Annals of the University of Craiova, Mathematics and Computer Science Series
Volume 49(2), 2022, Pages 208–249, DOI: 10.52846/ami.v49i2.1494
ISSN: 1223-6934

# $\Delta$-wavy probability distributions and Potts model

Udrea Păun

Abstract. We define the wavy probability distributions on a subset and $\Delta$-wavy probability distributions — two generalizations of the wavy probability distributions. A classification on the $\Delta$-waviness is given. For the $\Delta$-wavy probability distributions having normalization constant, we give a formula for this constant, to compute this constant. We show that the Potts model is a $\Delta$-wavy probability distribution, where $\Delta$ is a partition which will be specified. For the normalization constant of Potts model, we give formulas and bounds. As to the formulas for this constant, we give two general formulas, one of them is simple while the other is more complicated, and based on independent sets, a formula for the Potts model on connected separable graphs — closed-form expressions are then obtained in several cases —, and a formula for the Potts model on graphs with a vertex of degree 2 — a recurrence relation is then obtained for the normalization constant of Potts model on $\mathcal{C}_n$, the cycle graph with $n$ vertices; the normalization constant of Ising model on $\mathcal{C}_n$ is computed using this relation. As to the bounds for the normalization constant, we present two ways to obtain such bounds; we illustrate these ways giving a general lower bound, and a lower bound and an upper one when the model is the Potts model on $\mathcal{G}_{n,n}$, the square grid graph, $n = 6k$, $k \geq 1$ — two upper bounds for the free energy per site of this model are then obtained, one of them being in the limit. A sampling method for the $\Delta$-wavy probability distributions is given and, as a result, a sampling method for the Potts is given. This method — that for the Potts model too — has two steps, Step 1 and Step 2, when $|\Delta| > 1$ and one step, Step 2 only, when $|\Delta| = 1$. For the Potts model, Step 1 is, in general, difficult. As to Step 2, for the Potts model too, using the Gibbs sampler in a generalized sense, we obtain an exact (not approximate) sampling method having $p+1$ steps ($p+1$ substeps of Step 2), where $p = |I|$, $I$ is an independent set, better, a maximal independent set, best, a maximum independent set — for the Potts model on $\mathcal{G}_{n_1,n_2,...,n_d}$, the $d$-dimensional grid graph, $d \geq 1$, $n_1, n_2, ..., n_d \geq 1$, $n_1 n_2 ... n_d \geq 2$, we obtain an exact sampling method for half or half+1 vertices.

2020 *Mathematics Subject Classification.* 60J10, 60J20, 82B20, 82B31, 82-10, 05C69, 05C90, 60E05, 62D05, 65C05.
*Key words and phrases.* Wavy probability distribution, wavy probability distribution on a subset, $\Delta$-wavy probability distribution, normalization constant, Gibbs sampler in a generalilized sense, sampling, Ising model, Potts model, independent set, connected separable graph, graph with a vertex of degree 2, grid graph, bound.

## 1. $\Delta$-wavy probability distributions

In this section, we present some basic things on nonnegative matrices, products of stochastic matrices, the hybrid Metropolis-Hastings chain(s), the Gibbs sampler(s) in a generalized sense, the wavy probability distributions, the wavy probability distributions on subsets, and the $\Delta$-wavy probability distributions. The notions of wavy probability distribution on a subset and of $\Delta$-wavy probability distribution together with the things concerning them are new — the most important things obtained are

Received May 26, 2021. Accepted November 9, 2022.

for the $\Delta$-wavy probability distributions: 1) a formula for the normalization constant for the $\Delta$-wavy probability distributions which have normalization constant; 2) a sampling method. Moreover, two results, one on our hybrid Metropolis-Hastings chain and the other on our Gibbs sampler in a generalized sense, are improved and a classification on the $\Delta$-waviness is given.

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. $(E) \in \operatorname{Par}(E)$; $(E)$ is the improper (degenerate) partition of $E$.

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

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

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

$$\langle m\rangle = \{1, 2, ..., m\}\,(m \in \mathbb{N},\ m \geq 1),$$
$$\langle\langle m\rangle\rangle = \{0, 1, ..., m\}\ \ (m \in \mathbb{N}),$$
$$N_{m,n} = \{P \,|P \text{ is a nonnegative } m \times n \text{ matrix}\},$$
$$S_{m,n} = \{P \,|P \text{ is a stochastic } m \times n \text{ matrix}\},$$
$$N_n = N_{n,n},$$
$$S_n = S_{n,n}.$$

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

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

Set

$$(\{i\})_{i\in\{s_1,s_2,...,s_t\}} = (\{s_1\}, \{s_2\}, ..., \{s_t\});$$
$$(\{i\})_{i\in\{s_1,s_2,...,s_t\}} \in \operatorname{Par}(\{s_1, s_2, ..., s_t\})\ \ (t \geq 1).$$

*E.g.*,

$$(\{i\})_{i\in\langle\langle n\rangle\rangle} = (\{0\}, \{1\}, ..., \{n\}).$$

**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.** ([13].) Let $P \in N_{m,n}$. Let $\Delta \in \operatorname{Par}(\langle m\rangle)$ and $\Sigma \in \operatorname{Par}(\langle n\rangle)$. We say that $P$ is a $[\Delta]$-*stable matrix on* $\Sigma$ 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.** ([13].) Let $P \in N_{m,n}$. Let $\Delta \in \operatorname{Par}(\langle m\rangle)$ and $\Sigma \in \operatorname{Par}(\langle n\rangle)$. We say that $P$ is a $\Delta$-*stable matrix on* $\Sigma$ if $\Delta$ is the least fine partition for which $P$ is a $[\Delta]$-stable matrix on $\Sigma$. In particular, a $\Delta$-stable matrix on $(\{i\})_{i\in\langle n\rangle}$ is called $\Delta$-*stable* while a $(\langle m\rangle)$-stable matrix on $\Sigma$ is called *stable on* $\Sigma$ 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 [13] for $G_{\Delta_1,\Delta_2}$ and [14] 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 $G_{\Delta_1,\Delta_2}$ 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 an important result — a beautiful result — on products of stochastic matrices.

**Theorem 1.1.** ([13].) *Let* $P_1 \in G_{(\langle m_1\rangle),\Delta_2} \subseteq S_{m_1,m_2}$, $P_2 \in G_{\Delta_2,\Delta_3} \subseteq S_{m_2,m_3}$, ..., $P_{n-1} \in G_{\Delta_{n-1},\Delta_n} \subseteq S_{m_{n-1},m_n}$, $P_n \in G_{\Delta_n,(\{i\})_{i\in\langle m_{n+1}\rangle}} \subseteq S_{m_n,m_{n+1}}$. *Then*

$$P_1P_2...P_n$$

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

*Proof.* See [13]. □

**Definition 1.5.** (See, e.g., [21, p. 80].) Let $P \in N_{m,n}$. We say that $P$ is a row-allowable matrix if it has at least one positive entry in each row.

Let $P \in N_{m,n}$. Set

$$\overline{P} = \left(\overline{P}_{ij}\right) \in N_{m,n},\ \overline{P}_{ij} = \begin{cases} 1 \text{ if } P_{ij} > 0, \\ 0 \text{ if } P_{ij} = 0, \end{cases}$$

$\forall i \in \langle m\rangle, \forall j \in \langle n\rangle$. We call $\overline{P}$ the *incidence matrix of* $P$ (see, *e.g.*, [8, p. 222]).

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

In this article, some statements on the matrices hold eventually by permutation of rows and columns. For simplification, further, we omit to specify this fact.

Warning! In this article, if a Markov chain has the transition matrix $P = P_1P_2...P_s$, where $s \geq 1$ and $P_1$, $P_2$, ..., $P_s$ are stochastic matrices, then any 1-step transition of this chain is performed via $P_1$, $P_2$, ..., $P_s$, *i.e.*, doing $s$ transitions: one using $P_1$, one using $P_2$, ..., one using $P_s$.

Let $S$ be a finite set with $|S| = r$, where $r \geq 2$ ($|\cdot|$ is the cardinal; for "$r \geq 2$", see below). Let $\pi = (\pi_i)_{i\in S}$ be a positive probability distribution on $S$. One way to sample approximately or, at best, exactly from $S$ is by means of our hybrid Metropolis-Hastings chain from [14]. Below we define this chain.

Let $E$ be a nonempty set. Set $\Delta \succ \Delta'$ if $\Delta' \preceq \Delta$ and $\Delta' \neq \Delta$, where $\Delta$, $\Delta' \in$Par$(E)$.

Let $\Delta_1$, $\Delta_2$, ..., $\Delta_{t+1} \in$Par$(S)$ with $\Delta_1 = (S) \succ \Delta_2 \succ ... \succ \Delta_{t+1} = (\{i\})_{i\in S}$, where $t \geq 1$. ($\Delta_1 \succ \Delta_2$ implies $r \geq 2$.) Let $Q_1$, $Q_2$, ..., $Q_t \in S_r$, $Q_1 = \left((Q_1)_{ij}\right)_{i,j\in S}$, $Q_2 = \left((Q_2)_{ij}\right)_{i,j\in S}$, ..., $Q_t = \left((Q_t)_{ij}\right)_{i,j\in S}$, such that

(C1) $\overline{Q}_1$, $\overline{Q}_2$, ..., $\overline{Q}_t$ are symmetric matrices;

(C2) $(Q_l)_K^L = 0, \forall l \in \langle t\rangle - \{1\}, \forall K, L \in \Delta_l, K \neq L$ (this condition implies that $Q_l$ is a block diagonal matrix and $\Delta_l$-stable matrix on $\Delta_l, \forall l \in \langle t\rangle - \{1\}$);

(C3) $(Q_l)_K^U$ is a row-allowable matrix, $\forall l \in \langle t\rangle$, $\forall K \in \Delta_l$, $\forall U \in \Delta_{l+1}$, $U \subseteq K$.

Define the matrices

$$P_l = \left( (P_l)_{ij} \right)_{i,j \in S},$$

$$(P_l)_{ij} = \begin{cases} 0 & \text{if } j \neq i \text{ and } (Q_l)_{ij} = 0, \\ (Q_l)_{ij} \min\left(1, \frac{\pi_j (Q_l)_{ji}}{\pi_i (Q_l)_{ij}}\right) & \text{if } j \neq i \text{ and } (Q_l)_{ij} > 0, \\ 1 - \sum\limits_{k \neq i} (P_l)_{ik} & \text{if } j = i, \end{cases}$$

$\forall l \in \langle t \rangle$. Set $P = P_1 P_2 ... P_t$.

The next result — a basic result — is an improvement of Theorem 2.3 from [14].

**Theorem 1.2.** *Concerning $P$ above we have — two general good things —*

$$\pi P = \pi \text{ and } P > 0.$$

*If, moreover,*

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

*then*

$$P_l = Q_l, \ \forall l \in \langle t \rangle \ (\text{ and, therefore, } P = Q_1 Q_2 ... Q_t)$$

*If, moreover,*

$$Q_l \in G_{\Delta_l, \Delta_{l+1}}, \forall l \in \langle t \rangle,$$

*then*

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

*and, as a result,*

$$P = e'\pi$$

(*therefore, in this case, the Markov chain with transition matrix $P$ attains its stationarity at time* 1, *its stationary probability distribution* (*limit probability distribution*) *being, obviously,* $\pi$).

*Proof.* For the first statement, see [14, Theorem 2.3]. The second statement is obvious (see the definition of matrices $P_l$, $l \in \langle t \rangle$). It is also obvious that $P_l \in G_{\Delta_l, \Delta_{l+1}}$, $\forall l \in \langle t \rangle$, if, moreover, $Q_l \in G_{\Delta_l, \Delta_{l+1}}$, $\forall l \in \langle t \rangle$. Further, using Theorem 1.1, $\exists \psi$, $\psi$ is a probability distribution on $S$, such that

$$P = e'\psi.$$

Further,

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

So,

$$P = e'\pi.$$

□

By Theorem 1.2 (by $\pi P = \pi$ and $P > 0$), $P^n \to e'\pi$ as $n \to \infty$. We call the Markov chain with transition matrix $P$ the *hybrid Metropolis-Hastings chain.* In particular, we call this chain the *hybrid Metropolis chain* when $Q_1$, $Q_2$, ..., $Q_t$ are symmetric matrices.

An important example of hybrid Metropolis-Hastings chain is presented in the next result. This result is an improvement of Theorem 2.3 from [18].

**Theorem 1.3.** *Consider a hybrid Metropolis-Hastings chain with state space $S$ ($S$ above, so, $|S| = r \geq 2$) 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 $\forall l \in \langle t \rangle$, $\forall i, j \in S$,*

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

*(see above for $Q_l$, $l \in \langle t \rangle$, $\pi = (\pi_i)_{i \in S}$, ...). Then*

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

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

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

*then*

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

*If, moreover,*

$$Q_l \in G_{\Delta_l, \Delta_{l+1}}, \ \forall l \in \langle t \rangle,$$

*then*

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

*and, as a result,*

$$P = e' \pi.$$

*Proof.* Theorem 2.3 from [18] and Theorem 1.2. □

We call the hybrid Metropolis-Hastings chain from Theorem 1.3 the *cyclic Gibbs sampler in a generalized sense* — the *Gibbs sampler in a generalized sense* for short.

