G METHOD IN ACTION: NORMALIZATION CONSTANT, IMPORTANT PROBABILITIES, AND FAST EXACT SAMPLING FOR POTTS MODEL ON TREES

UDREA PĂUN

Communicated by Marius Iosifescu

Using G method, for the Potts model on (nondirected simple finite) trees, we compute the normalization constant and certain important probabilities and give a fast exact (not approximate) sampling Markovian method. The Ising model for n linear points (magnets) — no external field is allowed in our article — is a special Potts model, a model on the path graph (a tree) of above points. But this is not all — for the Potts model on connected (nondirected simple finite) graphs, we give bounds, nontrivial bounds, for the normalization constant, for the product of normalization constant and mean energy (this product is equal to the derivative of normalization constant), for the mean energy, for the mean energy per site, for the free energy per site, and for the limit free energy per site. These bounds help us to understand the Potts model better.

AMS 2010 Subject Classification: 60J10, 62Dxx, 68U20, 82B20.

Key words: G method, Gibbs sampler in a generalized sense, Potts model, Ising model, star graph, tree, connected graph, normalization constant, mean energy, mean energy per site, free energy per site, limit mean energy per site, limit free energy per site, bound, important probabilities, exact sampling.

1. SOME BASIC THINGS

In this section, we present some basic things from [8–10].

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 Z_{ij} or, if confusion can arise, $Z_{i\to j}$.

Set

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

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 ([8]). Let $P\in N_{m,n}$. Let $\Delta\in$Par$(\langle m\rangle)$ and $\Sigma\in$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 ([8]). Let $P\in N_{m,n}$. Let $\Delta\in$Par$(\langle m\rangle)$ and $\Sigma\in$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 $\Delta_1\in$Par$(\langle m\rangle)$ and $\Delta_2\in$Par$(\langle n\rangle)$. Set (see [8] for G_{Δ_1,Δ_2} and [9] 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.* G comes from the verb *to group* and its derivatives.

Below we give the basic result from [8] we need.

THEOREM 1.5 ([8]). *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*

$$P_1 P_2 ... P_n$$

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

Proof. See [8]. □

2. POTTS MODEL ON STAR GRAPHS

In this section, for the Potts model on star graphs, we compute the normalization constant and certain important probabilities and give a fast exact (not approximate) sampling Markovian method.

Set

$$\langle\langle m \rangle\rangle = \{0, 1, ..., m\} \ \ (m \geq 0).$$

Let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ be a (nondirected simple finite) graph with vertex set $\mathcal{V} = \{V_1, V_2, ..., V_n\}$ and edge set $\mathcal{E}$. Suppose that $|\mathcal{E}| \geq 1$ ($|\cdot|$ is the cardinal; $|\mathcal{E}| \geq 1 \Longrightarrow n \geq 2$). $[V_i, V_j]$ 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* (the *configurations of graph*). $\langle\langle h \rangle\rangle$ 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.*, [5, 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

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

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

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

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 [13]; see, *e.g.*, also [5, Chapter 6], [6], and [14]) — 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 [3]; see, *e.g.*, also [5, Chapter 6] and [7]; no external field is allowed in our article) — we also extend this notion considering $\theta \in \mathbb{R}^+$. Z is called the *normalization constant* (or the *normalizing constant,* or, extending the physical terminology, the *partition function*).

The next result is simple, but useful.

Theorem 2.1. *Let x, y, $z \in \langle\langle m-1 \rangle\rangle$, where $m \geq 1$ ($m \in \mathbb{N}$). Then*

$$x \neq y \iff x \oplus z \neq y \oplus z$$

(equivalently,

$$x = y \iff x \oplus z = y \oplus z),$$

where $\oplus$ is the addition modulo m. Moreover, the implication "$\Longleftarrow$" still holds if we replace x, y, $z \in \langle\langle m-1 \rangle\rangle$ with x, y, $z \in \mathbb{Z}$.

Proof. This is left to the reader. $\square$

Below we give a basic result about H, about the Potts model on graphs.

Theorem 2.2.

$$H(x_1, x_2, ..., x_n) = H(x_1 \oplus k, x_2 \oplus k, ..., x_n \oplus k), \ \forall x_1, x_2, ..., x_n, k \in \langle\langle h \rangle\rangle,$$

where $\oplus$ is the addition modulo $h+1$. (If $h = 1$, then $H(x_1, x_2, ..., x_n) = H(x_1 \oplus 1, x_2 \oplus 1, ..., x_n \oplus 1) = H(\overline{x}_1, \overline{x}_2, ..., \overline{x}_n)$, $\forall x_1, x_2, ..., x_n \in \langle\langle 1 \rangle\rangle$, where

$$\overline{x}_i = 1 - x_i = \begin{cases} 1 & \text{if } x_i = 0, \\ 0 & \text{if } x_i = 1, \end{cases}$$

$\forall i \in \langle n \rangle$.*)*

Proof. By Theorem 2.1. $\square$

Set

$$U_{(x_1)} = \{(y_1, y_2, ..., y_n) \mid y_1, y_2, ..., y_n \in \langle\langle h \rangle\rangle, \ y_1 = x_1\}, \ \forall x_1 \in \langle\langle h \rangle\rangle.$$

Below we give other basic results about the Potts model on graphs.

THEOREM 2.3. (i) $\left(U_{(0)},U_{(1)},...,U_{(h)}\right)$ *is a partition of* $\langle\langle h\rangle\rangle^{n}$.

(ii)

$$\left|U_{(v)}\right| = \left|U_{(w)}\right|,\ \forall v,w\in\langle\langle h\rangle\rangle.$$

(iii)

$$\pi_{(x_1,x_2,...,x_n)} = \pi_{(x_1\oplus k,x_2\oplus k,...,x_n\oplus k)},\ \forall x_1,x_2,...,x_n,k\in\langle\langle h\rangle\rangle,$$

where $\oplus$ *is the addition modulo* $h+1$.

(iv) *The function* $f_{j,j\oplus k}:U_{(j)}\longrightarrow U_{(j\oplus k)}$,

$$f_{j,j\oplus k}\left(j,x_2,...,x_n\right)=\left(j\oplus k,x_2\oplus k,...,x_n\oplus k\right),\forall x_2,x_3,...,x_n\in\langle\langle h\rangle\rangle,$$

is bijective, $\forall j,k\in\langle\langle h\rangle\rangle$, *and — do not forget this! —*

$$\pi_{f_{j,j\oplus k}(j,x_2,...,x_n)} = \pi_{(j,x_2,...,x_n)},\ \forall x_2,x_3,...,x_n,j,k\in\langle\langle h\rangle\rangle$$

($\oplus$ *is also the addition modulo* $h+1$).

(v)

$$P\left(U_{(v)}\right)=P\left(U_{(w)}\right),\ \forall v,w\in\langle\langle h\rangle\rangle,$$

and, therefore,

$$P\left(U_{(v)}\right)=\frac{1}{h+1},\ \forall v\in\langle\langle h\rangle\rangle,$$

where

$$P\left(U_{(v)}\right)=\sum_{x\in U_{(v)}}\pi_x,\ \forall v\in\langle\langle h\rangle\rangle.$$

(vi)

$$Z=(h+1)\sum_{x\in U_{(v)}}\theta^{H(x)},\ \forall v\in\langle\langle h\rangle\rangle,$$

and, as a result,

$$Z>h+1.$$

Proof. (i) and (ii) Obvious.

(iii) By Theorem 2.2.

(iv) $\left|U_{(j)}\right| = \left|U_{(j\oplus k)}\right|$ (by (ii)) and $f_{j,j\oplus k}$ is injective (by Theorem 2.1) imply $f_{j,j\oplus k}$ is bijective, $\forall j,k\in\langle\langle h\rangle\rangle$. The equations follow from (iii).

(v) By (iv).

(vi) Let $v\in\langle\langle h\rangle\rangle$. By (v) we have

$$\frac{1}{h+1}=P\left(U_{(v)}\right)=\sum_{x\in U_{(v)}}\pi_x=\sum_{x\in U_{(v)}}\frac{\theta^{H(x)}}{Z}=\frac{1}{Z}\sum_{x\in U_{(v)}}\theta^{H(x)}.$$

So,

$$Z=(h+1)\sum_{x\in U_{(v)}}\theta^{H(x)}.$$

Since $(v, v, ..., v) \in U_{(v)}$, $H(v, v, ..., v) = 0$, $|\langle\langle h\rangle\rangle| \geq 2$, and $|\mathcal{E}| \geq 1$, we have

$$\sum_{x \in U_{(v)}} \theta^{H(x)} > 1.$$

Consequently,

$$Z > h + 1. \quad \square$$

Let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ be a star graph, where $\mathcal{V} = \{V_1, V_2, ..., V_n\}$ $(n \geq 2)$, the vertex set, and $\mathcal{E} = \{[V_1, V_2], [V_1, V_3], ..., [V_1, V_n]\}$, the edge set.

THEOREM 2.4. *Consider the above star graph. Then*

$$H(x_1, x_2, ..., x_{t-1}, a, x_{t+1}, ..., x_n) =$$

$$= \begin{cases} H(x_1, x_2, ..., x_{t-1}, b, x_{t+1}, ..., x_n) & \text{if } a = b \text{ or } x_1 \neq a \neq b \neq x_1, \\ H(x_1, x_2, ..., x_{t-1}, b, x_{t+1}, ..., x_n) - 1 & \text{if } x_1 = a \neq b, \\ H(x_1, x_2, ..., x_{t-1}, b, x_{t+1}, ..., x_n) + 1 & \text{if } a \neq b = x_1, \end{cases}$$

$\forall t \in \langle n\rangle - \{1\}$ *(x_{t+1}, ..., x_n vanish when $t = n$),* $\forall x_1, x_2, ..., x_{t-1}, x_{t+1}, ..., x_n, a, b \in \langle\langle h\rangle\rangle$.

