the fact that f_c is bijective, etc.), and $f_c(x_1, x_2, ..., x_n), f_c(y_1, y_2, ..., y_n)$ are configurations of the star graph with n vertices. Note, moreover, that, due to Theorem 5.2, the Potts model on the tree is also a wavy probability distribution with respect to the 3 partitions from Theorem 5.2 and order relation $\leqslant$.

Below we give another application of our results about trees from [15] — we derive the normalization constant for the Ising model on $\mathcal{C}_n$, the cycle graph with n vertices, from the normalization constant for the Ising model on $\mathcal{P}_n$, the path graph with n vertices. For a different proof of the next result, see, *e.g.*, [8, pp. 31–35].

Theorem 5.3. *Let $Z_{\mathcal{C}_n}$ be the normalization constant for the Ising model on $\mathcal{C}_n$ (with the parameter θ; $h = 1$ in this case; $n \geq 3$). Then*

$$Z_{\mathcal{C}_n} = (1+\theta)^n + (1-\theta)^n.$$

Proof. Suppose that $\mathcal{C}_n$ has the vertices V_1, V_2, ..., V_n and edges $[V_1, V_2]$, $[V_2, V_3]$, ..., $[V_{n-1}, V_n]$, $[V_n, V_1]$. We cut the edge $[V_1, V_n]$ of $\mathcal{C}_n$ ($[V_1, V_n] = [V_n, V_1]$) and obtain $\mathcal{P}_n$, which is a special tree with n vertices. Denote the normalization constant for the Ising model on $\mathcal{P}_n$ by $Z_{\mathcal{P}_n}$. By [15] we have

$$Z_{\mathcal{P}_n} = 2\left(\theta + 1\right)^{n-1}.$$

Let $H_{\mathcal{C}_n}$ and $H_{\mathcal{P}_n}$ be the energies of Ising model on $\mathcal{C}_n$ and of that on $\mathcal{P}_n$, respectively. Any configuration of $\mathcal{P}_n$ is a (finite) sequence of $0'$s and $1'$s. If the ends of sequence are equal, 0, 0 or 1, 1, then, from left to right, the number of transitions from 0 to 1 or from 1 to 0 is even while if the ends of sequence are different, 0, 1 or 1, 0, then, from left to right too, the number of transitions from 0 to 1 or from 1 to 0 is odd. *E.g.*, for $n = 6$, the sequence 0, 1, 0, 0, 1, 0 has 4 transitions (from 0 to 1 or from 1 to 0, from left to right), so, $H_{\mathcal{P}_6}(0,1,0,0,1,0) = 4$ and, consequently, $H_{\mathcal{C}_6}(0,1,0,0,1,0) = 4$ while the sequence 0, 1, 1, 0, 1, 1 has 3 transitions, so, $H_{\mathcal{P}_6}(0,1,1,0,1,1) = 3$ and, consequently, $H_{\mathcal{C}_6}(0,1,1,0,1,1) = 3 + 1 = 4$. Due this fact, we have

$$\begin{aligned} Z_{\mathcal{C}_n} &= \sum_{x \in \langle\langle 1 \rangle\rangle^n} \theta^{H_{\mathcal{C}_n}(x)} \\ &= [\text{the sum of terms of } Z_{\mathcal{P}_n} \text{ having } \theta \text{ with even exponent}] \\ &\quad + \theta \times [\text{the sum of terms of } Z_{\mathcal{P}_n} \text{ having } \theta \text{ with odd exponent}] \\ &= 2\left[\left(1 + C_{n-1}^2\theta^2 + C_{n-1}^4\theta^4 + ...\right) + \theta\left(C_{n-1}^1\theta + C_{n-1}^3\theta^3 + ...\right)\right]. \end{aligned}$$

It is easy to see that (a known result)

$$C_{s-1}^{t-1} + C_{s-1}^t = C_s^t, \ \forall s, t, \ s \geq 2, \ 1 \leq t \leq s - 1.$$