It is worthy to note that Theorem 2.4 from [18] can also be improved; adding "*If, moreover,*

$$Q_l \in G_{\Delta_l, \Delta_{l+1}}, \ \forall l \in \langle t \rangle,$$

*then*

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

*and, as a result,*

$$P = e' \pi."$$

(see above for $Q_l$, $l \in \langle t \rangle$, $\Delta_l$, $l \in \langle t+1 \rangle$, ...), we obtain an improvement of it.

Further, we consider that $S = \{s_1, s_2, ..., s_r\}$, where $r \geq 2$ ($|S| = r$). Equip $S$ with an order relation, $\leqq$. Suppose that $s_1 \leqq s_2 \leqq ... \leqq s_r$. Let $\pi = (\pi_{s_i})_{i \in \langle r \rangle}$ be a positive probability distribution (on $S$). Let $\Delta_1$, $\Delta_2$, ..., $\Delta_{t+1}$ $\in$Par$(S)$ with $\Delta_1 = (S) \succ \Delta_2 \succ ... \succ \Delta_{t+1} = (\{s_i\})_{i \in \langle r \rangle}$, where $t \geq 1$ and $(\{s_i\})_{i \in \langle r \rangle} = (\{s_1\}, \{s_2\}, ..., \{s_r\})$. ($t \geq 1$ implies $t + 1 \geq 2$; further, $\Delta_1 \succ \Delta_2$ implies $r \geq 2$.) Consider that $\Delta_l = \left(K_1^{(l)}, K_2^{(l)}, ..., K_{u_l}^{(l)}\right)$, $K_1^{(l)}$ having the first $\left|K_1^{(l)}\right|$ elements of $S$, $K_2^{(l)}$ having the next

$\left|K_2^{(l)}\right|$ elements of $S$ (this condition and the next ones vanish when $l = 1$), ..., $K_{u_l}^{(l)}$ having the last $\left|K_{u_l}^{(l)}\right|$ elements of $S$, $\forall l \in \langle t+1 \rangle$. Consider that

(c1) $\left|K_1^{(l)}\right| = \left|K_2^{(l)}\right| = ... = \left|K_{u_l}^{(l)}\right|$, $\forall l \in \langle t+1 \rangle$ with $u_l \geq 2$;

(c2) $r = r_1 r_2 ... r_t$ with $r_1 r_2 ... r_l = |\Delta_{l+1}|$, $\forall l \in \langle t-1 \rangle$, and $r_t = \left|K_1^{(t)}\right|$.

We have

$$K_v^{(l)} = \bigcup_{w \in D_{v,b_l} \cup \{vb_l\}} K_w^{(l+1)}, \ \forall l \in \langle t \rangle, \ \forall v \in \langle u_l \rangle,$$

where

$$b_l = \frac{|\Delta_{l+1}|}{|\Delta_l|}, \ \forall l \in \langle t \rangle,$$

and

$$D_{v,b_l} = \{(v-1) b_l + 1, \ (v-1) b_l + 2, \ ..., \ vb_l - 1\}, \ \forall l \in \langle t \rangle, \ \forall v \in \langle u_l \rangle.$$

Suppose that $\forall l \in \langle t \rangle$, $\forall v \in \langle u_l \rangle$, $\forall w \in D_{v,b_l}$, $\exists \alpha_w^{(l,v)} > 0$ such that

$$\pi_{s_{i+d_w^{(l,v)}}} = \alpha_w^{(l,v)} \pi_{s_i} \text{ (direct proportionality)}, \forall i \in \langle r \rangle \text{ with } s_i \in K_{(v-1)b_l+1}^{(l+1)},$$

which, using vectors, is equivalent to

$$(\pi_{s_i})_{i \in \langle r \rangle, \ s_i \in K_{w+1}^{(l+1)}} = \alpha_w^{(l,v)} (\pi_{s_i})_{i \in \langle r \rangle, \ s_i \in K_{(v-1)b_l+1}^{(l+1)}},$$

where

$$d_w^{(l,v)} = \left|K_{(v-1)b_l+1}^{(l+1)}\right| + \left|K_{(v-1)b_l+2}^{(l+1)}\right| + ... + \left|K_w^{(l+1)}\right|,$$

$\forall l \in \langle t \rangle$, $\forall v \in \langle u_l \rangle$, $\forall w \in D_{v,b_l}$ — obviously,

$$(\pi_{s_i})_{i \in \langle r \rangle, \ s_i \in K_{(v-1)b_l+1}^{(l+1)}} = \left(\pi_{s_{j(l,v)}}, \pi_{s_{j(l,v)+1}}, ..., \pi_{s_{j(l,v)+d_{(v-1)b_l+1}^{(l,v)}-1}}\right),$$

$$(\pi_{s_i})_{i \in \langle r \rangle, \ s_i \in K_{w+1}^{(l+1)}} = \left(\pi_{s_{j(l,v)+d_w^{(l,v)}}}, \pi_{s_{j(l,v)+d_w^{(l,v)}+1}}, ..., \pi_{s_{j(l,v)+d_w^{(l,v)}+d_{(v-1)b_l+1}^{(l,v)}-1}}\right),$$

$\forall l \in \langle t \rangle$, $\forall v \in \langle u_l \rangle$, $\forall w \in D_{v,b_l}$, where $s_{j(l,v)}$ is the first element of $K_{(v-1)b_l+1}^{(l+1)}$, $\forall l \in \langle t \rangle$, $\forall v \in \langle u_l \rangle$, so,

$$j(l,v) = \begin{cases} 1 & \text{if } v = 1, \\ \left|K_1^{(l)}\right| + \left|K_2^{(l)}\right| + ... + \left|K_{v-1}^{(l)}\right| + 1 & \text{if } v \neq 1, \end{cases}$$

$\forall l \in \langle t \rangle$, $\forall v \in \langle u_l \rangle$.

**Definition 1.6.** ([19].) The probability distribution $\pi = (\pi_{s_i})_{i \in \langle r \rangle}$ having the above property (direct proportionality) we call the *wavy probability distribution* (*with respect to the order relation* $\leqq$ *and partitions* $\Delta_1, \Delta_2, ..., \Delta_{t+1}$).

For simple examples of wavy probability distributions, see [17]–[19].

In the next result, giving a wavy probability distribution, we construct a Gibbs sampler in a generalized sense which attains its stationarity at time 1... This chain is constructed using the $G$ method such that Theorem 1.1 can be applied.

**Theorem 1.4.** ([19].) *Let $\pi = (\pi_{s_i})_{i\in\langle r\rangle}$ be a wavy probability distribution (on $S$) with respect to the order relation $\leqq$ and partitions $\Delta_1, \Delta_2, ..., \Delta_{t+1}$ — for $S$,$\leqq$, ..., see Definition 1.6 and above this definition. Consider a Markov chain with state space $S$ and transition matrix $P = P_1P_2...P_t$ $(t \geq 1)$, where (the notation from Definition 1.6 and above this definition is again used)*

$$(P_l)_{s_{i+d_w^{(l,v)}}\to\xi} = \begin{cases} \dfrac{\pi_{s_{i+d_u^{(l,v)}}}}{\sum\limits_{z\in\{0\}\cup D_{v,b_l}} \pi_{s_{i+d_z^{(l,v)}}}} & \text{if } \xi = s_{i+d_u^{(l,v)}} \text{ for some } u \in \{0\}\cup D_{v,b_l}, \\ 0 & \text{if } \xi \neq s_{i+d_u^{(l,v)}},\ \forall u \in \{0\}\cup D_{v,b_l}, \end{cases}$$

*$\forall l \in \langle t\rangle$, $\forall v \in \langle u_l\rangle$, $\forall i \in \langle r\rangle$ with $s_i \in K^{(l+1)}_{(v-1)b_l+1}$, $\forall w \in \{0\}\cup D_{v,b_l}$, $\forall \xi \in S$, setting $d_0^{(l,v)} = 0$, $\forall l \in \langle t\rangle$, $\forall v \in \langle u_l\rangle$. Then this chain is a Gibbs sampler in a generalized sense and*

$$P = e'\pi$$

*(therefore, this chain attains its stationarity at time 1, its stationary probability distribution (limit probability distribution) being, obviously, $\pi$).*

*Proof.* See [19]. □

Theorem 1.4 leads to the next result.

**Theorem 1.5.** ([19].) *Let $\pi = (\pi_{s_i})_{i\in\langle r\rangle}$ be a wavy probability distribution (on $S$) with respect to the order relation $\leqq$ and partitions $\Delta_1, \Delta_2, ..., \Delta_{t+1}$ — for $S$, $\leqq$, ..., see Definition 1.6 and above this definition. Suppose that*

$$\pi_{s_i} = \frac{\nu_{s_i}}{Z},\ \forall i \in \langle r\rangle,$$

*where*

$$Z = \sum_{i\in\langle r\rangle} \nu_{s_i},$$

*$Z$ is the normalization constant ($Z \in \mathbb{R}^+$). Then*

$$Z = \nu_{s_1} \prod_{l\in\langle t\rangle} \left(1 + \sum_{w\in D_{1,b_l}} \alpha_w^{(l,1)}\right).$$

*Proof.* See [19]. □

Below we define two new notions, the wavy probability distribution on a subset and $\Delta$-wavy probability distribution, both being generalizations of the notion of wavy probability distribution.

**Definition 1.7.** Let $S$ be a finite set with $|S| \geq 2$. Let $\pi = (\pi_i)_{i\in S}$ be a positive probability distribution (on $S$). Let $A \subseteq S$ with $|A| \geq 2$. Equip $A$ with an order relation, $\overset{A}{\leqq}$. Let $\Delta_1, \Delta_2, ..., \Delta_{t+1} \in$Par$(A)$ with $\Delta_1 = (A) \succ \Delta_2 \succ ... \succ \Delta_{t+1} = (\{i\})_{i\in A}$, where $t \geq 1$. We say that $\pi$ is a *wavy probability distribution on $A$* (*with respect to the order relation $\overset{A}{\leqq}$ and partitions $\Delta_1, \Delta_2, ..., \Delta_{t+1}$*) if

$$\pi|_A = \left(\frac{\pi_i}{P(A)}\right)_{i\in A},$$

the normalized restriction of $\pi$ to $A$, is a wavy probability distribution (on $A$) with respect to the order relation $\overset{A}{\leqq}$ and partitions $\Delta_1$, $\Delta_2$, ..., $\Delta_{t+1}$, where

$$P(A) = \sum_{k \in A} \pi_k.$$

In particular, a wavy probability distribution on $S$ ($A = S$) is called *wavy probability distribution* for short.

In the above definition, we considered a subset $A$ of $S$ with $|A| \geq 2$. If $A$ is a subset of $S$ with $|A| = 1$, we say by convention that $\pi$ is a *wavy probability distribution on* $A$. We need these improper (degenerate) wavy probability distributions on subsets for the $\Delta$-wavy probability distributions, see below.

**Remark 1.1.** If $\pi = (\pi_i)_{i \in S}$ is a wavy probability distribution on $A$, $A \subseteq S$ with $|A| \geq 2$, then Theorem 1.4 holds, in particular, for $\pi|_A$. If, moreover, $\pi$ has the normalization constant, say, $Z$, then $ZP(A)$ is the normalization constant of $\pi|_A$, and, using Theorem 1.5, we can compute $ZP(A)$.

**Definition 1.8.** Let $S$ be a finite set with $|S| \geq 2$. Let $\pi = (\pi_i)_{i \in S}$ be a positive probability distribution (on $S$). Let $\Delta \in \mathrm{Par}(S)$. We say that $\pi$ is a $\Delta$-*wavy probability distribution* (*on* $S$) if $\pi$ is a wavy probability distribution on $L$, $\forall L \in \Delta$. In particular, a $(S)$-wavy probability distribution is called *wavy probability distribution* for short.

Consider that $\pi$ is a wavy probability distribution on $A$, $A \subseteq S$ ($|A| \geq 1$). Obviously, when $|A^c| \geq 2$ ($A^c$ is the complement of $A$), $\pi$ is a wavy probability distribution on $A^c$ with respect to the partitions $(A^c)$ and $(\{x\})_{x \in A^c}$ and an order relation on $A^c$ fixed — any order relation on $A^c$ fixed is good when the partitions are $(A^c)$ and $(\{x\})_{x \in A^c}$. (When $|A^c| = 1$, $\pi$ is, by convention, a wavy probability distribution on $A^c$.) Using these things, $\pi$ is a $(A, A^c)$-wavy probability distribution. This is a simple case in which from a wavy probability distribution on a subset we obtain a $\Delta$-wavy probability distribution.