Proof. Obvious. $\square$

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

Below we give the main result of this section.

THEOREM 2.5. *Let* $\pi = (\pi_x)_{x \in \langle\langle h\rangle\rangle^n}$ *be the Potts model on the above star graph. Consider a Markov chain with state space* $\langle\langle h\rangle\rangle^n$ *and transition matrix* $P = P_1 P_2 ... P_n$, *where* P_t, $t \in \langle n\rangle$, *are stochastic matrices on* $\langle\langle h\rangle\rangle^n$,

$$(P_1)_{(x_1, x_2, ..., x_n) \to \xi} =$$

$$= \begin{cases} \dfrac{\pi_{(x_1 \oplus k, x_2 \oplus k, ..., x_n \oplus k)}}{\sum\limits_{u \in \langle\langle h\rangle\rangle} \pi_{(x_1 \oplus u, x_2 \oplus u, ..., x_n \oplus u)}} & \begin{array}{l} \text{if } \xi = (x_1 \oplus k, x_2 \oplus k, ..., x_n \oplus k) \\ \text{for some } k \in \langle\langle h\rangle\rangle, \end{array} \\ 0 & \text{if } \xi \neq (x_1 \oplus k, x_2 \oplus k, ..., x_n \oplus k), \forall k \in \langle\langle h\rangle\rangle, \end{cases}$$

$\forall (x_1, x_2, ..., x_n), \xi \in \langle\langle h\rangle\rangle^n$, *where* $\oplus$ *is the addition modulo* $h + 1$,

$$(P_t)_{(x_1, x_2, ..., x_n) \to \xi} =$$

$$= \begin{cases} \dfrac{\pi_{(x_1, x_2, ..., x_{t-1}, k, x_{t+1}, ..., x_n)}}{\sum\limits_{u \in \langle\langle h\rangle\rangle} \pi_{(x_1, x_2, ..., x_{t-1}, u, x_{t+1}, ..., x_n)}} & \begin{array}{l} \text{if } \xi = (x_1, x_2, ..., x_{t-1}, k, x_{t+1}, ..., x_n) \\ \text{for some } k \in \langle\langle h\rangle\rangle, \end{array} \\ 0 & \begin{array}{l} \text{if } \xi \neq (x_1, x_2, ..., x_{t-1}, k, x_{t+1}, ..., x_n), \\ \forall k \in \langle\langle h\rangle\rangle, \end{array} \end{cases}$$

$\forall t \in \langle n \rangle - \{1\}$ *(x_{t+1}, ..., x_n vanish when $t = n$),* $\forall \left(x_1, x_2, ..., x_n\right), \ \xi \in \langle\langle h \rangle\rangle^n$. *Then*

$$P = e'\pi$$

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

Proof. Set

$$U_{(x_1,x_2,...,x_t)} = \left\{ (y_1, y_2, ..., y_n) \mid \ (y_1, y_2, ..., y_n) \in \langle\langle h \rangle\rangle^n, \ y_v = x_v, \forall v \in \langle t \rangle \right\},$$

$\forall t \in \langle n \rangle$, $\forall x_1, x_2, ..., x_t \in \langle\langle h \rangle\rangle$. Consider the partitions

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

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

$\forall t \in \langle n \rangle$. Obviously, we have $\Delta_{n+1} = (\{x\})_{x \in \langle\langle h \rangle\rangle^n}$.

By hypothesis and Theorems 2.3(iii) and 2.4 we have

$$\left(P_1\right)_{(x_1,x_2,...,x_n) \to \xi} =$$

$$= \begin{cases} \frac{1}{h+1} & \text{if } \xi = (x_1 \oplus k, x_2 \oplus k, ..., x_n \oplus k) \text{ for some } k \in \langle\langle h \rangle\rangle, \\ 0 & \text{if } \xi \neq (x_1 \oplus k, x_2 \oplus k, ..., x_n \oplus k), \forall k \in \langle\langle h \rangle\rangle, \end{cases}$$

$\forall \left(x_1, x_2, ..., x_n\right), \xi \in \langle\langle h \rangle\rangle^n$,

$$\left(P_t\right)_{(x_1,x_2,...,x_n) \to \xi} =$$

$$= \begin{cases} \frac{1}{h\theta+1} & \text{if } \xi = (x_1, x_2, ..., x_{t-1}, x_1, x_{t+1}, ..., x_n), \\ \frac{\theta}{h\theta+1} & \text{if } \xi = (x_1, x_2, ..., x_{t-1}, k, x_{t+1}, ..., x_n) \text{ for some } k \in \langle\langle h \rangle\rangle, k \neq x_1, \\ 0 & \text{if } \xi \neq (x_1, x_2, ..., x_{t-1}, k, x_{t+1}, ..., x_n), \ \forall k \in \langle\langle h \rangle\rangle, \end{cases}$$

$\forall t \in \langle n \rangle - \{1\}$, $\forall \left(x_1, x_2, ..., x_n\right), \xi \in \langle\langle h \rangle\rangle^n$. It follows that

$$P_t \in G_{\Delta_t, \Delta_{t+1}}, \ \forall t \in \langle n \rangle.$$

Since $P = P_1 P_2 ... P_n$, by Theorem 1.5, P is a stable matrix. Consequently, $\exists \psi$, ψ is a probability distribution on $\langle\langle h \rangle\rangle^n$, such that

$$P = e'\psi.$$

It is easy to see that

$$\pi_x \left(P_t\right)_{xy} = \pi_y \left(P_t\right)_{yx}, \forall t \in \langle n \rangle, \forall x, y \in \langle\langle h \rangle\rangle^n.$$

This thing implies

$$\pi P_t = \pi, \forall t \in \langle n \rangle,$$

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 2.5 and its proof.

1. Any 1-step of the chain with transition matrix $P = P_1 P_2 ... P_n$ is performed via P_1, P_2, ..., P_n, *i.e.*, doing n transitions: one using P_1, one using P_2, ..., one using P_n. This chain:

a) attains its stationarity at time 1 — one step due to P or n steps due to P_1, P_2, ..., P_n;

b) belongs to our collection of hybrid Metropolis-Hastings chains from [9] (this follows from $U_{(x_1,x_2,...,x_{t+1})} \subset U_{(x_1,x_2,...,x_t)}$, $\forall t \in \langle n-1 \rangle$, $\forall x_1, x_2, ..., x_{t+1} \in \langle\langle h \rangle\rangle$, etc.; for our collection, see also [10–12]);

c) is a cyclic Gibbs sampler in a generalized sense (the state space is $\langle\langle h \rangle\rangle^n$) because

c1) the ratios used to define the transition probabilities of P_1 are similar to those of (usual) cyclic Gibbs sampler used to update coordinate 1 while the ratios used to define the transition probabilities of matrices P_t, $t \in \langle n \rangle - \{1\}$, are identical to those of (usual) cyclic Gibbs sampler (on $\langle\langle h \rangle\rangle^n$ too) used to update coordinates 2, 3, ..., n, respectively (the (usual) cyclic Gibbs sampler with finite state space also belongs to our collection of hybrid Metropolis-Hastings chains from [9], see [10]), and

c2) matrices P_t, $t \in \langle n \rangle$, are used cyclically.

(For finite Markov chain theory, see, *e.g.*, [2], and for the Gibbs sampler, see, *e.g.*, [5].)

2. Since $P = e' \pi$, we have

$$p_0 P = \pi, \ \forall p_0, \ p_0 = \text{ initial probability distribution.}$$

Therefore, theoretically speaking, it does not count the initial probability we choose. But, practically speaking, when we run the chain, it is suitable to work with the initial state $x = (0, 0, ..., 0)$ (in this case, $(p_0)_{(0,0,...,0)} = 1$).

3. Consider that the initial state of chain is $(0, 0, ..., 0)$. To generate a state of the chain at time 1, we use, for P_1, the uniform probability distribution

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

see the computation of matrix P_1 in the proof of Theorem 2.5 (the $0'$s do not count), while, for P_t, $t \in \langle n \rangle - \{1\}$, the probability distribution

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

see the computation of matrices P_t, $t \in \langle n \rangle - \{1\}$, in the proof of Theorem 2.5 (the $0'$s do not count too). The above probability distributions, the former being a uniform probability distribution and the latter being an almost uniform probability distribution — we call it *almost uniform probability distribution* because its components are identical, excepting at most one of them (all the components are identical when $\theta = 1$) — are, concerning the implementation, the best ones. To see that this is also true for the (above) almost uniform probability distribution, we split this probability distribution into two blocks,

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

If

$$X > \frac{1}{h\theta + 1}, \ X \sim U(0, 1),$$

further, we work with the latter block, which, by normalization, leads to the uniform probability distribution

$$\left(\frac{1}{h}, \frac{1}{h}, ..., \frac{1}{h} \right).$$

Therefore, our exact sampling Markovian method, having n steps (due to P_1, P_2, ..., P_n), is simple and good, very good. Another case which uses uniform and almost uniform probability distributions is presented in [11].