Case 1. $n = 2k$, $k \geq 2$. We have

$$\begin{aligned} Z_{\mathcal{C}_{2k}} &= 2\left[\left(1 + C_{2k-1}^2\theta^2 + C_{2k-1}^4\theta^4 + ...\right) + \theta\left(C_{2k-1}^1\theta + C_{2k-1}^3\theta^3 + ...\right)\right] \\ &= 2C_{2k-1}^0 + 2\left(C_{2k-1}^1 + C_{2k-1}^2\right)\theta^2 + 2\left(C_{2k-1}^3 + C_{2k-1}^4\right)\theta^4 + ... + \\ &\quad + 2\left(C_{2k-1}^{2k-3} + C_{2k-1}^{2k-2}\right)\theta^{2k-2} + 2C_{2k-1}^{2k-1}\theta^{2k} \\ &= 2C_{2k}^0 + 2C_{2k}^2\theta^2 + 2C_{2k}^4\theta^4 + ... + 2C_{2k}^{2k-2}\theta^{2k-2} + 2C_{2k}^{2k}\theta^{2k} \\ &= \left(C_{2k}^0 + C_{2k}^1\theta + C_{2k}^2\theta^2 + ... + C_{2k}^{2k}\theta^{2k}\right) + \left(C_{2k}^0 - C_{2k}^1\theta + C_{2k}^2\theta^2 - ... + C_{2k}^{2k}\theta^{2k}\right) \\ &= (1 + \theta)^{2k} + (1 - \theta)^{2k}. \end{aligned}$$

Case 2. $n = 2k + 1$, $k \geq 1$. We have

$$\begin{aligned} Z_{\mathcal{C}_{2k+1}} &= 2\left[\left(1 + C_{2k}^2\theta^2 + C_{2k}^4\theta^4 + ...\right) + \theta\left(C_{2k}^1\theta + C_{2k}^3\theta^3 + ...\right)\right] \\ &= 2C_{2k}^0 + 2\left(C_{2k}^1 + C_{2k}^2\right)\theta^2 + 2\left(C_{2k}^3 + C_{2k}^4\right)\theta^4 + ... + 2\left(C_{2k}^{2k-1} + C_{2k}^{2k}\right)\theta^{2k} \\ &= 2C_{2k+1}^0 + 2C_{2k+1}^2\theta^2 + 2C_{2k+1}^4\theta^4 + ... + 2C_{2k+1}^{2k}\theta^{2k} \\ &= \left(C_{2k+1}^0 + C_{2k+1}^1\theta + C_{2k+1}^2\theta^2 + ... + C_{2k+1}^{2k+1}\theta^{2k+1}\right) \\ &\quad + \left(C_{2k+1}^0 - C_{2k+1}^1\theta + C_{2k+1}^2\theta^2 - ... - C_{2k+1}^{2k+1}\theta^{2k+1}\right) \\ &= (1 + \theta)^{2k+1} + (1 - \theta)^{2k+1}. \end{aligned}$$

From Cases 1 and 2, we have

$$Z_{\mathcal{C}_n} = (1 + \theta)^n + (1 - \theta)^n.$$

□

From Theorem 5.2, we know that the Ising model on $\mathcal{C}_3$ is a wavy probability distribution with respect to the order relation $\leqslant$ and 3 partitions from there in this special case. But the Ising model on $\mathcal{C}_3$ is also a wavy probability distribution in a generalized sense with respect to 3 partitions which will be specified in the next example (it is easy to give other examples for 3 or even more partitions).