The probability distributions on sets with one element are *improper* (*degenerate*). Further, we consider finite sets with at least two elements — let $S$ be a finite set with $|S| \geq 2$. Any probability distribution on $S$ is a wavy probability distribution with respect to (the partitions) $(S)$ and $(\{i\})_{i \in S}$. A wavy probability distribution on $S$ with respect to $(S)$ and $(\{i\})_{i \in S}$ is called *trivial* — no order relation on $S$ is mentioned; any order relation on $S$ can be used when the partitions are $(S)$ and $(\{i\})_{i \in S}$. A probability distribution on $S$ with $|S| \geq 3$ which is wavy only with respect to $(S)$ and $(\{i\})_{i \in S}$ is called *w-irregular* (warning! the notion is *w-irregular probability distribution,* not w-irregular wavy probability distribution; *w-irregular* comes from *wavy/waviness* and *irregular*). The class of wavy probability distributions with respect to three or more partitions (the class of nontrivial wavy probability distributions) is the best class of $\Delta$-wavy probability distributions — the probability distributions from this class can lead to good or very good results, see, *e.g.*, [19] and, here (for three partitions), Theorem 2.8, the first paragraph after the proof of Theorem 2.8, the last paragraph from Section 2, and Sections 3 and 4. Any probability distribution on $S$ is a wavy probability distribution on $A$ with respect to (the partitions) $(A)$ and $(\{i\})_{i \in A}$, $\forall A \subseteq S$ with $|A| \geq 2$. A wavy probability distribution on a subset, $A$, of $S$ with respect to $(A)$

and $(\{i\})_{i\in A}$ (this condition implies $|A| \geq 2$) is called *trivial*. A wavy probability distribution on a subset, $A$, of $S$ with $|A| = 1$ (see the convention after Definition 1.7) is by convention called *trivial*. A probability distribution which is wavy on a subset, $A$, of $S$ with $|A| \geq 3$ only with respect to $(A)$ and $(\{i\})_{i\in A}$ is called *w-irregular on* $A$. In particular, a w-irregular probability distribution on $S$ ($A = S$) is called *w-irregular probability distribution* for short. A $\Delta$-wavy probability distribution on $S$ is called *trivial* if it is a wavy probability distribution on $L$ with respect to $(L)$ and $(\{i\})_{i\in L}$, $\forall L \in \Delta$ with $|L| \geq 2$ (any $(\{i\})_{i\in S}$-wavy probability distribution (on $S$) is automatically trivial). A probability distribution on $S$ is called $\Delta$-*w-irregular* if it is a probability distribution which is wavy on $L$ only with respect to $(L)$ and $(\{i\})_{i\in L}$, $\forall L \in \Delta$ with $|L| \geq 3$; we consider that $\exists L \in \Delta$ such that $|L| \geq 3$. In particular, a $(S)$-w-irregular probability distribution is called *w-irregular probability distribution* for short. A probability distribution on $S$ with $|S| \geq 3$ is called *W-complicated* if it is a $\Delta$-w-irregular probability distribution, $\forall \Delta \in \mathrm{Par}(S)$ with the property that $\exists L \in \Delta$ such that $|L| \geq 3$. Excepting the case when $|S|$ is sufficiently small, the class of W-complicated probability distributions is, on the $\Delta$-waviness, the worst class of probability distributions. Some W-complicated probability distributions can be transformed into good $\Delta$-wavy probability distributions — do not forget this idea!; for an example, see the example for the alias method in [10, pp. 25−27] and [15, pp. 422−424]. For our interest (for sampling, ...), it is important that the probability distributions on finite sets with at least two elements be $\Delta$-wavy probability distributions with $|\Delta|$ as small as possible — some excepted cases can appear, see, *e.g.*, the first paragraph after the proof of Theorem 2.5 (in Section 2). To complete our classification on the $\Delta$-waviness (the waviness is a special case of the $\Delta$-waviness; the waviness on a subset can be considered, see the previous paragraph, as being a special case of the $\Delta$-waviness), we must say one thing more. From the above notions, using the prefix "non", we derive others: nontrivial wavy probability distribution, non-w-irregular probability distribution, etc.

Below we give a basic result to compute normalization constants.

**Theorem 1.6.** *Let* $\pi = (\pi_i)_{i\in S}$ *be a* $\Delta$*-wavy probability distribution* ($S$ *is a finite set with* $|S| \geq 2$; $\Delta \in \mathrm{Par}(S)$). *Suppose that* $\pi$ *has the normalization constant, say,* $Z$. *Then*

$$Z = \sum_{L\in\Delta} ZP(L),$$

*where*

$$P(L) = \sum_{k\in L} \pi_k, \ \forall L \in \Delta.$$

*Proof.* Since

$$\sum_{L\in\Delta} P(L) = 1,$$

we have

$$Z = \sum_{L\in\Delta} ZP(L).$$

□

**Remark 1.2.** If $\pi = (\pi_i)_{i \in S}$ is a $\Delta$-wavy probability distribution (on $S$), then Theorem 1.4 holds, in particular, for

$$\pi|_L = \left( \frac{\pi_i}{P(L)} \right)_{i \in L}, \forall L \in \Delta \text{ with } |L| \geq 2.$$

**Remark 1.3.** Let $\pi = (\pi_i)_{i \in S}$ be a $\Delta$-wavy probability distribution (on $S$) with normalization constant $Z$.

(a) By Theorem 1.5 we can compute $ZP(L)$ for some $L \in \Delta$ with $|L| \geq 2$ or for all $L \in \Delta$ with $|L| \geq 2$; $ZP(L) = \nu_i$ if $L = \{i\}$ and $\pi_i = \frac{\nu_i}{Z}$.

(b) If we can compute $ZP(L)$ for all $L \in \Delta$, then, using Theorem 1.6, we could compute $Z$.

(c) If we know $Z$ and $ZP(L)$ for some $L \in \Delta$, then we can compute the probability $P(L)$. If we know $Z$ and $ZP(L)$ for all $L \in \Delta$, then we can compute the probabilities $P(L)$, $L \in \Delta$.

For the $\Delta$-wavy probability distributions, below we give a sampling method having one step when $|\Delta| = 1$ and two steps when $|\Delta| > 1$.

Let $\pi = (\pi_i)_{i \in S}$ be a $\Delta$-wavy probability distribution (on $S$). Let $X$ be a random variable with probability distribution $\pi$. We generate the random variable $X$ as follows.

*Step* 1 (when $|\Delta| > 1$). Sample from $\Delta$ according to the probability distribution $\tau = (\tau_L)_{L \in \Delta}$ (on $\Delta$), where

$$\tau_L = P(L) = \sum_{k \in L} \pi_k, \ \forall L \in \Delta \ (\tau_L > 0, \ \forall L \in \Delta).$$

Suppose that the result of sampling is, say, $A$ $(A \in \Delta)$.

*Step* 2. Sample from $A$ according to the probability distribution

$$\pi|_A = \left( \frac{\pi_i}{\tau_A} \right)_{i \in A} = \left( \frac{\pi_i}{P(A)} \right)_{i \in A}.$$

Suppose that the result of sampling is, say, $j$ $(j \in A \subseteq S)$.

Set $X = j$ — this value of $X$ is generated according to the $\Delta$-wavy probability distribution $\pi$ ($j$ is the result of sampling from $S$ according to $\pi$) because by general multiplicative formula (see, *e.g.*, [8. p. 26])

$$P(X = j) = P(X \in \{j\}) = P(X \in \{j\} \cap A) = P(\{X \in \{j\}\} \cap \{X \in A\}) =$$
$$= P(X \in A) P(X \in \{j\} | X \in A) = \tau_A \cdot \frac{\pi_j}{\tau_A} = \pi_j.$$

To use the above sampling method exactly or approximately, we must use other exact or approximate sampling methods — examples of methods which could be used: the inversion method, rejection method, $G$ method, method based on our Gibbs sampler in a generalized sense (Theorem 1.4 could be used at Step 2 (because $\pi|_L$ is a wavy probability distribution, $\forall L \in \Delta$ with $|L| \geq 2$ — when $|L| = 1$, $\pi|_L$ is an improper probability distribution, so, no problem, no theorem (result) is necessary) and at Step 1, in the latter case when $\tau$ is a nontrivial wavy probability distribution or, more generally, when $\tau$ is a nontrivial $\Gamma$-wavy probability distribution, $|\Gamma|$ being sufficiently small), and method based on our hybrid Metropolis-Hastings chain with $P^*$ (see [14]–[15] for this chain). The last three methods are exact when Theorem 1.1, practically speaking, can be applied. The $G$ method is neither the method based on our Gibbs sampler in a generalized sense nor the method based on our hybrid Metropolis-Hastings chain (with or without $P^*$), but it together with the Gibbs sampler in a generalized sense or, more generally, with the hybrid Metropolis-Hastings chain can give good or very good results, see, *e.g.*, [16] — the Gibbs sampler in a generalized sense from there was constructed taking into account Theorem 1.1.

## 2. Potts model

In this section, we present a few things about graphs. We then consider the Potts model together with some basic results about it, the best results being: 1) the Potts model on an arbitrary but fixed graph is a $\Delta$-wavy probability distribution, where $\Delta$, which depends on the fixed graph, is a partition which will be specified; 2) two formulas for the normalization constant of Potts model (on an arbitrary but fixed graph), one of them is simple while the other is more complicated, and based on independent sets — we also give the steps we need to compute this constant by the more complicated formula.

In this article, we work with nondirected simple finite graphs excepting Section 4, where we will work with nondirected simple finite graphs and nondirected finite multigraphs without loops. (A *simple graph* is a graph without multiple edges and loops.) Moreover, we work with nonempty graphs, *i.e.*, with graphs which have at least one edge. (For the graph theory, see, *e.g.*, [4], [5], and [22].)

Let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ be a (nonempty nondirected simple finite) graph, where $\mathcal{V}$ is the vertex set ($|\mathcal{V}| \geq 2$) and $\mathcal{E}$ is the edge set ($|\mathcal{E}| \geq 1$). Below we give a few definitions, the most important for the Potts model being that of maximum independent set. Some simple and not too simple results are also considered.

Let $(X, Y) \in \text{Par}(\mathcal{V})$ (consequently, $X, Y \neq \emptyset$). $(X, Y)$ is called a *bipartition* (*of the graph* $\mathcal{G}$) if each edge of $\mathcal{G}$ has a vertex (end) in $X$ and a vertex (end) in $Y$. The graph $\mathcal{G}$ is called *bipartite* if it has at least one bipartition. (For the notions of bipartition and of bipartite graph and their definitions, see, *e.g.*, [22, p. 51].) If the graph $\mathcal{G}$ is connected and bipartite, then it has a unique bipartition (see, *e.g.*, [1, p. 8]). The graph $\mathcal{G}$ is bipartite if and only if it contains no odd cycles (see, *e.g.*, [4, p. 54]). If the graph $\mathcal{G}$ is bipartite, then it is isomorphic to a spanning subgraph of a complete bipartite graph. (For isomorphic graphs, see, *e.g.*, [4, p. 40].)

Let $\emptyset \neq I \subseteq \mathcal{V}$. $I$ is called an *independent set* (*of vertices of the graph* $\mathcal{G}$) if each edge of $\mathcal{G}$ has at most one vertex (end) in $I$. This notion is a central one in the graph theory. Obviously, if $I$ is an independent set, then $|I| < |\mathcal{V}|$ (equivalently, $I \subset \mathcal{V}$). Obviously, the graph $\mathcal{G}$ has at least $|\mathcal{V}|$ independent sets (because if $V \in \mathcal{V}$, then $\{V\}$ is an independent set). If $I$ is an independent set, then it is called a *maximal independent set* if $\forall J$, $I \subset J \subseteq \mathcal{V}$, $J$ is not an independent set. If $I$ is a maximal independent set of maximum cardinality, then it is called a *maximum independent set*. (For the above notions and their definitions, see, *e.g.*, [2, pp. 70−71], [6], and [7, pp. 461−462] — see also Internet (Wikipedia, etc.; some books are available).)

If the graph $\mathcal{G}$ is connected and bipartite, and has the bipartition $(X, Y)$, then

(i) $X$ and $Y$ are maximum independent sets if $|X| = |Y|$;

(ii) $X$ is the maximum independent set while $Y$ is a maximal independent set if $|X| > |Y|$.

An interesting example of connected and bipartite graph is $\mathcal{G}_{n_1,n_2,...,n_d}$, the $d$-dimensional grid graph (with dimensions $n_1$, $n_2$, ..., $n_d$), $d \geq 1$, $n_1$, $n_2$, ..., $n_d \geq 1$, $n_1 n_2 ... n_d \geq 2$.

Further, we consider a (nonempty nondirected simple finite) graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V} = \{V_1, V_2, ..., V_n\}$ is the vertex set and $\mathcal{E}$ is the edge set ($|\mathcal{E}| \geq 1$; $|\mathcal{E}| \geq 1 \Longrightarrow n \geq 2$). Moreover, since for the Potts model the isolated vertices count, but at least for the normalization constant it is sufficient to consider graphs without isolated vertices (see Remark 3.1), further, we consider, for simplification, that the graph $\mathcal{G}$ has no isolated vertices. (An *isolated vertex* is a vertex of degree 0, see, *e.g.*, [4, p. 20].) $[V_i, V_j]$ 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* $\mathcal{G}$). $\langle\langle h \rangle\rangle$ 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.*, [10, 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 *energy* or *Hamiltonian*; $H(x)$ represents the *energy* (or *Hamiltonian*) *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 [20], see, *e.g.*, also [10, Chapter 6], [11], and [23] — we extend this notion considering $\theta \in \mathbb{R}^+$. In particular, if $h = 1$ and $0 < \theta < 1$, $\pi$ is called the *Ising model* (*on the graph* $\mathcal{G}$), see [9], see, *e.g.*, also [10, Chapter 6] and [12] (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 *normalizing constant,* or, extending the physical terminology, *partition function* (*of* (or *for the*) *Potts model*). In the theory of Potts model, $Z$ is a central object (see, *e.g.*, also [3, p. 6]), so, its computation is a fundamental problem.

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

The next result is simple, but very useful — a basic result about $H$, about the Potts model on graphs.