4. By $P = e'\pi$ we can compute the normalization constant Z. Indeed, since $\langle\langle h \rangle\rangle^n \supset U_{(0)} \supset U_{(0,0)} \supset ... \supset U_{(0,0,...,0)} = \{(0, 0, ..., 0)\}$ (recall that $U_{(x_1,x_2,...,x_{t+1})} \subset U_{(x_1,x_2,...,x_t)}$, $\forall t \in \langle n - 1 \rangle$, $\forall x_1, x_2, ..., x_{t+1} \in \langle\langle h \rangle\rangle$), P_t is a block diagonal matrix (eventually by permutation of rows and columns — P_t is a block diagonal matrix when it is used, *e.g.*, the lexicographic order (on $\langle\langle h \rangle\rangle^n$)), $\forall t \in \langle n \rangle - \{1\}$, and $P_t \in G_{\Delta_t,\Delta_{t+1}}$, $\forall t \in \langle n \rangle$ (moreover, P_t is a Δ_t-stable matrix on Δ_{t+1}, $\forall t \in \langle n \rangle$), we have (using $P = e'\pi$)

$$\pi_{(0,0,...,0)} = \frac{1}{h+1} \cdot \left(\frac{1}{h\theta + 1} \right)^{n-1}.$$

On the other hand,

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

So,

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

(To compute Z, we can use

$$\pi_{(1,1,\ldots,1)} = \frac{1}{Z}, \text{ or } \pi_{(2,2,\ldots,2)} = \frac{1}{Z}, \ \ldots, \ \text{ or } \pi_{(h,h,\ldots,h)} = \frac{1}{Z}$$

instead of $\pi_{(0,0,\ldots,0)} = \frac{1}{Z}$, etc.) From the above formula for Z, we can obtain other important things, such as (extending the physical terminology), the mean energy

$$\overline{H} = \sum_{x \in \langle\langle h \rangle\rangle^n} H(x)\, \pi_x = \sum_{x \in \langle\langle h \rangle\rangle^n} H(x) \frac{\theta^{H(x)}}{Z} = \frac{1}{Z} \sum_{x \in \langle\langle h \rangle\rangle^n} H(x)\, \theta^{H(x)}$$

(see, *e.g.*, [1, p. 6]). Indeed, since $\theta \in \mathbb{R}^+$, $\exists \beta \in \mathbb{R}$ such that $\theta = e^\beta$. In this case,

$$Z = Z(\beta) = (h+1)\left(he^\beta + 1\right)^{n-1},$$

so, differentiating with respect to β,

$$Z' = Z'(\beta) = (h+1)(n-1)\, he^\beta \left(he^\beta + 1\right)^{n-2}.$$

On the other hand,

$$Z = Z(\beta) = \sum_{x \in \langle\langle h \rangle\rangle^n} e^{\beta H(x)},$$

so,

$$Z' = Z'(\beta) = \sum_{x \in \langle\langle h \rangle\rangle^n} H(x)\, e^{\beta H(x)} = Z\overline{H}.$$

Finally, we have

$$\overline{H} = \frac{Z'}{Z} = \frac{(h+1)(n-1)\, he^\beta \left(he^\beta + 1\right)^{n-2}}{(h+1)\left(he^\beta + 1\right)^{n-1}} = \frac{(n-1)\, he^\beta}{he^\beta + 1};$$

using θ instead of e^β,

$$\overline{H} = \frac{(n-1)\, h\theta}{h\theta + 1}.$$

5. Using Uniqueness Theorem from [10] (the presentation of this result is too long, so, we omit to give it here), we can compute certain important probabilities for the Potts model on the star graph from Theorem 2.5. Indeed, by Uniqueness Theorem we have

$$P\left(U_{(x_1)}\right) = \sum_{x \in U_{(x_1)}} \pi_x = \frac{1}{h+1}, \ \forall x_1 \in \langle\langle h \rangle\rangle$$

(a generalization (for any graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ with $|\mathcal{E}| \geq 1$ ($\mathcal{V}$ is the vertex set and $\mathcal{E}$ is the edge set)) of this result is given in Theorem 2.3(v)). Further, by

Uniqueness Theorem we have

$$\frac{P\left(U_{(x_1,x_2)}\right)}{P\left(U_{(x_1)}\right)}=\frac{\sum\limits_{x\in U_{(x_1,x_2)}}\pi_x}{\sum\limits_{x\in U_{(x_1)}}\pi_x}=\begin{cases}\frac{1}{h\theta+1} & \text{if } x_2=x_1,\\ \frac{\theta}{h\theta+1} & \text{if } x_2\neq x_1,\end{cases}$$

$\forall x_1,x_2\in\langle\langle h\rangle\rangle$, so,

$$P\left(U_{(x_1,x_2)}\right)=\begin{cases}\frac{1}{(h+1)(h\theta+1)} & \text{if } x_2=x_1,\\ \frac{\theta}{(h+1)(h\theta+1)} & \text{if } x_2\neq x_1,\end{cases}$$

$\forall x_1,x_2\in\langle\langle h\rangle\rangle$. To compute $P\left(U_{(x_1,x_2,x_3)}\right)$, etc., we use (see Uniqueness Theorem)

$$\frac{P\left(U_{(x_1,x_2,\ldots,x_t)}\right)}{P\left(U_{(x_1,x_2,\ldots,x_{t-1})}\right)}=\frac{\sum\limits_{x\in U_{(x_1,x_2,\ldots,x_t)}}\pi_x}{\sum\limits_{x\in U_{(x_1,x_2,\ldots,x_{t-1})}}\pi_x}=\begin{cases}\frac{1}{h\theta+1} & \text{if } x_t=x_1,\\ \frac{\theta}{h\theta+1} & \text{if } x_t\neq x_1,\end{cases}$$

$\forall x_1,x_2,\ldots,x_t\in\langle\langle h\rangle\rangle$ $(3\leq t\leq n)$. Using the Kronecker delta (symbol),

$$\delta_{ij}=\begin{cases}1 & \text{if } i=j,\\ 0 & \text{if } i\neq j,\end{cases}$$

$\forall i,j\in I$, and Hamming distance,

$$d(x,y)=1-\delta_{x_1y_1}+1-\delta_{x_2y_2}+\ldots+1-\delta_{x_my_m},\ \forall x,y\in J_1\times J_2\times\ldots\times J_m$$

$(m\geq 1;\ x=(x_1,x_2,\ldots,x_m),\ \ldots)$, where I, J_1, J_2, ..., J_m are nonempty sets, we conclude that

$$P\left(U_{(x_1,x_2,\ldots,x_t)}\right)=\begin{cases}\frac{1}{h+1} & \text{if } t=1,\\ \frac{1}{h+1}\cdot\frac{\theta^{1-\delta_{x_1x_2}}}{h\theta+1}\cdot\frac{\theta^{1-\delta_{x_1x_3}}}{h\theta+1}\ldots\cdot\frac{\theta^{1-\delta_{x_1x_t}}}{h\theta+1} & \text{if } t\in\langle n\rangle-\{1\},\end{cases}$$

$$=\begin{cases}\frac{1}{h+1} & \text{if } t=1,\\ \frac{\theta^{d((x_1,x_1,\ldots,x_1),(x_2,x_3,\ldots,x_t))}}{(h+1)(h\theta+1)^{t-1}} & \text{if } t\in\langle n\rangle-\{1\}\end{cases}$$

$(n\geq 2)$, $\forall x_1,x_2,\ldots,x_t\in\langle\langle h\rangle\rangle$, because

$$P\left(U_{(x_1,x_2,\ldots,x_t)}\right)=P\left(U_{(x_1)}\right)\cdot\frac{P\left(U_{(x_1,x_2)}\right)}{P\left(U_{(x_1)}\right)}\cdot\ldots\cdot\frac{P\left(U_{(x_1,x_2,\ldots,x_t)}\right)}{P\left(U_{(x_1,x_2,\ldots,x_{t-1})}\right)},$$

$\forall t\in\langle n\rangle$, $\forall x_1,x_2,\ldots,x_t\in\langle\langle h\rangle\rangle$.

To illustrate Theorem 2.5, its proof, and the above comments, we give an example.

Example 2.6. Consider the Ising model ($h = 1$) on the star graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V} = \{V_1, V_2, V_3\}$ ($n = 3$) and $\mathcal{E} = \{[V_1, V_2], [V_1, V_3]\}$. To write the matrices P_1, P_2, P_3, we use the lexicographic order: $(0,0,0)$, $(0,0,1)$, $(0,1,0)$, $(0,1,1)$, $(1,0,0)$, $(1,0,1)$, $(1,1,0)$, $(1,1,1)$. By Theorem 2.5 (and its proof) we have

$$P_1 = \begin{pmatrix} \frac{1}{2} & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{2} \\ 0 & \frac{1}{2} & 0 & 0 & 0 & 0 & \frac{1}{2} & 0 \\ 0 & 0 & \frac{1}{2} & 0 & 0 & \frac{1}{2} & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{2} & 0 & 0 & \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{2} & 0 & 0 & 0 & 0 & \frac{1}{2} & 0 \\ \frac{1}{2} & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{2} \end{pmatrix},$$

$$P_2 = \begin{pmatrix} \frac{1}{\theta+1} & 0 & \frac{\theta}{\theta+1} & 0 & & & & \\ 0 & \frac{1}{\theta+1} & 0 & \frac{\theta}{\theta+1} & & & & \\ \frac{1}{\theta+1} & 0 & \frac{\theta}{\theta+1} & 0 & & & & \\ 0 & \frac{1}{\theta+1} & 0 & \frac{\theta}{\theta+1} & & & & \\ & & & & \frac{\theta}{\theta+1} & 0 & \frac{1}{\theta+1} & 0 \\ & & & & 0 & \frac{\theta}{\theta+1} & 0 & \frac{1}{\theta+1} \\ & & & & \frac{\theta}{\theta+1} & 0 & \frac{1}{\theta+1} & 0 \\ & & & & 0 & \frac{\theta}{\theta+1} & 0 & \frac{1}{\theta+1} \end{pmatrix},$$

$$P_3 = \begin{pmatrix} \frac{1}{\theta+1} & \frac{\theta}{\theta+1} & & & & & & \\ \frac{1}{\theta+1} & \frac{\theta}{\theta+1} & & & & & & \\ & & \frac{1}{\theta+1} & \frac{\theta}{\theta+1} & & & & \\ & & \frac{1}{\theta+1} & \frac{\theta}{\theta+1} & & & & \\ & & & & \frac{\theta}{\theta+1} & \frac{1}{\theta+1} & & \\ & & & & \frac{\theta}{\theta+1} & \frac{1}{\theta+1} & & \\ & & & & & & \frac{\theta}{\theta+1} & \frac{1}{\theta+1} \\ & & & & & & \frac{\theta}{\theta+1} & \frac{1}{\theta+1} \end{pmatrix}.$$