Example 5.1. Let π be the Ising model on $\mathcal{C}_3$. Consider that the elements of $\langle\langle 1\rangle\rangle^3$ are in the order:

$$(0,0,0)\,,\ (1,1,1)\,,\ (0,0,1)\,,\ (1,1,0)\,,\ (0,1,0)\,,\ (1,0,1)\,,\ (0,1,1)\,,\ (1,0,0)\,.$$

By this order,

$$\pi=\left(\frac{1}{Z},\frac{1}{Z},\frac{\theta^2}{Z},\frac{\theta^2}{Z},\frac{\theta^2}{Z},\frac{\theta^2}{Z},\frac{\theta^2}{Z},\frac{\theta^2}{Z}\right).$$

Consider the partitions

$$\Delta_1=\left(\langle\langle 1\rangle\rangle^3\right),$$

$$\Delta_2=\left(\{(0,0,0)\,,(1,1,1)\}\,,\{(0,0,1)\,,(1,1,0)\}\,,\{(0,1,0)\,,(1,0,1)\,,(0,1,1)\,,(1,0,0)\}\right),$$

$$\Delta_3=\left(\{x\}\right)_{x\in\langle\langle 1\rangle\rangle^3}\,.$$

Consider the Gibbs sampler in a generalized sense with transition matrix $P=P_1P_2$, P_1, P_2 corresponding to Q_1, Q_2, respectively, where (we use Theorems 2.2 and 2.4 — these results, among other things, help us to construct examples of Gibbs samplers in a generalized sense more quickly)

$$P_1=Q_1=\begin{pmatrix}
\frac{1}{1+3\theta^2} & 0 & \frac{\theta^2}{1+3\theta^2} & 0 & \frac{\theta^2}{1+3\theta^2} & \frac{\theta^2}{1+3\theta^2} & 0 & 0\\
0 & \frac{1}{1+3\theta^2} & 0 & \frac{\theta^2}{1+3\theta^2} & 0 & 0 & \frac{\theta^2}{1+3\theta^2} & \frac{\theta^2}{1+3\theta^2}\\
\frac{1}{1+3\theta^2} & 0 & \frac{\theta^2}{1+3\theta^2} & 0 & \frac{\theta^2}{1+3\theta^2} & \frac{\theta^2}{1+3\theta^2} & 0 & 0\\
0 & \frac{1}{1+3\theta^2} & 0 & \frac{\theta^2}{1+3\theta^2} & 0 & 0 & \frac{\theta^2}{1+3\theta^2} & \frac{\theta^2}{1+3\theta^2}\\
\frac{1}{1+3\theta^2} & 0 & \frac{\theta^2}{1+3\theta^2} & 0 & \frac{\theta^2}{1+3\theta^2} & \frac{\theta^2}{1+3\theta^2} & 0 & 0\\
\frac{1}{1+3\theta^2} & 0 & \frac{\theta^2}{1+3\theta^2} & 0 & \frac{\theta^2}{1+3\theta^2} & \frac{\theta^2}{1+3\theta^2} & 0 & 0\\
0 & \frac{1}{1+3\theta^2} & 0 & \frac{\theta^2}{1+3\theta^2} & 0 & 0 & \frac{\theta^2}{1+3\theta^2} & \frac{\theta^2}{1+3\theta^2}\\
0 & \frac{1}{1+3\theta^2} & 0 & \frac{\theta^2}{1+3\theta^2} & 0 & 0 & \frac{\theta^2}{1+3\theta^2} & \frac{\theta^2}{1+3\theta^2}
\end{pmatrix},$$

$$P_2=Q_2=\begin{pmatrix}
\frac{1}{2} & \frac{1}{2} & & & & & & \\
\frac{1}{2} & \frac{1}{2} & & & & & & \\
& & \frac{1}{2} & \frac{1}{2} & & & & \\
& & \frac{1}{2} & \frac{1}{2} & & & & \\
& & & & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4}\\
& & & & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4}\\
& & & & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4}\\
& & & & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4}
\end{pmatrix}$$