**Theorem 2.1.**

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

*For* $h = 1$, *we have*

$$H(x_1, x_2, ..., x_n) = H(x_1 \oplus 1, x_2 \oplus 1, ..., x_n \oplus 1) = H(\bar{x}_1, \bar{x}_2, ..., \bar{x}_n),$$

$\forall x_1, x_2, ..., x_n \in \langle\langle 1 \rangle\rangle$, *where*

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

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

*Proof.* See [16]. □

Let $x = (x_1, x_2, ..., x_n) \in \mathbb{R}^n$. Let $a \in \mathbb{R}$. Let $k \in \langle n \rangle$. Set

$$x^{(a|k)} = \left(x_1^{(a|k)}, x_2^{(a|k)}, ..., x_n^{(a|k)}\right) \in \mathbb{R}^n,$$

where

$$x_i^{(a|k)} = \begin{cases} x_i & \text{if } i \neq k, \\ a & \text{if } i = k, \end{cases}$$

$\forall i \in \langle n \rangle$. Therefore, $x^{(a|k)} = (x_1, x_2, ..., x_{k-1}, a, x_{k+1}, ..., x_n)$ ($x_1$, $x_2$, ..., $x_{k-1}$ vanish if $k = 1$ and $x_{k+1}$, ..., $x_n$ vanish if $k = n$).

Let $V_j$, $V_k \in \mathcal{V}$, $j$, $k \in \langle n \rangle$, $j \neq k$. We say that $V_k$ is *adjacent to* $V_j$ if $[V_j, V_k] \in \mathcal{E}$. Obviously, $V_k$ is adjacent to $V_j$ if and only if $V_j$ is adjacent to $V_k$ (because $[V_j, V_k] = [V_k, V_j]$ — the graph $\mathcal{G}$ is not directed).

Fix $V_i \in \mathcal{V}$ ($i \in \langle n \rangle$). Suppose that the vertices adjacent to $V_i$ are $V_{w_1}, V_{w_2}, ..., V_{w_{s(i)}}$. Consider the subgraph

$$\mathcal{G}(V_i) = (\mathcal{V}(V_i), \mathcal{E}(V_i))$$

of $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V}(V_i) = \left\{V_i, V_{w_1}, V_{w_2}, ..., V_{w_{s(i)}}\right\}$ (the vertex set) and $\mathcal{E}(V_i) = \left\{[V_i, V_{w_1}], [V_i, V_{w_2}], ..., \left[V_i, V_{w_{s(i)}}\right]\right\}$ (the edge set). Obviously, $\mathcal{G}(V_i)$ is a star graph (a star subgraph of $\mathcal{G}$); $V_i$ is its internal vertex (node). We call $\mathcal{G}(V_i)$ the $V_i$*-star subgraph* (*of* $\mathcal{G}$). Set

$$H_{\mathcal{G}(V_i)}(x) = \sum_{k \in \langle s(i) \rangle} \mathbf{1}[x_i \neq x_{w_k}], \ \forall x \in \langle\langle h \rangle\rangle^n$$

($x$ is a configuration of the graph $\mathcal{G}$; for $\mathbf{1}[x_i \neq x_{w_k}]$, see the definition of $H$). We call $H_{\mathcal{G}(V_i)}$ the *energy* or *Hamiltonian of* ($V_i$*-star subgraph*) $\mathcal{G}(V_i)$. $H_{\mathcal{G}(V_i)}(x)$ is the *energy* (or *Hamiltonian*) *of configuration* $x$ *on* $\mathcal{G}(V_i)$.

The next result is a generalization of Theorem 2.4 in [16].

**Theorem 2.2.** *Consider the above graph* $\mathcal{G}$. *Let* $I$ *be an independent set of* $\mathcal{G}$ ($I$ *does not contain isolated vertices*). *Suppose that* $I = \left\{V_{i_1}, V_{i_2}, ..., V_{i_p}\right\}$, *where* $p \in \langle n-1 \rangle$, $i_1$, $i_2$, ..., $i_p \in \langle n \rangle$, $i_u \neq i_v$, $\forall u, v \in \langle p \rangle$, $u \neq v$. *Then*

$$H\left(x^{(a|i_t)}\right) - H\left(x^{(b|i_t)}\right) = c_{(a,b), i_t, x_{w_1}, x_{w_2}, ..., x_{w_{s(i_t)}}},$$

*$\forall x \in \langle\langle h\rangle\rangle^n$, $\forall a, b \in \langle\langle h\rangle\rangle$, $\forall t \in \langle p\rangle$, where, $\forall x \in \langle\langle h\rangle\rangle^n$, $\forall a, b \in \langle\langle h\rangle\rangle$, $\forall t \in \langle p\rangle$, $c_{(a,b),i_t,x_{w_1},x_{w_2},\dots,x_{w_{s(i_t)}}}$ is a quantity which depends on $(a,b)$, $i_t$ (equivalently, $V_{i_t}$), and $x_{w_1}$, $x_{w_2}$, ..., $x_{w_{s(i_t)}}$, $x_{w_1}$, $x_{w_2}$, ..., $x_{w_{s(i_t)}}$ being the colors of vertices $V_{w_1}$, $V_{w_2}$, ..., $V_{w_{s(i_t)}}$, respectively, $V_{w_1}$, $V_{w_2}$, ..., $V_{w_{s(i_t)}}$ being the vertices adjacent to $V_{i_t}$ — $c_{(a,b),i_t,x_{w_1},x_{w_2},\dots,x_{w_{s(i_t)}}}$ is a constant when $a, b, i_t, x_{w_1}, x_{w_2}, \dots, x_{w_{s(i_t)}}$ are fixed. The difference $c_{(a,b),i_t,x_{w_1},x_{w_2},\dots,x_{w_{s(i_t)}}}$, $\forall x \in \langle\langle h\rangle\rangle^n$, $\forall a, b \in \langle\langle h\rangle\rangle$, $\forall t \in \langle p\rangle$, can be computed using the formula — a simple formula —*

$$c_{(a,b),i_t,x_{w_1},x_{w_2},\dots,x_{w_{s(i_t)}}} = H_{\mathcal{G}(V_{i_t})}\left(x^{(a|i_t)}\right) - H_{\mathcal{G}(V_{i_t})}\left(x^{(b|i_t)}\right),$$

*$\forall x \in \langle\langle h\rangle\rangle^n$, $\forall a, b \in \langle\langle h\rangle\rangle$, $\forall t \in \langle p\rangle$.*

*Proof.* Let $x \in \langle\langle h\rangle\rangle^n$. Let $a, b \in \langle\langle h\rangle\rangle$. Let $t \in \langle p\rangle$. If $[V_i, V_j] \in \mathcal{E}$ ($i, j \in \langle n\rangle$, $i \neq j$) and $V_i$, $V_j \neq V_{i_t}$, then

$$\mathbf{1}\left[x_i^{(a|i_t)} \neq x_j^{(a|i_t)}\right] = \mathbf{1}\left[x_i^{(b|i_t)} \neq x_j^{(b|i_t)}\right]$$

(see the definitions of $\mathbf{1}[x_i \neq x_j]$ and $x^{(a|k)}$), so,

$$\mathbf{1}\left[x_i^{(a|i_t)} \neq x_j^{(a|i_t)}\right] - \mathbf{1}\left[x_i^{(b|i_t)} \neq x_j^{(b|i_t)}\right] = 0.$$

Now, it is easy, it is obvious — for the difference $H\left(x^{(a|i_t)}\right) - H\left(x^{(b|i_t)}\right)$, use the definitions of $H$ and $x^{(a|k)}$ and the previous equation while, for the formula for $c_{(a,b),i_t,x_{w_1},x_{w_2},\dots,x_{w_{s(i_t)}}}$, use the definitions of $H_{\mathcal{G}(V_i)}$ and $x^{(a|k)}$ and the previous equation. □

**Remark 2.1.** (a) From Theorem 2.2, we have

$$c_{(a,b),i_t,x_{w_1},x_{w_2},\dots,x_{w_{s(i_t)}}} + c_{(b,a),i_t,x_{w_1},x_{w_2},\dots,x_{w_{s(i_t)}}} = 0,$$

$\forall a, b \in \langle\langle h\rangle\rangle$, $\forall t \in \langle p\rangle$, $\forall x_{w_1}, x_{w_2}, \dots, x_{w_{s(i_t)}} \in \langle\langle h\rangle\rangle$.
(b) Replacing "$a, b \in \langle\langle h\rangle\rangle$" with "$a, b \in \langle\langle h\rangle\rangle$, $a \leq b$" (or with "$a, b \in \langle\langle h\rangle\rangle$, $a \geq b$") and "$c_{(a,b),i_t,x_{w_1},x_{w_2},\dots,x_{w_{s(i_t)}}}$" with "$c_{a,b,i_t,x_{w_1},x_{w_2},\dots,x_{w_{s(i_t)}}}$" in Theorem 2.2, we obtain a version of this theorem. The reader, if he/she wishes, can use this version instead of Theorem 2.2.

Below we generalize some things from [18, Section 5].
Set

$$U_{(x_{i_1},x_{i_2},\dots,x_{i_l})} = \{(y_1, y_2, \dots, y_n) | \, (y_1, y_2, \dots, y_n) \in \langle\langle h\rangle\rangle^n \text{ and } y_{i_m} = x_{i_m}, \ \forall m \in \langle l\rangle\},$$

$\forall l \in \langle n\rangle$, $\forall i_1, i_2, \dots, i_l \in \langle n\rangle$, $i_u \neq i_v$, $\forall u, v \in \langle l\rangle$, $u \neq v$, $\forall x_{i_1}, x_{i_2}, \dots, x_{i_l} \in \langle\langle h\rangle\rangle$, and, more generally,

$$U_{(x_{i_1}\oplus k,x_{i_2}\oplus k,\dots,x_{i_l}\oplus k)} = \\ = \{(y_1, y_2, \dots, y_n) \mid \ (y_1, y_2, \dots, y_n) \in \langle\langle h\rangle\rangle^n \text{ and } y_{i_m} = x_{i_m} \oplus k, \ \forall m \in \langle l\rangle\},$$

$\forall l \in \langle n\rangle$, $\forall i_1, i_2, \dots, i_l \in \langle n\rangle$, $i_u \neq i_v$, $\forall u, v \in \langle l\rangle$, $u \neq v$, $\forall x_{i_1}, x_{i_2}, \dots, x_{i_l}, k \in \langle\langle h\rangle\rangle$.
Set

$$S_{(x_{i_1},x_{i_2},\dots,x_{i_l})} = \bigcup_{k \in \langle\langle h\rangle\rangle} U_{(x_{i_1}\oplus k,x_{i_2}\oplus k,\dots,x_{i_l}\oplus k)},$$

$\forall l \in \langle n \rangle$, $\forall i_1, i_2, ..., i_l \in \langle n \rangle$, $i_u \neq i_v$, $\forall u, v \in \langle l \rangle$, $u \neq v$, $\forall \left(x_{i_1}, x_{i_2}, ..., x_{i_l}\right) \in \{0\} \times \langle\langle h \rangle\rangle^{l-1}$ (warning! $x_{i_1} \in \{0\}$ only — $x_{i_2}, ..., x_{i_l} \in \langle\langle h \rangle\rangle$). We will construct an order relation on $S_{\left(x_{i_1}, x_{i_2}, ..., x_{i_l}\right)}$, $\forall l \in \langle n \rangle$, $\forall i_1, i_2, ..., i_l \in \langle n \rangle$, $i_u \neq i_v$, $\forall u, v \in \langle l \rangle$, $u \neq v$, $\forall \left(x_{i_1}, x_{i_2}, ..., x_{i_l}\right) \in \{0\} \times \langle\langle h \rangle\rangle^{l-1}$. To make this, we need the next result, a generalization of Theorem 5.1 from [18].

**Theorem 2.3.** *We have*

$$U_{\left(x_{i_1} \oplus k, x_{i_2} \oplus k, ..., x_{i_l} \oplus k\right)} = U_{\left(x_{i_1}, x_{i_2}, ..., x_{i_l}\right)} \oplus (k, k, ..., k),$$

$\forall l \in \langle n \rangle$, $\forall i_1, i_2, ..., i_l \in \langle n \rangle$, $i_u \neq i_v$, $\forall u, v \in \langle l \rangle$, $u \neq v$, $\forall x_{i_1}, x_{i_2}, ..., x_{i_l}, k \in \langle\langle h \rangle\rangle$, *where*

$$\begin{aligned} & U_{\left(x_{i_1}, x_{i_2}, ..., x_{i_l}\right)} \oplus (k, k, ..., k) = \\ & \quad = \left\{ (y_1, y_2, ..., y_n) \oplus (k, k, ..., k) \;\middle|\; (y_1, y_2, ..., y_n) \in U_{\left(x_{i_1}, x_{i_2}, ..., x_{i_l}\right)} \right\} = \\ & \quad = \left\{ (y_1 \oplus k, y_2 \oplus k, ..., y_n \oplus k) \;\middle|\; (y_1, y_2, ..., y_n) \in U_{\left(x_{i_1}, x_{i_2}, ..., x_{i_l}\right)} \right\}, \end{aligned}$$

$\forall l \in \langle n \rangle$, $\forall i_1, i_2, ..., i_l \in \langle n \rangle$, $i_u \neq i_v$, $\forall u, v \in \langle l \rangle$, $u \neq v$, $\forall x_{i_1}, x_{i_2}, ..., x_{i_l}, k \in \langle\langle h \rangle\rangle$ *(the vector* $(k, k, ..., k)$ *has dimension* $n$*).*

*Proof.* Let $l \in \langle n \rangle$. Let $i_1, i_2, ..., i_l \in \langle n \rangle$, $i_u \neq i_v$, $\forall u, v \in \langle l \rangle$, $u \neq v$. Let $x_{i_1}, x_{i_2}, ..., x_{i_l}, k \in \langle\langle h \rangle\rangle$.