We have

$$U_{(0)} = \{(0,0,0),\ (0,0,1),\ (0,1,0),\ (0,1,1)\},$$
$$U_{(1)} = \{(1,0,0),\ (1,0,1),\ (1,1,0),\ (1,1,1)\},$$
$$U_{(0,0)} = \{(0,0,0),\ (0,0,1)\},\ U_{(0,1)} = \{(0,1,0),\ (0,1,1)\},$$
$$U_{(1,0)} = \{(1,0,0),\ (1,0,1)\},\ U_{(1,1)} = \{(1,1,0),\ (1,1,1)\},$$

$$U_{(0,0,0)} = \{(0,0,0)\}\,,\ U_{(0,0,1)} = \{(0,0,1)\}\,,\ ...,\ U_{(1,1,1)} = \{(1,1,1)\}\,,$$
$$\Delta_1 = \left(\langle\langle 1\rangle\rangle^3\right),\ \Delta_2 = \left(U_{(0)},\ U_{(1)}\right),$$
$$\Delta_3 = \left(U_{(0,0)},\ U_{(0,1)},\ U_{(1,0)},\ U_{(1,1)}\right),$$
$$\Delta_4 = \left(U_{(0,0,0)},\ U_{(0,0,1)},\ ...,\ U_{(1,1,1)}\right).$$

Obviously, $\Delta_4 \preceq \Delta_3 \preceq \Delta_2 \preceq \Delta_1$, $\Delta_4 \neq \Delta_3 \neq \Delta_2 \neq \Delta_1$. It is easy to see that $P_1 \in G_{\Delta_1,\Delta_2}$, $P_2 \in G_{\Delta_2,\Delta_3}$, $P_3 \in G_{\Delta_3,\Delta_4}$, and $\pi_x (P_t)_{xy} = \pi_y (P_t)_{yx}$, $\forall t \in \langle 3\rangle$, $\forall x,y \in \langle\langle 1\rangle\rangle^3$. By Theorem 2.5 or direct computation, $P = e'\pi$. Since $\pi_{(0,0,0)} = \frac{1}{Z}$, it is easy to see, using $P = e'\pi$, that $Z = 2(\theta+1)^2$ (Z can also be obtained by direct computation). Obviously, P_2 and P_3 are block diagonal matrices; P_2 is a Δ_2-stable matrix on Δ_2, P_3 is a Δ_3-stable matrix on Δ_3. Moreover, P_2 is a Δ_2-stable matrix on Δ_3, P_3 is a Δ_3-stable matrix. P_1 is a stable matrix both on Δ_1 and on Δ_2. By Uniqueness Theorem from [10] or direct computation we have

$$P\left(U_{(0)}\right) = P\left(U_{(1)}\right) = \frac{1}{2}$$

(this result also follows from Theorem 2.3(v)),

$$P\left(U_{(0,0)}\right) = P\left(U_{(1,1)}\right) = \frac{1}{2(\theta+1)},$$
$$P\left(U_{(0,1)}\right) = P\left(U_{(1,0)}\right) = \frac{\theta}{2(\theta+1)},$$
$$P\left(U_{(0,0,0)}\right) = \pi_{(0,0,0)} = \frac{1}{Z} = \frac{1}{2(\theta+1)^2},$$
$$P\left(U_{(0,0,1)}\right) = \pi_{(0,0,1)} = \frac{\theta}{Z} = \frac{\theta}{2(\theta+1)^2},\ \text{etc.}$$

Obviously, our exact sampling Markovian method has, here, 3 steps (due to P_1, P_2, P_3).

3. POTTS MODEL ON TREES

In this section, the study of Potts model on trees is reduced to that on star graphs.

Below we consider a tree and a star graph, but the reader could consider, first, a tree, the star graph being obtained from this tree deleting some edges and adding other ones — similar things hold for configurations, but deleting and adding numbers.

Let $n \geq 2$ ($n \in \mathbb{N}$). Let $\mathcal{G}_T = (\mathcal{V}_T, \mathcal{E}_T)$ be a tree, where $\mathcal{V}_T = \{N_1, N_2, ..., N_n\}$, the vertex (node) set, and $\mathcal{E}_T$ is the edge set. Consider the star graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V} = \{V_1, V_2, ..., V_n\}$, the vertex set, and $\mathcal{E} = \{[V_1, V_2], [V_1, V_3], ...,$

$[V_1, V_n]\}$, the edge set. We associate the tree with the star graph as follows. Consider $f = (f_v,\ f_e,\ f_c)$, where

$$f_v : \mathcal{V}_T \longrightarrow \mathcal{V},\ f_v(N_i) = V_i,\ \forall i \in \langle n \rangle$$

$(|\mathcal{V}_T| = |\mathcal{V}| = n)$,

$$f_e : \mathcal{E}_T \longrightarrow \mathcal{E},\ f_e([N_i, N_j]) = [V_1, V_j],$$

$\forall i, j \in \langle n \rangle$ with $[N_i, N_j] \in \mathcal{E}_T$ and $d(N_j, N_1) = d(N_i, N_1) + 1$ ("$[N_i, N_j] \in \mathcal{E}_T$ and $d(N_j, N_1) = d(N_i, N_1) + 1$" is equivalent to "$[N_i, N_j] \in \mathcal{E}_T$ and between N_i and N_j, N_i is the nearest to N_1"; d is the (usual) distance on graphs),

$$f_c : \langle\langle h \rangle\rangle^n \longrightarrow \langle\langle h \rangle\rangle^n\ \ (h \geq 1),\ f_c(x_1, x_2, ..., x_n) = (y_1, y_2, ..., y_n),$$

where the values $x_1, x_2, ..., x_n$ are associated with $N_1, N_2, ..., N_n$, respectively (x_1 is the color of N_1, etc.), and the values $y_1, y_2, ..., y_n$ are associated with $V_1, V_2, ..., V_n$, respectively,

$$y_1 = x_1,$$
$$y_s = (x_1 + x_{i_s} - x_s) \bmod (h+1),\ \forall s \in \langle n \rangle - \{1\},$$

i_s, which depends on $s \in \langle n \rangle - \{1\}$, is from $\langle n \rangle$ such that $[N_{i_s}, N_s] \in \mathcal{E}_T$ and $d(N_s, N_1) = d(N_{i_s}, N_1) + 1$ (i_s is unique with this property).

THEOREM 3.1. (i) *f_v, the function for vertices, is bijective.*

(ii) *f_e, the function for edges, is bijective.*

(iii) *f_c, the function for configurations, is bijective and, moreover, has an important property for energies, namely,*

$$H(f_c(x)) = H_T(x),\ \forall x \in \langle\langle h \rangle\rangle^n$$

($x = (x_1, x_2, ..., x_n)$), where H_T and H are energies for the tree and for its associate star graph, respectively.

Proof. (i) Obvious.

(ii) First, we show that f_e is injective. Let $[N_{i_1}, N_{j_1}], [N_{i_2}, N_{j_2}] \in \mathcal{E}_T$, $[N_{i_1}, N_{j_1}] \neq [N_{i_2}, N_{j_2}]$, $d(N_{j_1}, N_1) = d(N_{i_1}, N_1) + 1$, $d(N_{j_2}, N_1) = d(N_{i_2}, N_1) + 1$. We must show that

$$f_e([N_{i_1}, N_{j_1}]) \neq f_e([N_{i_2}, N_{j_2}]).$$

Since $f_e([N_{i_1}, N_{j_1}]) = [V_1, V_{j_1}]$ and $f_e([N_{i_2}, N_{j_2}]) = [V_1, V_{j_2}]$, we must show that

$$V_{j_1} \neq V_{j_2}\ \text{(equivalently, } j_1 \neq j_2\text{)}.$$

Since $N_{j_1} \neq N_{j_2}$ implies $V_{j_1} \neq V_{j_2}$, we show that

$$N_{j_1} \neq N_{j_2}.$$

Suppose that $N_{j_1} = N_{j_2}$.

Case 1. $N_{i_1} = N_{i_2}$. From $N_{i_1} = N_{i_2}$ and $N_{j_1} = N_{j_2}$, we have $[N_{i_1}, N_{j_1}] = [N_{i_2}, N_{j_2}]$. Contradiction.

Case 2. $N_{i_1} \neq N_{i_2}$. In a tree, any two vertices are connected by a unique path (a known result). Since $\mathcal{G}_T$ is a tree, it follows that there exists a unique path with ends N_1 and N_{j_1} ($N_{j_2} = N_{j_1}$; $N_{j_1} \neq N_1$ because $d(N_{j_1}, N_1) = d(N_{i_1}, N_1) + 1 \geq 1 > 0$). Consider this path. Since $[N_{i_1}, N_{j_1}] \in \mathcal{E}_T$ and $d(N_{j_1}, N_1) = d(N_{i_1}, N_1) + 1$, it follows that N_{i_1} belongs to this path. N_{i_2} belongs to this path too because $N_{j_2} = N_{j_1}$, $[N_{i_2}, N_{j_2}] \in \mathcal{E}_T$ and $d(N_{j_2}, N_1) = d(N_{i_2}, N_1) + 1$. Since $N_{j_1} = N_{j_2}$, we have $d(N_{j_1}, N_1) = d(N_{j_2}, N_1)$ and, further, $d(N_{i_1}, N_1) = d(N_{i_2}, N_1)$. Therefore, $N_{i_1} = N_{i_2}$. Contradiction.