(the rows and columns of P_1 and P_2 are labeled using the given order relation). We have

$$P_1\in G_{\Delta_1,\Delta_2},\ P_2\in G_{\Delta_2,\Delta_3}.$$

By Theorem 4.1, π is a wavy probability distribution in a generalized sense with respect to the partitions considered. The normalization constant Z can be obtained in this case by direct computation, Theorem 5.3, or Theorem 4.2. If we use Theorem 4.2, since

$$\pi_{(0,0,0)} = \frac{1}{Z},$$

we have

$$Z = \frac{1}{(P_1)^{-+}_{\langle\langle 1\rangle\rangle^3 \{(0,0,0),\ (1,1,1)\}} (P_2)^{-+}_{\{(0,0,0),\ (1,1,1)\}\{(0,0,0)\}}} = \frac{1}{\frac{1}{1+3\theta^2} \cdot \frac{1}{2}} = 2\left(1 + 3\theta^2\right)$$

(we can use the notations

$$(P_1)^{-+}_{\langle\langle 1\rangle\rangle^3 \to \{(0,0,0),\ (1,1,1)\}} \text{ and } (P_2)^{-+}_{\{(0,0,0),\ (1,1,1)\} \to \{(0,0,0)\}}$$

instead of

$$(P_1)^{-+}_{\langle\langle 1\rangle\rangle^3 \{(0,0,0),\ (1,1,1)\}} \text{ and } (P_2)^{-+}_{\{(0,0,0),\ (1,1,1)\}\{(0,0,0)\}},$$

respectively, see [16]).

Remark 5.2. For the Ising model on $\mathcal{C}_n$, from the normalization constant computed in Theorem 5.3, we can obtain other things, such as, the mean energy (see, *e.g.*, [1, p. 6] or [15] for the computation method for this) and free energy per site (see, *e.g.*, [7] or [15] for this notion).

References

[1] J. Beck, *Inevitable Randomness in Discrete Mathematics*, AMS, Providence, Rhode Island, 2009.

[2] S. Geman, D. Geman, Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images, *IEEE Pattern Anal. Mach. Intell.* **6** (1984), 721–741.

[3] W.K. Hastings, Monte Carlo sampling methods using Markov chains and their applications, *Biometrika* **57** (1970), 97–109.

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

[5] E. Ising, Beitrag zur Theorie des Ferromagnetismus, Z. *Physik* **31** (1925), 253–258.

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

[7] P. Martin, *Potts Models and Related Problems in Statistical Mechanics*, World Scientific, Singapore, 1991.

[8] B.M. McCoy, T.T. Wu, *The Two-Dimensional Ising Model*, 2nd Edition, Dover, Mineola, N.Y., 2014.

[9] N. Metropolis, A.W. Rosenbluth, M.N. Rosenbluth, A.H. Teller, E. Teller, Equations of state calculations by fast computing machines, *J. Chem. Phys.* **21** (1953), 1087–1092.

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

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

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

[13] 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.

[14] U. Păun, G method in action: fast exact sampling from set of permutations of order n according to Mallows model through Kendall metric, *Rev. Roumaine Math. Pures Appl.* **63** (2018), 259–280.

[15] U. Păun, G method in action: normalization constant, important probabilities, and fast exact sampling for Potts model on trees, *Submitted.*

[16] U. Păun, A Gibbs sampler in a generalized sense, *Ann. Univ. Craiova Math. Comp. Sci. Ser.* **43** (2016), 62–71.
([10]−[16] are in the order in which they were written, leaving the improvements aside.)
[17] R.B. Potts, Some generalized order-disorder transitions, *Proc. Cambridge Philos. Soc.* **48** (1952), 106–109.
[18] C. Robert, G. Casella, A short history of Markov chain Monte Carlo: subjective recollections from incomplete data, *Statist. Sci.* **26** (2011), 102–115.
[19] F.Y. Wu, The Potts model, *Rev. Mod. Phys.* **54** (1982), 235–268.

(Udrea Păun) Romanian Academy, "Gheorghe Mihoc - Caius Iacob" Institute of Mathematical Statistics and Applied Mathematics
Calea 13 Septembrie nr. 13, 050711 Bucharest, Romania
E-mail address: paun@csm.ro