"$\subseteq$" Let $z = (z_1, z_2, ..., z_n) \in U_{\left(x_{i_1} \oplus k, x_{i_2} \oplus k, ..., x_{i_l} \oplus k\right)}$. It follows that $z_{i_1} = x_{i_1} \oplus k$, $z_{i_2} = x_{i_2} \oplus k$, ..., $z_{i_l} = x_{i_l} \oplus k$. Let $t \in \langle n \rangle$, $t \neq i_1, i_2, ..., i_l$. We have

$$z_t = \begin{cases} (h + 1 + z_t - k) \oplus k & \text{if } z_t < k, \\ (z_t - k) \oplus k & \text{if } z_t \geq k \end{cases}$$

$(0 \leq z_t, k \leq h)$. Further, we have $h + 1 + z_t - k \in \langle\langle h \rangle\rangle$ if $z_t < k$ and $z_t - k \in \langle\langle h \rangle\rangle$ if $z_t \geq k$. We conclude that $z \in U_{\left(x_{i_1}, x_{i_2}, ..., x_{i_l}\right)} \oplus (k, k, ..., k)$.

"$\supseteq$" Let $z = (z_1, z_2, ..., z_n) \in U_{\left(x_{i_1}, x_{i_2}, ..., x_{i_l}\right)} \oplus (k, k, ..., k)$. We have $z_{i_1} = x_{i_1} \oplus k$, $z_{i_2} = x_{i_2} \oplus k$, ..., $z_{i_l} = x_{i_l} \oplus k$. Therefore, $z \in U_{\left(x_{i_1} \oplus k, x_{i_2} \oplus k, ..., x_{i_l} \oplus k\right)}$. □

Consider $U_{\left(x_{i_1}, x_{i_2}, ..., x_{i_l}\right)}$ ($x_{i_1}, x_{i_2}, ..., x_{i_l} \in \langle\langle h \rangle\rangle$, ...) equipped with the lexicographic order, $\overset{lex}{\leq}$. Let $k \in \langle h \rangle = \langle\langle h \rangle\rangle - \{0\}$. Consider $U_{\left(x_{i_1} \oplus k, x_{i_2} \oplus k, ..., x_{i_l} \oplus k\right)}$ equipped with the order relation $\overset{k}{\leq}$ defined as follows (see the formula for $U_{\left(x_{i_1} \oplus k, x_{i_2} \oplus k, ..., x_{i_l} \oplus k\right)}$ from Theorem 2.3):

$$(a_1 \oplus k, a_2 \oplus k, ..., a_n \oplus k) \overset{k}{\leq} (b_1 \oplus k, b_2 \oplus k, ..., b_n \oplus k)$$

if

$$(a_1, a_2, ..., a_n) \overset{lex}{\leq} (b_1, b_2, ..., b_n),$$

where $(a_1, a_2, ..., a_n), (b_1, b_2, ..., b_n) \in U_{\left(x_{i_1}, x_{i_2}, ..., x_{i_l}\right)}$.

Set $\overset{0}{\leq} = \overset{lex}{\leq}$.

Consider $S_{\left(x_{i_1},x_{i_2},...,x_{i_l}\right)}$ $(x_{i_1} \in \{0\}, x_{i_2}, ..., x_{i_l} \in \langle\langle h\rangle\rangle, ...)$ equipped with the order relation $\leqslant$ defined as follows (see the definition of $S_{\left(x_{i_1},x_{i_2},...,x_{i_l}\right)}$ again):

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

if

$$(u_1, u_2, ..., u_n) \in U_{\left(x_{i_1}\oplus k_1, x_{i_2}\oplus k_1,...,x_{i_l}\oplus k_1\right)} \text{ and}$$

$$(v_1, v_2, ..., v_n) \in U_{\left(x_{i_1}\oplus k_2, x_{i_2}\oplus k_2,...,x_{i_l}\oplus k_2\right)} \text{ for some } k_1, k_2 \in \langle\langle h\rangle\rangle, \ k_1 < k_2,$$

or if

$$(u_1, u_2, ..., u_n), (v_1, v_2, ..., v_n) \in U_{\left(x_{i_1}\oplus k, x_{i_2}\oplus k,...,x_{i_l}\oplus k\right)} \text{ and}$$

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

where $(u_1, u_2, ..., u_n), (v_1, v_2, ..., v_n) \in S_{\left(x_{i_1},x_{i_2},...,x_{i_l}\right)}$.

Set

$$\begin{aligned} &U_{\left(x_{s_1}\oplus k, x_{s_2}\oplus k,...,x_{s_l}\oplus k, x_{t_1}=k_1, x_{t_2}=k_2,...,x_{t_m}=k_m\right)} = \\ &\quad = \left\{(y_1, y_2, ..., y_n) \mid (y_1, y_2, ..., y_n) \in \langle\langle h\rangle\rangle^n, \ y_{s_g} = x_{s_g} \oplus k, \ \forall g \in \langle l\rangle, \text{ and} \right. \\ &\qquad\qquad \left. y_{t_i} = x_{t_i} = k_i, \forall i \in \langle m\rangle\right\}, \end{aligned}$$

$\forall l, m \in \langle n\rangle$, $l + m \leq n$, $\forall s_1, s_2, ..., s_l, t_1, t_2, ..., t_m \in \langle n\rangle$, $s_u \neq s_v$, $\forall u, v \in \langle l\rangle$, $u \neq v$, $t_w \neq t_z$, $\forall w, z \in \langle m\rangle$, $w \neq z$, $\{s_1, s_2, ..., s_l\} \cap \{t_1, t_2, ..., t_m\} = \emptyset$, $\forall x_{s_1}, x_{s_2}, ..., x_{s_l}, k, k_1, k_2, ..., k_m \in \langle\langle h\rangle\rangle$.

Consider $I = \left\{V_{i_1}, V_{i_2}, ..., V_{i_p}\right\} \subseteq \mathcal{V}$, an independent set of $\mathcal{G}$. Consider $I^c = \left\{V_{j_1}, V_{j_2}, ..., V_{j_q}\right\}$, the complement of $I$ $(q \geq 1,\ p + q = n = |\mathcal{V}|)$. Fix $x_{j_1}$, $x_{j_2}$, ..., $x_{j_q}$; $x_{j_1}, x_{j_2}, ..., x_{j_q} \in \langle\langle h\rangle\rangle$, $x_{j_1}, x_{j_2}, ..., x_{j_q}$ are the colors of $V_{j_1}, V_{j_2}, ..., V_{j_q}$ in the configurations from $U_{\left(x_{j_1},x_{j_2},...,x_{j_q}\right)} \subset S_{\left(x_{j_1},x_{j_2},...,x_{j_q}\right)}$, respectively ($x_{j_1} \oplus k$, $x_{j_2} \oplus k$, ..., $x_{j_q} \oplus k$ are the colors of $V_{j_1}, V_{j_2}, ..., V_{j_q}$ in the configurations from $U_{\left(x_{j_1}\oplus k, x_{j_2}\oplus k,...,x_{j_q}\oplus k\right)} \subset S_{\left(x_{j_1},x_{j_2},...,x_{j_q}\right)}$, respectively, where $k \in \langle\langle h\rangle\rangle$). Define the partitions of $S_{\left(x_{j_1},x_{j_2},...,x_{j_q}\right)} \subseteq \langle\langle h\rangle\rangle^n$:

$$\Delta_1^{\left(x_{j_1},x_{j_2},...,x_{j_q}\right)} = \left(S_{\left(x_{j_1},x_{j_2},...,x_{j_q}\right)}\right)$$

($\Delta_1^{\left(x_{j_1},x_{j_2},...,x_{j_q}\right)}$ is the improper partition of $S_{\left(x_{j_1},x_{j_2},...,x_{j_q}\right)}$; the elements (configurations) of $S_{\left(x_{j_1},x_{j_2},...,x_{j_q}\right)}$ are in the order determined by $\leqslant$),

$$\begin{aligned} \Delta_2^{\left(x_{j_1},x_{j_2},...,x_{j_q}\right)} &= \left(U_{\left(x_{j_1}\oplus k, x_{j_2}\oplus k,...,x_{j_q}\oplus k\right)}\right)_{k\in\langle\langle h\rangle\rangle} = \\ &= \left(U_{\left(x_{j_1},x_{j_2},...,x_{j_q}\right)}, U_{\left(x_{j_1}\oplus 1, x_{j_2}\oplus 1,...,x_{j_q}\oplus 1\right)}, ..., U_{\left(x_{j_1}\oplus h, x_{j_2}\oplus h,...,x_{j_q}\oplus h\right)}\right) \end{aligned}$$

($U_{\left(x_{j_1},x_{j_2},...,x_{j_q}\right)}$ contains the first $\left|U_{\left(x_{j_1},x_{j_2},...,x_{j_q}\right)}\right|$ elements of $S_{\left(x_{j_1},x_{j_2},...,x_{j_q}\right)}$, $U_{\left(x_{j_1}\oplus 1, x_{j_2}\oplus 1,...,x_{j_q}\oplus 1\right)}$ contains the second $\left|U_{\left(x_{j_1}\oplus 1, x_{j_2}\oplus 1,...,x_{j_q}\oplus 1\right)}\right|$ elements of

$S_{\left(x_{j_1},x_{j_2},\ldots,x_{j_q}\right)}$, ..., $U_{\left(x_{j_1}\oplus h,x_{j_2}\oplus h,\ldots,x_{j_q}\oplus h\right)}$ contains the last $\left|U_{\left(x_{j_1}\oplus h,x_{j_2}\oplus h,\ldots,x_{j_q}\oplus h\right)}\right|$ elements of $S_{\left(x_{j_1},x_{j_2},\ldots,x_{j_q}\right)}$(see the definition of $\preccurlyeq$)),

$$\Delta_3^{\left(x_{j_1},x_{j_2},\ldots,x_{j_q}\right)} = \left(U_{\left(x_{j_1}\oplus k,x_{j_2}\oplus k,\ldots,x_{j_q}\oplus k,x_{i_1}=k_1\right)}\right)_{k,k_1\in\langle\langle h\rangle\rangle} =$$
$$= \Big(U_{\left(x_{j_1},x_{j_2},\ldots,x_{j_q},x_{i_1}=0\right)}, U_{\left(x_{j_1},x_{j_2},\ldots,x_{j_q},x_{i_1}=1\right)}, \ldots, U_{\left(x_{j_1},x_{j_2},\ldots,x_{j_q},x_{i_1}=h\right)},$$
$$U_{\left(x_{j_1}\oplus 1,x_{j_2}\oplus 1,\ldots,x_{j_q}\oplus 1,x_{i_1}=0\oplus 1\right)}, U_{\left(x_{j_1}\oplus 1,x_{j_2}\oplus 1,\ldots,x_{j_q}\oplus 1,x_{i_1}=1\oplus 1\right)}, \ldots,$$
$$U_{\left(x_{j_1}\oplus 1,x_{j_2}\oplus 1,\ldots,x_{j_q}\oplus 1,x_{i_1}=h\oplus 1\right)}, \ldots, \ldots, U_{\left(x_{j_1}\oplus h,x_{j_2}\oplus h,\ldots,x_{j_q}\oplus h,x_{i_1}=0\oplus h\right)},$$
$$U_{\left(x_{j_1}\oplus h,x_{j_2}\oplus h,\ldots,x_{j_q}\oplus h,x_{i_1}=1\oplus h\right)}, \ldots, U_{\left(x_{j_1}\oplus h,x_{j_2}\oplus h,\ldots,x_{j_q}\oplus h,x_{i_1}=h\oplus h\right)}\Big)$$

$(U_{\left(x_{j_1},x_{j_2},\ldots,x_{j_q},x_{i_1}=k_1\right)} \subset U_{\left(x_{j_1},x_{j_2},\ldots,x_{j_q}\right)}$, $\forall k_1 \in \langle\langle h\rangle\rangle$, and

$$U_{\left(x_{j_1},x_{j_2},\ldots,x_{j_q}\right)} = \bigcup_{k_1\in\langle\langle h\rangle\rangle} U_{\left(x_{j_1},x_{j_2},\ldots,x_{j_q},x_{i_1}=k_1\right)}, \quad \text{etc.;}$$

$U_{\left(x_{j_1},x_{j_2},\ldots,x_{j_q},x_{i_1}=0\right)}$ contains the first $\left|U_{\left(x_{j_1},x_{j_2},\ldots,x_{j_q},x_{i_1}=0\right)}\right|$ elements of $S_{\left(x_{j_1},x_{j_2},\ldots,x_{j_q}\right)}$ (equivalently (here), of $U_{\left(x_{j_1},x_{j_2},\ldots,x_{j_q}\right)}$), ... (see the definitions of $\overset{k}{\leq}$ and $\preccurlyeq$)),

$$\Delta_4^{\left(x_{j_1},x_{j_2},\ldots,x_{j_q}\right)} = \left(U_{\left(x_{j_1}\oplus k,x_{j_2}\oplus k,\ldots,x_{j_q}\oplus k,x_{i_1}=k_1,x_{i_2}=k_2\right)}\right)_{k,k_1,k_2\in\langle\langle h\rangle\rangle} =$$
$$= \Big(U_{\left(x_{j_1},x_{j_2},\ldots,x_{j_q},x_{i_1}=0,x_{i_2}=0\right)}, U_{\left(x_{j_1},x_{j_2},\ldots,x_{j_q},x_{i_1}=0,x_{i_2}=1\right)}, \ldots, \ldots,$$
$$U_{\left(x_{j_1}\oplus h,x_{j_2}\oplus h,\ldots,x_{j_q}\oplus h,x_{i_1}=h\oplus h,x_{i_2}=h\oplus h\right)}\Big),$$
$$\vdots$$
$$\Delta_{p+2}^{\left(x_{j_1},x_{j_2},\ldots,x_{j_q}\right)} = \left(U_{\left(x_{j_1}\oplus k,x_{j_2}\oplus k,\ldots,x_{j_q}\oplus k,x_{i_1}=k_1,x_{i_2}=k_2,\ldots,x_{i_p}=k_p\right)}\right)_{k,k_1,k_2,\ldots,k_p\in\langle\langle h\rangle\rangle}.$$

Obviously,

$$\Delta_{p+2}^{\left(x_{j_1},x_{j_2},\ldots,x_{j_q}\right)} = (\{x\})_{x\in S_{\left(x_{j_1},x_{j_2},\ldots,x_{j_q}\right)}}$$

and

$$\Delta_1^{\left(x_{j_1},x_{j_2},\ldots,x_{j_q}\right)} \succ \Delta_2^{\left(x_{j_1},x_{j_2},\ldots,x_{j_q}\right)} \succ \ldots \succ \Delta_{p+2}^{\left(x_{j_1},x_{j_2},\ldots,x_{j_q}\right)}.$$

The next two results are about two basic properties of the Potts model, about the structure of Potts model.