From Cases 1 and 2, we conclude that $N_{j_1} \neq N_{j_2}$. Since $N_{j_1} \neq N_{j_2}$, we have $j_1 \neq j_2$ and, further,

$$f_e([N_{i_1}, N_{j_1}]) = [V_1, V_{j_1}] \neq [V_1, V_{j_2}] = f_e([N_{i_2}, N_{j_2}]).$$

Therefore, f_e is injective.

Second, since $f_e : \mathcal{E}_T \longrightarrow \mathcal{E}$, $|\mathcal{E}_T| = |\mathcal{E}|$, and f_e is injective, it follows that f_e is bijective.

(iii) First, we show that f_c is injective. Let a, $b \in \langle\langle h \rangle\rangle^n$, $a \neq b$ ($a = (a_1, a_2, ..., a_n)$, ...). We must show that

$$f_c(a) \neq f_c(b).$$

As $a \neq b$, it follows that $\exists s \in \langle n \rangle$ such that $a_s \neq b_s$.

Case 1. $a_1 \neq b_1$. Obviously, $f_c(a) \neq f_c(b)$.

Case 2. $a_1 = b_1$. As $a \neq b$, it follows that $\exists s \in \langle n \rangle - \{1\}$ such that $a_s \neq b_s$. Since $f_c : \langle\langle h \rangle\rangle^n \longrightarrow \langle\langle h \rangle\rangle^n$, it follows that $\exists u, w \in \langle\langle h \rangle\rangle^n$ ($u = (u_1, u_2, ..., u_n)$, ...) such that

$$f_c(a) = u, \ f_c(b) = w.$$

Recall that, in a tree, any two vertices are connected by a unique path. It follows that there exists a unique path with ends N_1 and N_s ($N_s \neq N_1$ because $s \in \langle n \rangle - \{1\}$). Consider this path. Consider that the vertices of this path are $N_{j_0} = N_1$, N_{j_1}, ..., $N_{j_{k-1}}$, $N_{j_k} = N_s$ ($k \geq 1$). Since $N_{j_0} = N_1$, we have $a_{j_0} = a_1$ and $b_{j_0} = b_1$. Set

$$t = \min_{1 \leq v \leq k} \{ j_v \mid a_{j_v} \neq b_{j_v} \}.$$

It follows that $t = j_v$ for some $v \in \langle k \rangle$. By the definition of f_c,

$$u_t = \left(a_1 + a_{j_{v-1}} - a_t\right) \bmod (h+1),$$
$$w_t = \left(b_1 + b_{j_{v-1}} - b_t\right) \bmod (h+1).$$

Suppose that $u_t = w_t$. It follows that

$$\left(a_1 + a_{j_{v-1}} - a_t\right) \bmod (h+1) = \left(b_1 + b_{j_{v-1}} - b_t\right) \bmod (h+1).$$

As $a_1 = b_1$ and $a_{j_{v-1}} = b_{j_{v-1}}$, we have

$$\left(a_1 + a_{j_{v-1}} - a_t\right) \bmod (h+1) = \left(a_1 + a_{j_{v-1}} - b_t\right) \bmod (h+1).$$

Further, we have

$$a_1 + a_{j_{v-1}} - a_t - q_1 (h+1) = a_1 + a_{j_{v-1}} - b_t - q_2 (h+1)$$

for some q_1, $q_2 \in \{-1, 0, 1\}$. Further, we have

$$b_t - a_t = (q_1 - q_2)(h+1)$$

for some q_1, $q_2 \in \{-1, 0, 1\}$. It follows that

$$h + 1 \mid b_t - a_t.$$

Contradiction (because

$$0 < |b_t - a_t| < h + 1).$$

Therefore, $u_t \neq w_t$. Further, we obtain $u \neq w$. Finally, we obtain $f_c(a) \neq f_c(b)$.

From Cases 1 and 2, we have $f_c(a) \neq f_c(b)$, so, f_c is injective.

Second, since $f_c : \langle\langle h \rangle\rangle^n \longrightarrow \langle\langle h \rangle\rangle^n$ and f_c is injective, it follows that f_c is bijective.

Third, we show that $H(f_c(x)) = H_T(x)$, $\forall x \in \langle\langle h \rangle\rangle^n$. Fix $x \in \langle\langle h \rangle\rangle^n$. Set

$$y = f_c(x)$$

($y = (y_1, y_2, ..., y_n)$ is the configuration of star graph corresponding to the configuration $x = (x_1, x_2, ..., x_n)$ of tree). Fix $i, j \in \langle n \rangle$ such that $[N_i, N_j] \in \mathcal{E}_T$ and $d(N_j, N_1) = d(N_i, N_1) + 1$. Then $f_e([N_i, N_j]) = [V_1, V_j]$. We show that

$$x_i \neq x_j \Longleftrightarrow y_1 \neq y_j.$$

"$\Longrightarrow$" Suppose that $y_1 = y_j$. We have

$$y_j = (x_1 + x_i - x_j) \bmod (h+1).$$

So, $\exists q \in \{-1, 0, 1\}$ such that

$$x_1 + x_i - x_j = q(h+1) + y_j.$$

Further, since $y_j = y_1 = x_1$, we have

$$x_i - x_j = q(h+1)$$

for some $q \in \{-1, 0, 1\}$, so,

$$h + 1 \mid x_i - x_j.$$

Contradiction (because

$$0 < |x_i - x_j| < h + 1).$$

"$\Longleftarrow$" Suppose that $x_i = x_j$. We have

$$y_j = (x_1 + x_i - x_j) \bmod (h+1) = x_1 \bmod (h+1) .$$

So,

$$y_j = x_1$$

because $x_1 \in \langle\langle h\rangle\rangle$. As $y_1 = x_1$ (see the definition of f_c), we have

$$y_1 = y_j .$$

Contradiction.

From the above equivalence, we obtain

$$H(y) = H_T(x) ,$$

i.e.,

$$H(f_c(x)) = H_T(x) . \quad \square$$

It is interesting to find functions for configurations, if any, better, computationally speaking, than f_c. If we replace

$$y_s = (x_1 + x_{i_s} - x_s) \bmod (h+1) , \ \forall s \in \langle n\rangle - \{1\} ,$$

from the definition of f_c with

$$y_s = (x_1 + x_{i_s} + x_s) \bmod (h+1) , \ \forall s \in \langle n\rangle - \{1\} ,$$

and keep the other things from the definition of f_c, we obtain another function for configurations. Is this function bijective? Does it have a property for energies similar to that for f_c from Theorem 3.1(iii)? The answers to these questions (when $h \geq 1$, when $h = 1$, ...) are left to the reader.

Let $f : A \longrightarrow B$ be a function. Let $C \subseteq B$. Recall that

$$f^{-1}(C) = \{x \mid x \in A \text{ and } f(x) \in C\} .$$

Below we give the main result about the Potts model on trees.

THEOREM 3.2. *Consider the above tree and its associate star graph. Suppose that*

$$\rho = (\rho_x)_{x \in \langle\langle h\rangle\rangle^n} , \ \rho_x = \frac{\theta^{H_T(x)}}{Z_T}, \ \forall x \in \langle\langle h\rangle\rangle^n ,$$

is the Potts model on the tree and

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

is the Potts model on its associate star graph, where H_T and H are the energies for the Potts model on the tree and for that on its associate star graph, respectively, and Z_T and Z are the normalization constants for the Potts model on the tree and for that on its associate star graph, respectively,

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

Then

(i)

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

(*since the normalization constant Z_T depends on n, h, and θ only, it follows that it is an invariant with respect to transformations which transform a tree into a tree with the same number of vertices — this is interesting*!);

(ii)

$$\rho_x = \pi_{f_c(x)}, \forall x \in \langle\langle h\rangle\rangle^n;$$

(iii)

$$f_c^{-1}\left(U_{(x_1)}\right) = U_{(x_1)}, \forall x_1 \in \langle\langle h\rangle\rangle,$$

where $U_{(x_1)}$ is the set defined in the proof of Theorem 2.5, $\forall x_1 \in \langle\langle h\rangle\rangle$;

(iv)

$$P\left(f_c^{-1}\left(U_{(x_1,x_2,...,x_t)}\right)\right) = P\left(U_{(x_1,x_2,...,x_t)}\right), \forall t \in \langle n\rangle, \forall x_1, x_2, ..., x_t \in \langle\langle h\rangle\rangle,$$

where $U_{(x_1,x_2,...,x_t)}$ is the set defined in the proof of Theorem 2.5, $\forall t \in \langle n\rangle$, $\forall x_1, x_2, ..., x_t \in \langle\langle h\rangle\rangle$,

$$P\left(U_{(x_1,x_2,...,x_t)}\right) = \sum_{y\in U_{(x_1,x_2,...,x_t)}} \pi_y, \ \forall t \in \langle n\rangle, \ \forall x_1, x_2, ..., x_t \in \langle\langle h\rangle\rangle,$$

and

$$P\left(f_c^{-1}\left(U_{(x_1,x_2,...,x_t)}\right)\right) = \sum_{z\in f_c^{-1}\left(U_{(x_1,x_2,...,x_t)}\right)} \rho_z, \ \forall t \in \langle n\rangle, \ \forall x_1, x_2, ..., x_t \in \langle\langle h\rangle\rangle.$$

Proof. (i) By Theorem 3.1(iii),

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

Since $f_c : \langle\langle h\rangle\rangle^n \to \langle\langle h\rangle\rangle^n$ and f_c is bijective (see Theorem 3.1(iii)), setting $y = f_c(x)$, we have

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

For $Z = (h+1)(h\theta+1)^{n-1}$, see Comment 4 from Section 2.

(ii) By (i) and Theorem 3.1(iii),

$$\rho_x = \frac{\theta^{H_T(x)}}{Z_T} = \frac{\theta^{H(f_c(x))}}{Z} = \pi_{f_c(x)}, \ \forall x \in \langle\langle h\rangle\rangle^n.$$

(iii) Let $x_1 \in \langle\langle h\rangle\rangle$. By the definition of f_c,

$$f_c^{-1}\left(U_{(x_1)}\right) \subseteq U_{(x_1)},$$

and, further, by Theorem 3.1(iii) (the fact that f_c is bijective),

$$\left|f_c^{-1}\left(U_{(x_1)}\right)\right| = \left|U_{(x_1)}\right|,$$

so,

$$f_c^{-1}\left(U_{(x_1)}\right) = U_{(x_1)}.$$

(iv) Let $t \in \langle n \rangle$ and $x_1, x_2, ..., x_t \in \langle\langle h \rangle\rangle$. By (ii) and Theorem 3.1(iii) (the fact that f_c is bijective),

$$P\left(f_c^{-1}\left(U_{(x_1,x_2,...,x_t)}\right)\right) = \sum_{z \in f_c^{-1}\left(U_{(x_1,x_2,...,x_t)}\right)} \rho_z = \sum_{z \in f_c^{-1}\left(U_{(x_1,x_2,...,x_t)}\right)} \pi_{f_c(z)} =$$

$$= \sum_{y=f_c(z),\ z \in f_c^{-1}\left(U_{(x_1,x_2,...,x_t)}\right)} \pi_y = \sum_{z=f_c^{-1}(y),\ z \in f_c^{-1}\left(U_{(x_1,x_2,...,x_t)}\right)} \pi_y =$$

$$= \sum_{y \in U_{(x_1,x_2,...,x_t)}} \pi_y = P\left(U_{(x_1,x_2,...,x_t)}\right). \quad \square$$

To generate a configuration of the above tree according to the Potts model on this tree, we proceed as follows. We first generate a configuration $(y_1, y_2, ..., y_n)$ of its associate star graph according to the Potts model on this star graph, see Section 2 — the generation method has n steps. $f_c^{-1}(y_1, y_2, ..., y_n)$ is the generated configuration of tree. Set $(x_1, x_2, ..., x_n) = f_c^{-1}(y_1, y_2, ..., y_n)$. Consider that N_1 is the root of tree. Using the levels of tree with respect to this root, one way to compute $x_2, x_3, ..., x_n$ $(x_1 = y_1)$ is: we first compute the x_i's of level 1 of tree, then compute the x_i's of level 2 of tree, etc. (see the definition of f_c). Now, we consider the preorder traversal of tree — suppose that this traversal is $N_1, N_{i_2}, N_{i_3}, ..., N_{i_n}$ (N_1 is the root of tree, $i_2, i_3, ..., i_n \in \langle n \rangle - \{1\}$, ...). In this case, we have another way to compute $x_2, x_3, ..., x_n$: we compute x_{i_2}, then compute x_{i_3}, ..., then compute x_{i_n} (see the definition of f_c again). (For the root, levels, and preorder traversal of a tree, see, *e.g.*, [4].)

Example 3.3. Consider the Ising model for n linear points (magnets) [3]. This model is a special Potts model, a model on the path graph (a tree) of above points — we label these points $N_1, N_2, ..., N_n$; $[N_1, N_2], [N_2, N_3], ...,$ $[N_{n-1}, N_n]$ are the edges of path graph. By Theorem 3.2(i), the normalization constant for this Ising model is

$$Z_T = 2(\theta + 1)^{n-1}.$$

(Z_T, in this special case, can be computed by a different method, see, *e.g.*, [7, pp. 33–36].) Further, proceeding (since Z_T is known) as in Comment 4 from Section 2, we have

$$\overline{H}_T = \frac{(n-1)\theta}{\theta + 1}$$

($\overline{H}_T$ = the mean energy for the above Ising model). To compute the probabilities $P\left(f_c^{-1}\left(U_{(x_1,x_2,...,x_t)}\right)\right)$, $t \in \langle n \rangle$, $x_1, x_2, ..., x_t \in \langle\langle 1 \rangle\rangle$, see Theorem 3.2(iv) and Comment 5 from Section 2. Note that, here, besides $f_c^{-1}\left(U_{(x_1)}\right) = U_{(x_1)}$, $\forall x_1 \in \langle\langle 1 \rangle\rangle$ (by Theorem 3.2(iii)), we have $f_c^{-1}\left(U_{(x_1,x_2)}\right) = U_{(x_1,x_2)}$, $\forall x_1, x_2 \in \langle\langle 1 \rangle\rangle$ (for the proof, use the fact that $[N_1, N_2]$ is an edge of the path graph and cases: $x_1 = x_2 = 0$; $x_1 = 0$, $x_2 = 1$; $x_1 = 1$, $x_2 = 0$; $x_1 = x_2 = 1$). To generate a configuration of this path graph according to the Ising model on this path graph, see before this example.

4. OTHER RESULTS

Section 1 refers, especially, to stochastic matrices — implicitly, to Markov chains. Sections 2 and 3 contain applications of Markov chains to the Potts model on trees (the star graphs are special trees). In this section, we give applications of the Potts on trees — for the Potts model on connected (nondirected simple finite) graphs, we give bounds, nontrivial bounds, for the normalization constant, for the product of normalization constant and mean energy (this product is equal to the derivative of normalization constant), for the mean energy, for the mean energy per site, for the free energy per site, and for the limit free energy per site. These bounds help us to understand the Potts model better.

In this section, as in Sections 2 and 3, we work with parameters h and θ in all cases on the Potts model(s).

The first result refers to the normalization constant.

THEOREM 4.1. *Let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ be a connected (nondirected simple finite) graph with vertex set $\mathcal{V} = \{V_1, V_2, ..., V_n\}$ and edge set $\mathcal{E}$. Suppose that $n \geq 2$ (equivalently, $|\mathcal{E}| \geq 1$ (because the graph is connected)). Let $\mathcal{G}_T = (\mathcal{V}_T, \mathcal{E}_T)$ be a spanning tree of $\mathcal{G}$ ($\mathcal{V}_T$ is the vertex set of $\mathcal{G}_T$ and $\mathcal{E}_T$ is the edge set of $\mathcal{G}_T$). Let Z and Z_T be the normalization constants for the Potts model on $\mathcal{G}$ and for that on $\mathcal{G}_T$, respectively.*

(i) *If $0 < \theta < 1$, then*

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

(ii) *If $\theta \geq 1$, then*

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

Proof. Since $Z = Z_T$ if $\mathcal{G}$ is a tree (in this case, $\mathcal{G} = \mathcal{G}_T$), further, we consider that $\mathcal{G}$ is not a tree. Now, since $\mathcal{G}$ is connected and is not a tree, we have $n \geq 3$. Consider the graph $\mathcal{G}_1 = (\mathcal{V}_1, \mathcal{E}_1)$, where $\mathcal{V}_1 = \mathcal{V}$ and $\mathcal{E}_1 = \mathcal{E} - \mathcal{E}_T$ ($|\mathcal{V}_1| = |\mathcal{V}| = n \geq 3$, $|\mathcal{E}_1| \geq 1$). Let H, H_T, and H_1 be the energies for the Potts model on $\mathcal{G}$, for that on $\mathcal{G}_T$, and for that on $\mathcal{G}_1$, respectively. (For the energy, see Section 2.) Obviously,

$$H = H_T + H_1.$$

(i) Since $0 < \theta < 1$, we have

$$\theta^{H(x)} \le \theta^{H_T(x)}, \ \forall x \in \langle\langle h \rangle\rangle^n .$$

So, by Theorem 3.2(i),

$$Z \le Z_T = (h+1)\,(h\theta+1)^{n-1} .$$

(ii) Since $\theta \ge 1$, we have

$$\theta^{H(x)} \ge \theta^{H_T(x)}, \ \forall x \in \langle\langle h \rangle\rangle^n .$$

So, by Theorem 3.2(i),

$$Z \ge Z_T = (h+1)\,(h\theta+1)^{n-1} . \quad \square$$

Example 4.2. Consider the d-dimensional grid graph $\mathcal{G}_{n_1,n_2,\dots,n_d}$ (d, n_1, n_2, ..., $n_d \ge 1$). Suppose that $0 < \theta < 1$ and $n_1 n_2 \dots n_d \ge 2$. By Theorem 4.1 we have

$$Z \le (h+1)\,(h\theta+1)^{n_1 n_2 \dots n_d - 1} .$$

For $h = 1$ (for the Ising model), we have

$$Z \le 2\,(\theta+1)^{n_1 n_2 \dots n_d - 1} .$$

To give bounds for the product $Z\overline{H}$ (Z and $\overline{H}$ will be specified below) and for the mean energy, we need the next two theorems.

THEOREM 4.3. *Let $\xi \in \mathbb{R}^+$ and s, $t \in \mathbb{N}$.*
(i) *If $0 < \xi \le \frac{1}{2}$ and $s \ge t \ge 1$, then*

$$s\xi^s \le t\xi^t.$$

(ii) *If $\xi \ge 1$ and $s \ge t \ge 0$, then*

$$s\xi^s \ge t\xi^t.$$

Proof. (i) Consider the sequence $(a_n)_{n\ge1}$, $a_n = n\xi^n$, $\forall n \ge 1$. This sequence is decreasing because

$$\frac{a_{n+1}}{a_n} = \frac{n+1}{n}\xi \le \frac{2n}{n}\xi = 2\xi \le 1, \forall n \ge 1.$$

Consequently, (i) holds.
(ii) Obvious. $\square$

Remark 4.4. (a) For a different proof of Theorem 4.3(i), one can consider the function $f : \mathbb{R} \longrightarrow \mathbb{R}$, $f(x) = x\xi^x$, then its derivative, etc.