**Theorem 2.4.** *Under the above conditions the Potts model on the graph $\mathcal{G}$ is a wavy probability distribution on $S_{\left(x_{j_1},x_{j_2},\ldots,x_{j_q}\right)}$ with respect to the order relation $\preccurlyeq$ and partitions*

$$\Delta_1^{\left(x_{j_1},x_{j_2},\ldots,x_{j_q}\right)},\ \Delta_2^{\left(x_{j_1},x_{j_2},\ldots,x_{j_q}\right)},\ \ldots,\ \Delta_{p+2}^{\left(x_{j_1},x_{j_2},\ldots,x_{j_q}\right)}.$$

*Proof.* We must show that $\pi|_{S_{\left(x_{j_1},x_{j_2},\ldots,x_{j_q}\right)}}$ is a wavy probability distribution with respect to the order relation $\preccurlyeq$ and partitions $\Delta_1^{\left(x_{j_1},x_{j_2},\ldots,x_{j_q}\right)}$, $\Delta_2^{\left(x_{j_1},x_{j_2},\ldots,x_{j_q}\right)}$, ..., $\Delta_{p+2}^{\left(x_{j_1},x_{j_2},\ldots,x_{j_q}\right)}$. (For $\pi|_{S_{\left(x_{j_1},x_{j_2},\ldots,x_{j_q}\right)}}$, see Definition 1.7.)

Recall that (see above) $\Delta_l^{(x_{j_1},x_{j_2},\dots,x_{j_q})} \succ \Delta_{l+1}^{(x_{j_1},x_{j_2},\dots,x_{j_q})}$, $\forall l \in \langle p+1 \rangle$. The conditions (c1) and (c2) also hold. (See the definition of wavy probability distribution again.)

Consider the partitions $\Delta_1^{(x_{j_1},x_{j_2},\dots,x_{j_q})}$ and $\Delta_2^{(x_{j_1},x_{j_2},\dots,x_{j_q})}$. The first set of $\Delta_2^{(x_{j_1},x_{j_2},\dots,x_{j_q})}$ is $U_{(x_{j_1},x_{j_2},\dots,x_{j_q})}$, the second one is $U_{(x_{j_1}\oplus 1,x_{j_2}\oplus 1,\dots,x_{j_q}\oplus 1)}$, ..., the last one is $U_{(x_{j_1}\oplus h,x_{j_2}\oplus h,\dots,x_{j_q}\oplus h)}$. Let $k \in \langle\langle h \rangle\rangle$, $k \neq 0$. Let $z \in U_{(x_{j_1}\oplus k,x_{j_2}\oplus k,\dots,x_{j_q}\oplus k)}$. By Theorem 2.3, $\exists y \in U_{(x_{j_1},x_{j_2},\dots,x_{j_q})}$ such that $z = y \oplus (k,k,\dots,k)$. By Theorem 2.1, $H(z) = H(y)$. Suppose that $z$ is the $s$th element of $U_{(x_{j_1}\oplus k,x_{j_2}\oplus k,\dots,x_{j_q}\oplus k)}$, $1 \leq s \leq \left| U_{(x_{j_1}\oplus k,x_{j_2}\oplus k,\dots,x_{j_q}\oplus k)} \right|$ (see the definition of wavy probability distribution again). It follows that $y$ is the $s$th element of $U_{(x_{j_1},x_{j_2},\dots,x_{j_q})}$ (see the definitions of $\overset{k}{\leq}$ and $\preccurlyeq$ again). Finally, we obtain

$$\left(\pi|_{S_{(x_{j_1},x_{j_2},\dots,x_{j_q})}}\right)_z = \frac{\pi_z}{P\left(S_{(x_{j_1},x_{j_2},\dots,x_{j_q})}\right)} = \frac{\theta^{H(z)}}{ZP\left(S_{(x_{j_1},x_{j_2},\dots,x_{j_q})}\right)} =$$

$$= \frac{\theta^{H(y)}}{ZP\left(S_{(x_{j_1},x_{j_2},\dots,x_{j_q})}\right)} = \frac{\pi_y}{P\left(S_{(x_{j_1},x_{j_2},\dots,x_{j_q})}\right)} = \left(\pi|_{S_{(x_{j_1},x_{j_2},\dots,x_{j_q})}}\right)_y$$

(the proportionality factor is 1).

Now, we consider the partitions $\Delta_{l+1}^{(x_{j_1},x_{j_2},\dots,x_{j_q})}$ and $\Delta_{l+2}^{(x_{j_1},x_{j_2},\dots,x_{j_q})}$, where $1 \leq l \leq p$. Let $K \in \Delta_{l+1}^{(x_{j_1},x_{j_2},\dots,x_{j_q})}$. We have

$$K = \begin{cases} U_{(x_{j_1}\oplus k,x_{j_2}\oplus k,\dots,x_{j_q}\oplus k)} \\ \quad \text{for some } k \in \langle\langle h \rangle\rangle \text{ if } l = 1, \\ U_{(x_{j_1}\oplus k,x_{j_2}\oplus k,\dots,x_{j_q}\oplus k,x_{i_1}=k_1,x_{i_2}=k_2,\dots,x_{i_{l-1}}=k_{l-1})} \\ \quad \text{for some } k,k_1,k_2,\dots,k_{l-1} \in \langle\langle h \rangle\rangle \text{ if } 2 \leq l \leq p. \end{cases}$$

Using the order relation $\preccurlyeq$, the first subset of $K$ belonging to $\Delta_{l+2}^{(x_{j_1},x_{j_2},\dots,x_{j_q})}$ is ($x_{i_1} = k_1$, $x_{i_2} = k_2$, ..., $x_{i_{l-1}} = k_{l-1}$ vanish when $l = 1$)

$$U_{(x_{j_1}\oplus k,x_{j_2}\oplus k,\dots,x_{j_q}\oplus k,x_{i_1}=k_1,x_{i_2}=k_2,\dots,x_{i_{l-1}}=k_{l-1},x_{i_l}=0\oplus k)},$$

the second one is

$$U_{(x_{j_1}\oplus k,x_{j_2}\oplus k,\dots,x_{j_q}\oplus k,x_{i_1}=k_1,x_{i_2}=k_2,\dots,x_{i_{l-1}}=k_{l-1},x_{i_l}=1\oplus k)}, \ \dots,$$

the last one is

$$U_{(x_{j_1}\oplus k,x_{j_2}\oplus k,\dots,x_{j_q}\oplus k,x_{i_1}=k_1,x_{i_2}=k_2,\dots,x_{i_{l-1}}=k_{l-1},x_{i_l}=h\oplus k)}.$$

Let $g \in \langle h \rangle$. Let $v \in U_{(x_{j_1}\oplus k,x_{j_2}\oplus k,\dots,x_{j_q}\oplus k,x_{i_1}=k_1,x_{i_2}=k_2,\dots,x_{i_{l-1}}=k_{l-1},x_{i_l}=g\oplus k)}$. Let $u \in U_{(x_{j_1}\oplus k,x_{j_2}\oplus k,\dots,x_{j_q}\oplus k,x_{i_1}=k_1,x_{i_2}=k_2,\dots,x_{i_{l-1}}=k_{l-1},x_{i_l}=0\oplus k)}$, where

$$u = (u_i)_{i\in\langle n \rangle}, \ u_i = \begin{cases} v_i & \text{if } i \neq i_l, \\ 0 \oplus k & \text{if } i = i_l, \end{cases}$$

$\forall i \in \langle n \rangle$ . Suppose that $v$ is the $f$th element of

$$U_{\left(x_{j_1}\oplus k, x_{j_2}\oplus k, \ldots, x_{j_q}\oplus k, x_{i_1}=k_1, x_{i_2}=k_2, \ldots, x_{i_{l-1}}=k_{l-1}, x_{i_l}=g\oplus k\right)}.$$

It follows that $u$ is the $f$th element of

$$U_{\left(x_{j_1}\oplus k, x_{j_2}\oplus k, \ldots, x_{j_q}\oplus k, x_{i_1}=k_1, x_{i_2}=k_2, \ldots, x_{i_{l-1}}=k_{l-1}, x_{i_l}=0\oplus k\right)}$$

because the first element of

$$U_{\left(x_{j_1}\oplus k, x_{j_2}\oplus k, \ldots, x_{j_q}\oplus k, x_{i_1}=k_1, x_{i_2}=k_2, \ldots, x_{i_{l-1}}=k_{l-1}, x_{i_l}=0\oplus k\right)}$$

is $d = (d_1, d_2, ..., d_n)$ , where

$$d_{j_e} = x_{j_e} \oplus k, \ \forall e \in \langle q \rangle ,$$

$$d_{i_t} = x_{i_t} = k_{i_t}, \ \forall t \in \langle l-1 \rangle , \text{ if } 2 \leq l \leq p,$$

$$d_r = 0 \oplus k, \ \forall r \in \langle n \rangle - F,$$

where

$$F = \begin{cases} \{j_1, j_2, ..., j_q\} & \text{if } l = 1, \\ \{j_1, j_2, ..., j_q\} \cup \{i_1, i_2, ..., i_{l-1}\} & \text{if } 2 \leq l \leq p, \end{cases}$$

while the first element of

$$U_{\left(x_{j_1}\oplus k, x_{j_2}\oplus k, \ldots, x_{j_q}\oplus k, x_{i_1}=k_1, x_{i_2}=k_2, \ldots, x_{i_{l-1}}=k_{l-1}, x_{i_l}=g\oplus k\right)}$$

is $\overline{d} = \left(\overline{d}_1, \overline{d}_2, ..., \overline{d}_n\right)$ , where

$$\overline{d}_m = \begin{cases} d_m & \text{if } m \neq i_l, \\ g \oplus k & \text{if } m = i_l, \end{cases}$$

$\forall m \in \langle n \rangle$ (see the definitions of $\preccurlyeq$ and $\overset{k}{\leq}$ again), etc. Finally, using Theorem 2.2, we have

$$\left(\pi|_{S_{\left(x_{j_1}, x_{j_2}, \ldots, x_{j_q}\right)}}\right)_v = \frac{\pi_v}{P\left(S_{\left(x_{j_1}, x_{j_2}, \ldots, x_{j_q}\right)}\right)} = \frac{\theta^{H(v)}}{ZP\left(S_{\left(x_{j_1}, x_{j_2}, \ldots, x_{j_q}\right)}\right)} =$$

$$= \frac{\theta^{H(u)+c_{(g\oplus k, 0\oplus k), i_l, x_{w_1}, x_{w_2}, \ldots, x_{w_{s(i_l)}}}}}{ZP\left(S_{\left(x_{j_1}, x_{j_2}, \ldots, x_{j_q}\right)}\right)} =$$

$$= \theta^{c_{(g\oplus k, 0\oplus k), i_l, x_{w_1}, x_{w_2}, \ldots, x_{w_{s(i_l)}}}} \cdot \frac{\theta^{H(u)}}{ZP\left(S_{\left(x_{j_1}, x_{j_2}, \ldots, x_{j_q}\right)}\right)} =$$

$$= \theta^{c_{(g\oplus k, 0\oplus k), i_l, x_{w_1}, x_{w_2}, \ldots, x_{w_{s(i_l)}}}} \cdot \frac{\pi_u}{P\left(S_{\left(x_{j_1}, x_{j_2}, \ldots, x_{j_q}\right)}\right)} =$$

$$= \theta^{c_{(g\oplus k, 0\oplus k), i_l, x_{w_1}, x_{w_2}, \ldots, x_{w_{s(i_l)}}}} \cdot \left(\pi|_{S_{\left(x_{j_1}, x_{j_2}, \ldots, x_{j_q}\right)}}\right)_u$$