(b) If $\frac{1}{2} < \xi < 1$, neither

$$s\xi^s \leq t\xi^t, \ \forall s, t, \ s \geq t \geq 1,$$

nor

$$s\xi^s \geq t\xi^t, \ \forall s, t, \ s \geq t \geq 1,$$

holds. Indeed, if $s = 2t$, we have

$$\frac{s\xi^s}{t\xi^t} = \frac{s}{t}\xi^{s-t} = 2\xi^t,$$

so, when $\xi = \frac{3}{4}$, for $t = 1$, we have

$$2\xi^t = 2 \cdot \frac{3}{4} = \frac{6}{4} > 1$$

while, for $t = 3$, we have

$$2\xi^t = 2 \cdot \left(\frac{3}{4}\right)^3 = 2 \cdot \frac{27}{64} = \frac{54}{64} < 1.$$

THEOREM 4.5. *Let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ be a connected graph with vertex set $\mathcal{V} = \{V_1, V_2, ..., V_n\}$ and edge set $\mathcal{E}$. Suppose that $n \geq 2$. Let $\mathcal{G}_T = (\mathcal{V}_T, \mathcal{E}_T)$ be a spanning tree of $\mathcal{G}$. Let H and H_T be the energies for the Potts model on $\mathcal{G}$ and for that on $\mathcal{G}_T$, respectively. Let $(x_1, x_2, ..., x_n) \in \langle\langle h\rangle\rangle^n$. Then*

(i)

$$H(x_1, x_2, ..., x_n) = 0 \Longleftrightarrow x_1 = x_2 = ... = x_n;$$

(ii)

$$H(x_1, x_2, ..., x_n) = 0 \Longleftrightarrow H_T(x_1, x_2, ..., x_n) = 0.$$

Proof. (i) "$\Longrightarrow$" Suppose that $\exists i, j \in \langle n\rangle$, $i \neq j$, such that $x_i \neq x_j$. Consider a path with ends V_i and V_j (the values $x_1, x_2, ..., x_n$ are associated with $V_1, V_2, ..., V_n$, respectively (x_1 is the color of V_1, etc.)). Consider that the vertices of this path are $V_{t_0} = V_i$, $V_{t_1}, ..., V_{t_{k-1}}$, $V_{t_k} = V_j$ $(k \geq 1)$. Since $x_i \neq x_j$, $\exists s \in \langle k\rangle$ such that $x_{t_{s-1}} \neq x_{t_s}$. It follows that $H(x_1, x_2, ..., x_n) > 0$. Contradiction.

"$\Longleftarrow$" Obvious.

(ii) $\mathcal{G}_T$ is also a connected graph. By (i),

$$H(x_1, x_2, ..., x_n) = 0 \Longleftrightarrow x_1 = x_2 = ... = x_n \Longleftrightarrow H_T(x_1, x_2, ..., x_n) = 0. \qquad \square$$

We now give bounds for the product $Z\overline{H}$ and for the mean energy.

THEOREM 4.6. *Let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ be a connected graph with vertex set $\mathcal{V} = \{V_1, V_2, ..., V_n\}$ and edge set $\mathcal{E}$. Suppose that $n \geq 2$. Let $\mathcal{G}_T = (\mathcal{V}_T, \mathcal{E}_T)$ be a* *spanning tree of $\mathcal{G}$. Let Z and Z_T be the normalization constants for the Potts model on $\mathcal{G}$ and for that on $\mathcal{G}_T$, respectively. Let H and H_T be the energies for the Potts model on $\mathcal{G}$ and for that on $\mathcal{G}_T$, respectively. Let $\overline{H}$ and $\overline{H}_T$ be the mean energies for the Potts model on $\mathcal{G}$ and for that on $\mathcal{G}_T$, respectively.*

(i) *If $0 < \theta \leq \frac{1}{2}$, then*

$$Z\overline{H} \leq Z_T\overline{H}_T = (n-1)(h+1)\,h\theta\,(h\theta+1)^{n-2},$$

and, as a result,

$$\overline{H} < (n-1)\,h\theta\,(h\theta+1)^{n-2}.$$

(ii) *If $\theta \geq 1$, then*

$$Z\overline{H} \geq Z_T\overline{H}_T = (n-1)(h+1)\,h\theta\,(h\theta+1)^{n-2}.$$

Proof. Obviously (see the proof of Theorem 4.1 ($H = H_T$ if $\mathcal{G}$ is a tree, ...)),

$$H(x) \geq H_T(x), \forall x \in \langle\langle h\rangle\rangle^n.$$

By Theorem 3.2(i), and proceeding as in Comment 4 from Section 2, we have

$$\overline{H}_T = \frac{(n-1)\,h\theta}{h\theta+1}$$

($\overline{H}_T$ is also an invariant with respect to transformations which transform a tree into a tree with the same number of vertices).

(i) By Theorems 3.2(i), 4.3(i), and 4.5(ii) we have

$$Z\overline{H} = Z\sum_{x\in\langle\langle h\rangle\rangle^n} H(x)\frac{\theta^{H(x)}}{Z} = \sum_{x\in\langle\langle h\rangle\rangle^n} H(x)\,\theta^{H(x)} \leq \sum_{x\in\langle\langle h\rangle\rangle^n} H_T(x)\,\theta^{H_T(x)} =$$

$$= Z_T\sum_{x\in\langle\langle h\rangle\rangle^n} H_T(x)\frac{\theta^{H_T(x)}}{Z_T} = Z_T\overline{H}_T =$$

$$= (h+1)(h\theta+1)^{n-1}\cdot\frac{(n-1)\,h\theta}{h\theta+1} = (n-1)(h+1)\,h\theta\,(h\theta+1)^{n-2}.$$

Since $Z > h+1$ (see Theorem 2.3(vi)), we have

$$\overline{H} \leq \frac{1}{Z}\cdot(n-1)(h+1)\,h\theta\,(h\theta+1)^{n-2} < (n-1)\,h\theta\,(h\theta+1)^{n-2}.$$

(ii) By Theorems 3.2(i) and 4.3(ii) we have

$$Z\overline{H} = \sum_{x\in\langle\langle h\rangle\rangle^n} H(x)\,\theta^{H(x)} \geq \sum_{x\in\langle\langle h\rangle\rangle^n} H_T(x)\,\theta^{H_T(x)} = Z_T\overline{H}_T =$$

$$= (n-1)(h+1)\,h\theta\,(h\theta+1)^{n-2}. \quad \square$$

Remark 4.7. (a) The case when $\theta = e^{-\frac{1}{kT}}$ is of interest to statistical physics, where k is the Boltzmann constant, $k = 1.38064852 \times 10^{-23}$, and T is the absolute temperature. This case leads to

$$0 < \theta \leq \frac{1}{2} \Longleftrightarrow 0 < T \leq \frac{1}{k \ln 2} \simeq 1.04 \times 10^{23}.$$

Consequently, the condition $0 < \theta \leq \frac{1}{2}$ from Theorem 4.6(i) is not too restrictive. (This condition is also in (b), Example 4.8, Theorem 4.10, and Remark 4.11.)

(b) (An application of Theorem 4.6.) By Theorem 4.6 we can find impossible cases for the couple $\left(Z, \overline{H}\right)$. *E.g.*, if $0 < \theta \leq \frac{1}{2}$, it is impossible that, *e.g.*,

$$Z > (h+1)(h\theta + 1)^{n-2} \text{ and } \overline{H} > (n-1)\, h\theta$$

simultaneously.

(c) (Another application of Theorem 4.6.) Since, replacing θ with e^{β}, the product $Z\overline{H}$ is equal to the derivative of Z with respect to β (hint: see Comment 4 from Section 2 again), the bounds for $Z\overline{H}$ from Theorem 4.6 lead to bounds for Z', then to information on the graph of function $\beta \longmapsto Z(\beta)$, $\beta \in \mathbb{R}$ $(0 < \theta < 1 \Longleftrightarrow \beta < 0;\ \theta \geq 1 \Longleftrightarrow \beta \geq 0)$.

Example 4.8. Consider the d-dimensional grid graph $\mathcal{G}_{n_1,n_2,\dots,n_d}$. Suppose that $0 < \theta \leq \frac{1}{2}$ and $n_1 n_2 \dots n_d \geq 2$. By Theorem 4.6 we have

$$\overline{H} < (n_1 n_2 \dots n_d - 1)\, h\theta\, (h\theta + 1)^{n_1 n_2 \dots n_d - 2}.$$

For $h = 1$, we have

$$\overline{H} < (n_1 n_2 \dots n_d - 1)\, \theta\, (\theta + 1)^{n_1 n_2 \dots n_d - 2}.$$

Below we give four definitions, two in the finite case and two in the limit. For other terminologies on the four notions below, see, *e.g.*, [6–7] and [14].

Let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ be a (connected or not) graph with vertex set $\mathcal{V} = \{V_1, V_2, \dots, V_n\}$ and edge set $\mathcal{E}$. Suppose that $|\mathcal{E}| \geq 1$. Consider the Potts model on $\mathcal{G}$. Set

$$f_{|\mathcal{V}|} = \frac{\ln Z}{|\mathcal{V}|}$$

(Z is the normalization constant). Since $|\mathcal{V}| = n$, we have

$$f_{|\mathcal{V}|} = f_n = \frac{\ln Z}{n}.$$

We call $f_{|\mathcal{V}|}$ (f_n) the *free energy per site.* Set

$$\overline{h}_{|\mathcal{V}|} = \frac{\overline{H}}{|\mathcal{V}|}$$

($\overline{H}$ is the mean energy). Since $|\mathcal{V}| = n$, we have

$$\overline{h}_{|\mathcal{V}|} = \overline{h}_n = \frac{\overline{H}}{n}.$$

We call $\overline{h}_{|\mathcal{V}|}$ ($\overline{h}_n$) the *mean energy per site*. Further, we consider a sequence of (connected or not) graphs $(\mathcal{G}_m)_{m\geq 1}$, $\mathcal{G}_m = (\mathcal{V}_m, \mathcal{E}_m)$, $\forall m \geq 1$, with $|\mathcal{E}_m| \geq 1$, $\forall m \geq 1$, and $|\mathcal{V}_m| \to \infty$ as $m \to \infty$ ($|\mathcal{V}_m| \to \infty$ as $m \to \infty \not\Rightarrow |\mathcal{E}_m| \to \infty$ as $m \to \infty$), such as, $(\mathcal{C}_n)_{n\geq 3}$, $\mathcal{C}_n$ is the cycle graph (with n vertices), $(\mathcal{G}_{n_1,n_2,...,n_d})_{n_1,n_2,...,n_d\geq 1,\ n_1 n_2 ... n_d \geq 2}$, $\mathcal{G}_{n_1,n_2,...,n_d}$ is the d-dimensional grid graph ($n_1 n_2 ... n_d \to \infty$ as $n_1, n_2, ..., n_d \to \infty$), $(\mathcal{K}_n)_{n\geq 2}$, $\mathcal{K}_n$ is the complete graph (with n vertices), $(\mathcal{K}_{m,n})_{m,n\geq 1}$, $\mathcal{K}_{m,n}$ is the complete bipartite graph (with $m+n$ vertices; $m, n \geq 1 \Longrightarrow m+n \geq 2$; $m+n \to \infty$ as $m, n \to \infty$), the sequences of triangular lattices which satisfy the above conditions, and the sequences of hexagonal lattices which satisfy the above conditions. (The framework on the graphs from this place is also more general than that from statistical physics.) For each $m \geq 1$, we consider the Potts model on $\mathcal{G}_m$. Set

$$f = \lim_{m\to\infty} f_{|\mathcal{V}_m|} \text{ if } \lim_{m\to\infty} f_{|\mathcal{V}_m|} \text{ exists}$$

and

$$\overline{h} = \lim_{m\to\infty} \overline{h}_{|\mathcal{V}_m|} \text{ if } \lim_{m\to\infty} \overline{h}_{|\mathcal{V}_m|} \text{ exists.}$$

We call, extending the physical terminology, f the *limit free energy per site* and $\overline{h}$ the *limit mean energy per site*.

The definitions of f and $\overline{h}$ make sense if for each of them, there exists at least one example. For such examples, see Theorems 4.12(iii) and 4.13.

THEOREM 4.9. *Let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ be a connected graph with vertex set $\mathcal{V} = \{V_1, V_2, ..., V_n\}$ and edge set $\mathcal{E}$. Suppose that $n \geq 2$ (equivalently, $|\mathcal{E}| \geq 1$). Consider the Potts model on $\mathcal{G}$.*

(i) *If $0 < \theta < 1$, then* ($|\mathcal{V}| = n$)

$$f_n \leq \frac{\ln\left[(h+1)(h\theta+1)^{n-1}\right]}{n}.$$

(ii) *If $\theta \geq 1$, then*

$$f_n \geq \frac{\ln\left[(h+1)(h\theta+1)^{n-1}\right]}{n}.$$

(iii) *If $\mathcal{G}$ is a tree, then* ($\theta \in \mathbb{R}^+$)

$$f_n = \frac{\ln\left[(h+1)(h\theta+1)^{n-1}\right]}{n}.$$

Proof. By Theorem 4.1. $\square$

THEOREM 4.10. *Let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ be a connected graph with vertex set $\mathcal{V} = \{V_1, V_2, ..., V_n\}$ and edge set $\mathcal{E}$. Suppose that $n \geq 2$. Consider the Potts model on $\mathcal{G}$.*

(i) *If $0 < \theta \leq \frac{1}{2}$, then*

$$\overline{h}_n < \frac{(n-1)\, h\theta\, (h\theta + 1)^{n-2}}{n}.$$

(ii) *If $\mathcal{G}$ is a tree, then $(\theta \in \mathbb{R}^+)$*

$$\overline{h}_n = \frac{(n-1)\, h\theta}{n\,(h\theta + 1)}.$$

Proof. (i) By Theorem 4.6.

(ii) By the proof of Theorem 4.6. $\square$

Remark 4.11. By Theorem 4.6 we can also find impossible cases for the couple $\left(f_n, \overline{h}_n\right)$. *E.g.*, for $0 < \theta \leq \frac{1}{2}$, it is impossible that, *e.g.* (see Remark 4.7(b)),

$$f_n > \frac{\ln\left[(h+1)\,(h\theta+1)^{n-2}\right]}{n} \text{ and } \overline{h}_n > \frac{(n-1)\, h\theta}{n}$$

simultaneously.

THEOREM 4.12. *Consider a sequence of connected graphs $(\mathcal{G}_m)_{m\geq 1}$, $\mathcal{G}_m = (\mathcal{V}_m, \mathcal{E}_m)$, $\forall m \geq 1$, with $|\mathcal{V}_m| \geq 2$ (equivalently, $|\mathcal{E}_m| \geq 1$), $\forall m \geq 1$, and $|\mathcal{V}_m| \to \infty$ as $m \to \infty$. For each $m \geq 1$, we consider the Potts model on $\mathcal{G}_m$. Suppose that $\lim_{m\to\infty} f_{|\mathcal{V}_m|}$ exists.*

(i) *If $0 < \theta < 1$, then $(f = \lim_{m\to\infty} f_{|\mathcal{V}_m|})$*

$$f \leq \ln(h\theta + 1).$$

(If $h = 1$, then

$$f \leq \ln(\theta + 1).)$$

(ii) *If $\theta \geq 1$, then*

$$f \geq \ln(h\theta + 1).$$

(iii) *If $\mathcal{G}_m$ is a tree, $\forall m \geq 1$, then*

$$f = \ln(h\theta + 1).$$

For $h = 1$ and $\mathcal{G}_m = \mathcal{P}_{m+1}$, $\forall m \geq 1$, $\mathcal{P}_{m+1}$ is the path graph (with $m+1$ vertices) – for the 1-dimensional Ising model –, we have

$$f = \ln(\theta + 1).$$

Proof. (i) By Theorem 4.9(i) we have

$$f_{|\mathcal{V}_m|} \leq \frac{\ln\left[(h+1)(h\theta+1)^{|\mathcal{V}_m|-1}\right]}{|\mathcal{V}_m|} = \\ = \frac{\ln(h+1)}{|\mathcal{V}_m|} + \left(1 - \frac{1}{|\mathcal{V}_m|}\right)\ln(h\theta+1),$$

so,

$$f = \lim_{|\mathcal{V}_m|\to\infty} f_{|\mathcal{V}_m|} \leq \ln(h\theta+1).$$

(ii) Similar to (i) (using Theorem 4.9(ii)).

(iii) By Theorem 4.9(iii). □

THEOREM 4.13. *Consider a sequence of trees* $(\mathcal{G}_m)_{m\geq 1}$, $\mathcal{G}_m = (\mathcal{V}_m, \mathcal{E}_m)$, $\forall m \geq 1$, *with* $|\mathcal{V}_m| \geq 2$, $\forall m \geq 1$, *and* $|\mathcal{V}_m| \to \infty$ *as* $m \to \infty$. *For each* $m \geq 1$, *we consider the Potts model on* $\mathcal{G}_m$. *Then*

$$\overline{h} = \frac{h\theta}{h\theta+1}.$$

For $h = 1$ *and* $\mathcal{G}_m = \mathcal{P}_{m+1}$, $\forall m \geq 1$, $\mathcal{P}_{m+1}$ *is the path graph — for the* 1-*dimensional Ising model —, we have*

$$\overline{h} = \frac{\theta}{\theta+1}.$$

Proof. By Theorem 4.10(ii). □

Remark 4.14. Other results can also be obtained:

a) the normalization constant, mean energy, etc. for the Potts model on forests (if $\mathcal{G}_F = (\mathcal{V}_F, \mathcal{E}_F)$ is a forest having the trees $\mathcal{G}_T^{(i)} = \left(\mathcal{V}_T^{(i)}, \mathcal{E}_T^{(i)}\right)$, $i \in \langle k \rangle$, with $\left|\mathcal{V}_T^{(i)}\right| \geq 2$, $\forall i \in \langle k \rangle$, where $k \geq 1$ ($k \in \mathbb{N}$), then

$$Z_F = Z_T^{(1)} Z_T^{(2)} \ldots Z_T^{(k)},$$

where Z_F is the normalization constant for the Potts model on $\mathcal{G}_F$ and $Z_T^{(i)}$ is the normalization constant for the Potts model on $\mathcal{G}_T^{(i)}$, $\forall i \in \langle k \rangle$);

b) bounds for the normalization constant, etc. for the Potts model on non-connected (nondirected simple finite) graphs using the Potts model on spanning forests.

REFERENCES

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

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

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

[4] D.E. Knuth, *The Art of Computer Programming.* Volume 1. *Fundamental Algorithms*, 3rd Edition. Addison-Wesley, Reading, 1997.

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

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

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

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

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

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

[11] 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.

[12] 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.

[13] R.B. Potts, *Some generalized order-disorder transitions.* Proc. Cambridge Philos. Soc. **48** (1952), 106–109.

[14] F.Y. Wu, *The Potts model.* Rev. Mod. Phys. **54** (1982), 235–268.

Received 8 February 2019

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