(the proportionality factor is $\theta^{c_{(g\oplus k, 0\oplus k), i_l, x_{w_1}, x_{w_2}, \ldots, x_{w_{s(i_l)}}}}$ ), where $x_{w_1}$, $x_{w_2}$, ..., $x_{w_{s(i_l)}}$ are the colors of vertices $V_{w_1}$, $V_{w_2}$, ..., $V_{w_{s(i_l)}}$, respectively, $V_{w_1}$, $V_{w_2}$, ..., $V_{w_{s(i_l)}}$ being the vertices adjacent to $V_{i_l}$. $\square$

**Theorem 2.5.** *Consider the above graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$. Consider $I = \{V_{i_1}, V_{i_2}, ..., V_{i_p}\} \subseteq \mathcal{V}$, an independent set of $\mathcal{G}$. Consider $I^c = \{V_{j_1}, V_{j_2}, ..., V_{j_q}\}$ ($I^c$ is the complement of $I$, so, $q \geq 1$, $p + q = n = |\mathcal{V}|$). Then the Potts model on the graph $\mathcal{G}$ is a $\Delta$-wavy probability distribution, where*

$$\Delta = \left(S_{\left(x_{j_1}, x_{j_2}, ..., x_{j_q}\right)}\right)_{\left(x_{j_1}, x_{j_2}, ..., x_{j_q}\right) \in \{0\} \times \langle\langle h \rangle\rangle^{q-1}},$$

*and, as a result,*

$$|\Delta| = (h+1)^{q-1}$$

*— this model is a wavy probability distribution on $S_{\left(x_{j_1}, x_{j_2}, ..., x_{j_q}\right)}$ with respect to the order relation $\leqslant$ and partitions $\Delta_1^{\left(x_{j_1}, x_{j_2}, ..., x_{j_q}\right)}$, $\Delta_2^{\left(x_{j_1}, x_{j_2}, ..., x_{j_q}\right)}$, ..., $\Delta_{p+2}^{\left(x_{j_1}, x_{j_2}, ..., x_{j_q}\right)}$ specified before Theorem 2.4, $\forall \left(x_{j_1}, x_{j_2}, ..., x_{j_q}\right) \in \{0\} \times \langle\langle h \rangle\rangle^{q-1}$.*

*Proof.* Definition 1.8 and Theorem 2.4 □

Concerning Theorem 2.5, there exists a one-to-one correspondence between independent sets and partitions. It is important for the computation of normalization constant, etc. that $|\Delta|$ be as small as possible, but not in all cases, see, *e.g.*, Remark 3.2 and the proof of Theorem 4.1. $|\Delta|$ is minimum if and only if $I$ is a maximum independent set.

The next result is about the Potts model too, and is useful for sampling (see Section 6) and for the computation of normalization constant (see Theorem 2.7).

**Theorem 2.6.** *Consider the Potts model on the (above) graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$. Consider $I = \{V_{i_1}, V_{i_2}, ..., V_{i_p}\} \subseteq \mathcal{V}$, an independent set of $\mathcal{G}$. Consider $I^c = \{V_{j_1}, V_{j_2}, ..., V_{j_q}\}$ (recall that $I^c$ is the complement of $I$, so, $q \geq 1$, $p + q = n = |\mathcal{V}|$). Then*

$$ZP\left(S_{\left(x_{j_1}, x_{j_2}, ..., x_{j_q}\right)}\right) = (h+1)\,\theta^{H\left(y^{\left(x_{j_1}, x_{j_2}, ..., x_{j_q}\right)}\right)} \times$$
$$\times \prod_{l \in \langle p \rangle} \left(1 + \sum_{w \in \langle h \rangle} \theta^{c_{(w,0), i_l, x_{w_1}, x_{w_2}, ..., x_{w_{s(i_l)}}}}\right),$$

*$\forall \left(x_{j_1}, x_{j_2}, ..., x_{j_q}\right) \in \{0\} \times \langle\langle h \rangle\rangle^{q-1}$, where $y^{\left(x_{j_1}, x_{j_2}, ..., x_{j_q}\right)}$ is the first element of $S_{\left(x_{j_1}, x_{j_2}, ..., x_{j_q}\right)}$ (equivalently, of $U_{\left(x_{j_1}, x_{j_2}, ..., x_{j_q}\right)}$), $\forall \left(x_{j_1}, x_{j_2}, ..., x_{j_q}\right) \in \{0\} \times \langle\langle h \rangle\rangle^{q-1}$,*

$$y^{\left(x_{j_1}, x_{j_2}, ..., x_{j_q}\right)} = \left(y_i^{\left(x_{j_1}, x_{j_2}, ..., x_{j_q}\right)}\right)_{i \in \langle n \rangle}, \ \forall \left(x_{j_1}, x_{j_2}, ..., x_{j_q}\right) \in \{0\} \times \langle\langle h \rangle\rangle^{q-1},$$

$$y_i^{\left(x_{j_1}, x_{j_2}, ..., x_{j_q}\right)} = \begin{cases} 0 & \text{if } i \notin \{j_1, j_2, , ..., j_q\}, \\ x_{j_k} & \text{if } i = j_k \text{ for some } k \in \langle q \rangle, \end{cases}$$

*$\forall \left(x_{j_1}, x_{j_2}, ..., x_{j_q}\right) \in \{0\} \times \langle\langle h \rangle\rangle^{q-1}$, $\forall i \in \langle n \rangle$,*

$$c_{(w,0), i_l, x_{w_1}, x_{w_2}, ..., x_{w_{s(i_l)}}} = H\left(y^{\left(x_{j_1}, x_{j_2}, ..., x_{j_q}\right)(w|i_l)}\right) - H\left(y^{\left(x_{j_1}, x_{j_2}, ..., x_{j_q}\right)(0|i_l)}\right)$$
$$= H_{\mathcal{G}(V_{i_l})}\left(y^{\left(x_{j_1}, x_{j_2}, ..., x_{j_q}\right)(w|i_l)}\right) - H_{\mathcal{G}(V_{i_l})}\left(y^{\left(x_{j_1}, x_{j_2}, ..., x_{j_q}\right)(0|i_l)}\right),$$

*$\forall \left(x_{j_1}, x_{j_2}, ..., x_{j_q}\right) \in \{0\} \times \langle\langle h\rangle\rangle^{q-1}$, $\forall w \in \langle h\rangle$, $\forall l \in \langle p\rangle$, $x_{w_1}$, $x_{w_2}$, ..., $x_{w_{s(i_l)}}$ are the colors of $V_{w_1}$, $V_{w_2}$, ..., $V_{w_{s(i_l)}}$, respectively, $V_{w_1}$, $V_{w_2}$, ..., $V_{w_{s(i_l)}}$ being the vertices adjacent to $V_{i_l}$ (see Theorem 2.2; $y^{(x_{j_1},x_{j_2},...,x_{j_q})(w|i_l)} \in U_{(x_{j_1},x_{j_2},...,x_{j_q})} \subset S_{(x_{j_1},x_{j_2},...,x_{j_q})}$ — therefore, the colors of $V_{j_1}$, $V_{j_2}$, ..., $V_{j_q}$ are $x_{j_1}$, $x_{j_2}$, ..., $x_{j_q}$, respectively —, $\forall \left(x_{j_1}, x_{j_2}, ..., x_{j_q}\right) \in \{0\} \times \langle\langle h\rangle\rangle^{q-1}$, $\forall w \in \langle\langle h\rangle\rangle = \{0\} \cup \langle h\rangle$, $\forall l \in \langle p\rangle$; obviously, $y^{(x_{j_1},x_{j_2},...,x_{j_q})(0|i_l)} = y^{(x_{j_1},x_{j_2},...,x_{j_q})}$, $\forall \left(x_{j_1}, x_{j_2}, ..., x_{j_q}\right) \in \{0\} \times \langle\langle h\rangle\rangle^{q-1}$, $\forall l \in \langle p\rangle$).*

*Proof.* Let $\left(x_{j_1}, x_{j_2}, ..., x_{j_q}\right) \in \{0\} \times \langle\langle h\rangle\rangle^{q-1}$. By Definition 1.7 and Theorem 2.4, $\pi|_{S_{(x_{j_1},x_{j_2},...,x_{j_q})}}$ is a wavy probability distribution (on $S_{(x_{j_1},x_{j_2},...,x_{j_q})}$) with respect to the order relation $\leqslant$ and partitions $\Delta_l^{(x_{j_1},x_{j_2},...,x_{j_q})}$, $l \in \langle p+2\rangle$, its normalization constant being $ZP\left(S_{(x_{j_1},x_{j_2},...,x_{j_q})}\right)$. To compute this constant, we will use Theorem 1.5. We have

$$\pi_{y^{(x_{j_1},x_{j_2},...,x_{j_q})}} = \frac{\theta^{H\left(y^{(x_{j_1},x_{j_2},...,x_{j_q})}\right)}}{Z},$$

$$b_l = b_l^{(x_{j_1},x_{j_2},...,x_{j_q})} = \frac{\left|\Delta_{l+1}^{(x_{j_1},x_{j_2},...,x_{j_q})}\right|}{\Delta_l^{(x_{j_1},x_{j_2},...,x_{j_q})}} = \frac{(h+1)^l}{(h+1)^{l-1}} = h+1,\ \forall l \in \langle p+1\rangle,$$

$$D_{1,b_l} = D_{1,b_l^{(x_{j_1},x_{j_2},...,x_{j_q})}} = \{1,2,...,h\} = \langle h\rangle,\ \forall l \in \langle p+1\rangle,$$

and (see the proof of Theorem 2.4)

$$\alpha_w^{(l,1)} = \alpha_w^{(l,1),(x_{j_1},x_{j_2},...,x_{j_q})} = \begin{cases} 1 & \text{if } l = 1, \\ \theta^{c(w,0),i_{l-1},x_{w_1},x_{w_2},...,x_{w_{s(i_{l-1})}}} & \text{if } l \in \langle p+1\rangle - \{1\}, \end{cases}$$

$\forall l \in \langle p+1\rangle$, $\forall w \in \langle h\rangle$, where $x_{w_1}$, $x_{w_2}$, ..., $x_{w_{s(i_{l-1})}}$ are the colors of (vertices) $V_{w_1}$, $V_{w_2}$, ..., $V_{w_{s(i_{l-1})}}$, respectively, $V_{w_1}$, $V_{w_2}$, ..., $V_{w_{s(i_{l-1})}}$ being the vertices adjacent to $V_{i_{l-1}}$ if $l \in \langle p+1\rangle - \{1\}$. So, by Theorem 1.5,

$$ZP\left(S_{(x_{j_1},x_{j_2},...,x_{j_q})}\right) = \theta^{H\left(y^{(x_{j_1},x_{j_2},...,x_{j_q})}\right)} \prod_{l\in\langle p+1\rangle}\left(1 + \sum_{w\in\langle h\rangle}\alpha_w^{(l,1)}\right)$$
$$= (h+1)\,\theta^{H\left(y^{(x_{j_1},x_{j_2},...,x_{j_q})}\right)} \prod_{l\in\langle p+1\rangle-\{1\}}\left(1 + \sum_{w\in\langle h\rangle}\theta^{c(w,0),i_{l-1},x_{w_1},x_{w_2},...,x_{w_{s(i_{l-1})}}}\right)$$
$$= (h+1)\,\theta^{H\left(y^{(x_{j_1},x_{j_2},...,x_{j_q})}\right)} \prod_{l\in\langle p\rangle}\left(1 + \sum_{w\in\langle h\rangle}\theta^{c(w,0),i_l,x_{w_1},x_{w_2},...,x_{w_{s(i_l)}}}\right).$$

□

The next result is another main result about the Potts model, a connection between two central notions, independent set and normalization constant.

**Theorem 2.7.** *Under the same conditions as in Theorem 2.6 we have*

$$Z=(h+1)\sum_{\left(x_{j_1},x_{j_2},\ldots,x_{j_q}\right)\in\{0\}\times\langle\langle h\rangle\rangle^{q-1}}\left[\theta^{H\left(y^{\left(x_{j_1},x_{j_2},\ldots,x_{j_q}\right)}\right)}\cdot\prod_{l\in\langle p\rangle}\left(1+\sum_{w\in\langle h\rangle}\theta^{c(w,0),i_l,x_{w_1},x_{w_2},\ldots,x_{w_{s(i_l)}}}\right)\right]$$

($x_{j_1}$, $x_{j_2}$, ..., $x_{j_q}$ *are the colors of vertices* $V_{j_1}$, $V_{j_2}$, ..., $V_{j_q}$, *respectively*).

*Proof.* Theorems 1.6, 2.5, and 2.6. □

Based on the above result, we now can give the steps we need to compute the normalization constant $Z$ for the Potts model on $\mathcal{G}$ — an arbitrary but fixed graph (this graph can be connected or not). Our interest is to obtain for $Z$ a formula (an expression) as good as possible (as simple as possible, ...) — this fact is taken into account in these steps, the last step can be performed or not.

*Step* 1 (*Graph level*). Determine an independent set of $\mathcal{G}$ as large as possible (the larger the independent set is, the smaller the numbers of terms of sum for $Z$ from Theorem 2.7 is), better, a maximal independent set of $\mathcal{G}$ as large as possible, best, a maximum independent set of $\mathcal{G}$ (the last problem is NP-hard, but in some cases it can easy be solved, see, *e.g.*, [2, Chapter 4]).

*Step* 2 (*Markov chain* (*or* Δ*-wavy probability distribution*) *level*). For the independent set of graph $\mathcal{G}$ found at Step 1, determine $Z$ using the formula from Theorem 2.7.

*Step* 3 (*Algebraic level*, if possible). Simplify, if possible, the formula (expression) for $Z$ found at Step 2 — determine, if possible, the identical products, use, if possible, algebraic identities, ...

**Example 2.1.** (A simple case.) Consider the Potts model on the complete bipartite graph $\mathcal{K}_{2,n-2}$, where $n\geq 3$. Consider that the bipartition of this graph is $(X,Y)$, where $X=\{V_1,V_2\}$, $Y=\{V_3,V_4,...,V_n\}$ ($\{V_1,V_2,...,V_n\}$ is the vertex set of this graph).

*Step* 1 (for the Potts model on $\mathcal{K}_{2,n-2}$). We take $I=Y$. $I$ is an independent set (of $\mathcal{K}_{2,n-2}$) if $n=3$, a maximum independent set if $n=4$, and the maximum independent set if $n\geq 5$ (if $n=3$, $X$ is a maximum independent set, and we can take $I=X$).

*Steps* 2-3. By Theorem 2.7 we have

$$Z=(h+1)\sum_{x_1=0,x_2\in\langle\langle h\rangle\rangle}\theta^{H\left(y^{(x_1,x_2)}\right)}\prod_{i=3}^{n}\left(1+\sum_{w\in\langle h\rangle}\theta^{c(w,0),i,x_1,x_2}\right).$$

Since

$$H\left(y^{(x_1,x_2)}\right)=\begin{cases}0 & \text{if } x_1=x_2=0,\\ n-2 & \text{if } x_1=0,\ x_2\in\langle h\rangle\,,\end{cases}$$

and

$$c_{(w,0),i,x_1,x_2} = \begin{cases} 2 & \text{if } x_1 = x_2 = 0, \\ 0 & \text{if } x_1 = 0,\ x_2 \in \langle h \rangle,\ x_2 = w, \\ 1 & \text{if } x_1 = 0,\ x_2 \in \langle h \rangle,\ x_2 \neq w, \end{cases}$$

$\forall w \in \langle h \rangle$, $\forall i \in \{3, 4, ..., n\}$ ($I = Y = \{V_3, V_4, ..., V_n\}$), finally, we have

$$Z = (h+1) \left\{ \theta^0 \left(1 + h\theta^2\right)^{n-2} + h\theta^{n-2} \left[1 + \theta^0 + (h-1)\theta\right]^{n-2} \right\} =$$

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

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

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

**Remark 2.2.** (a) Using Theorem 2.7, it is also easy to compute the normalization constant of Potts model on an arbitrary but fixed nonempty spanning subgraph (without isolated vertices) of the complete bipartite graph $\mathcal{K}_{2,n-2}$, or, more generally, of the graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, $|\mathcal{V}| = n$ ($|\mathcal{E}| \geq 1$), which has an independent set $I$ with $|I| = n-2$, where $n \geq 3$ in both cases. Note, moreover, that the Potts model on the above graph $\mathcal{G}$ is a $\Delta$-wavy probability distribution with $|\Delta| = h+1$, where $\Delta = ...$ — see Theorem 2.5.
(b) Using Theorem 2.7, it is also easy to compute the normalization constant for the Potts model on the complete bipartite graph $\mathcal{K}_{1,n-1}$ (the star graph with $n$ vertices), where $n \geq 2$. In this case, $Z = (h+1)(h\theta + 1)^{n-1}$. This formula for $Z$ was also obtained in [16] and, moreover, will be also obtained in Section 3 by a different method.

Set

$$U_{\left(x_{s_1}=a_1, x_{s_2}=a_2, ..., x_{s_l}=a_l\right)} =$$
$$= \left\{ (y_1, y_2, ..., y_n) \mid (y_1, y_2, ..., y_n) \in \langle\langle h \rangle\rangle^n \text{ and } y_{s_b} = x_{s_b} = a_b,\ \forall b \in \langle l \rangle \right\},$$

$\forall l \in \langle n \rangle$, $\forall s_1, s_2, ..., s_l \in \langle n \rangle$, $s_u \neq s_v$, $\forall u, v \in \langle l \rangle$, $u \neq v$, $\forall a_1, a_2, ..., a_l \in \langle\langle h \rangle\rangle$.

Below we give another general formula for the normalization constant of Potts model.

**Theorem 2.8.** (Based on Theorem 2.3(vi) in [16]; see also Remark 5.1 in [18].) *Consider the Potts model on the graph $\mathcal{G}$. Then*

$$Z = (h+1) \sum_{x \in U_{(x_s=a)}} \theta^{H(x)},\ \forall s \in \langle n \rangle,\ \forall a \in \langle\langle h \rangle\rangle.$$

*Proof.* See the proof of Theorem 2.3(vi) in [16]. □

Theorems 2.7 and 2.8 are somehow related because the former is based on $\Delta$-wavy probability distributions while the latter is based on wavy probability distributions (see Theorem 1.5 and, in [18], Theorem 5.2 and Remark 5.1). In some cases, the above simple formula for the normalization constant is or seem better than that from Theorem 2.7. We illustrate this fact in the next example.

**Example 2.2.** Consider the Ising model on $\mathcal{K}_n$, the complete graph with $n$ vertices ($n \geq 2$). Consider that the vertices of $\mathcal{K}_n$ are $V_1$, $V_2$, ..., $V_n$. By Theorem 2.8 we have ($h = 1$ and we work with $s = 1$ and $a = 0$)

$$\begin{aligned} Z &= 2 \sum_{x \in U_{(x_1=0)}} \theta^{H(x)} \\ &= 2 \sum_{i \in \langle\langle n-1 \rangle\rangle} [\text{the term of } Z \text{ dues to the configurations with 0 in } V_1 \text{ and} \\ &\qquad\qquad i \text{ 1s in the other vertices}] \\ &= 2 \sum_{i \in \langle\langle n-1 \rangle\rangle} C_{n-1}^{i} \theta^{i(n-i)}. \end{aligned}$$

Now, we apply Theorem 2.7 for the independent set $\{V_n\}$ (this is a maximum independent set). Setting

$$T = \{k \mid k \in \langle n-1 \rangle - \{1\} \text{ and } x_k = 1\},$$

we have

$$\begin{aligned} Z &= 2 \sum_{(x_1,x_2,\ldots,x_{n-1}) \in \{0\} \times \langle\langle 1 \rangle\rangle^{n-2}} \theta^{H\left(y^{(x_1,x_2,\ldots,x_{n-1})}\right)} \left(1 + \theta^{c_{(1,0),n,x_1,x_2,\ldots,x_{n-1}}}\right) = \\ &= 2 \sum_{i \in \langle\langle n-2 \rangle\rangle} \sum_{\substack{(x_1,x_2,\ldots,x_{n-1}) \in \{0\} \times \langle\langle 1 \rangle\rangle^{n-2} \\ |T|=i}} \theta^{i(n-i)} \left(1 + \theta^{n-2i-1}\right) \\ &= 2 \sum_{i \in \langle\langle n-2 \rangle\rangle} C_{n-2}^{i} \theta^{i(n-i)} \left(1 + \theta^{n-2i-1}\right). \end{aligned}$$

This formula is a bit more complicated than the former one, but its sum has $n-1$ terms.

Based on the formulas from Theorems 2.7 and 2.8 we will give other ways to compute normalization constants for the Potts model in the next two sections. Sometimes, both formulas will be used, the results obtained being good or very good in some cases — for an example, see Theorem 4.1 and its proof (see also Theorems 4.2 and 4.3); another example is in Remark 3.2.

## 3. Potts model on connected separable graphs

In this section, we give a formula for the normalization constant of Potts model on a connected separable graph. This formula can be used to compute the normalization constant for the Ising or Potts model in many cases — we give a few examples, for trees, for the friendship graphs, for the windmill graphs, for the bull graph, and for others.

When we work with two or more subgraphs or graphs — sometimes, even when we work with one graph —, we will use subscripts or superscripts in the case when the energies are used, in that when the normalization constants are used, etc. *E.g.*, $H_{\mathcal{G}}$ is (denote) the energy of graph $\mathcal{G}$.

**Remark 3.1.** (See also Remark 4.14 in [16].) Consider the Potts model on a nonempty graph $\mathcal{G}$ with connected components $\mathcal{G}_1$, $\mathcal{G}_2$, ..., $\mathcal{G}_m$. Suppose that $\mathcal{G}_k$ is a nonempty subgraph of $\mathcal{G}$, $\forall k \in \langle m \rangle$. Consider the Potts model on $\mathcal{G}_k$, $\forall k \in \langle m \rangle$. It is easy to prove that

$$Z_{\mathcal{G}} = Z_{\mathcal{G}_1} Z_{\mathcal{G}_2} ... Z_{\mathcal{G}_m}$$

($Z_{\mathcal{G}}$ is the normalization constant of Potts model on $\mathcal{G}$, ...). The condition that $\mathcal{G}_k$ be a nonempty subgraph of $\mathcal{G}$, $\forall k \in \langle m \rangle$, can be removed. Indeed, removing this condition and setting by convention

$$Z_{\mathcal{G}_k} = h + 1$$

when $\mathcal{G}_k$ is a trivial subgraph (a vertex-subgraph) of $\mathcal{G}$, where $k \in \langle m \rangle$, we have

$$Z_{\mathcal{G}} = Z_{\mathcal{G}_1} Z_{\mathcal{G}_2} ... Z_{\mathcal{G}_m}.$$

Due to the above remark, it is sufficient to compute the normalization constant(s) for the Potts model on connected graphs.

**Definition 3.1.** (See, *e.g.*, [22, p. 54].) Let $\mathcal{G}$ be a connected graph. Let $\mathcal{H}$ and $\mathcal{K}$ be two subgraphs of $\mathcal{G}$. $(\mathcal{H}, \mathcal{K})$ is called a 1-*separation of* $\mathcal{G}$ if $\mathcal{H} \cup \mathcal{K} = \mathcal{G}$, $\mathcal{H} \cap \mathcal{K}$ is a vertex-graph, and $\mathcal{H}$ and $\mathcal{K}$ have each at least one edge. The vertex of $\mathcal{H} \cap \mathcal{K}$ is called the *cut-vertex of* 1-*separation* $(\mathcal{H}, \mathcal{K})$.

**Definition 3.2.** (See, *e.g.*, [22, p. 54].) Let $\mathcal{G}$ be a (connected or not) graph. $\mathcal{G}$ is called a *separable graph* if it is disconnected (nonconnected) or has a (at least one) 1-separation when it is connected.

**Definition 3.3.** (See, *e.g.*, [22, pp. 54 and 60].) Let $\mathcal{G}$ be a (connected or not) graph. Let $\mathcal{B}$ be a subgraph of $\mathcal{G}$. $\mathcal{B}$ is called a *block of* (*graph*) $\mathcal{G}$ if it is a maximal nonseparable subgraph of $\mathcal{G}$. (If $\mathcal{B}$ is a block of $\mathcal{G}$, it follows from Definition 3.2 that it is a connected subgraph of $\mathcal{G}$.)

**Definition 3.4.** (See, *e.g.*, [22, p. 64].) Let $\mathcal{G}$ be a graph. Let $\mathcal{B}$ be a block of $\mathcal{G}$. $\mathcal{B}$ is called an *extremal block* (*of* $\mathcal{G}$) if it includes exactly one cut-vertex of $\mathcal{G}$.

**Theorem 3.1.** (See, *e.g.*, [22, p. 64].) *Let $\mathcal{G}$ be a connected separable graph. Then it has at least two extremal blocks.*

*Proof.* See, *e.g.*, [22, p. 64]. (This result is based on the fact that Blk($\mathcal{G}$), the block-graph of $\mathcal{G}$, is a tree, see, *e.g.*, [22, pp. 63–64].) □

Below we give the main result of this section on the Potts model.

**Theorem 3.2.** *Consider a connected separable graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ with $\mathcal{V} = \{V_1, V_2, ...$ $..., V_n\}$ $(n \geq 3)$. Consider that its blocks are $\mathcal{B}_1$, $\mathcal{B}_2$, ..., $\mathcal{B}_m$, where $m \geq 2$. Consider the Potts model on $\mathcal{G}$. Consider the Potts model on $\mathcal{B}_k$, $\forall k \in \langle m \rangle$. Then*

$$Z_{\mathcal{G}} = \frac{1}{(h+1)^{m-1}} Z_{\mathcal{B}_1} Z_{\mathcal{B}_2} ... Z_{\mathcal{B}_m}.$$

*Proof.* Induction on $m$.

$m = 2$. In this case, $\exists j_1 \in \langle n \rangle$ such that $(\mathcal{B}_1, \mathcal{B}_2)$ is a 1-separation of $\mathcal{G}$ with cut-vertex $V_{j_1}$. Consider that $\mathcal{B}_k = (\mathcal{V}_k, \mathcal{E}_k)$, $\forall k \in \langle 2 \rangle$. We have

$$H_{\mathcal{G}}(x_{\mathcal{G}}) = H_{\mathcal{B}_1}(x_{\mathcal{B}_1}) + H_{\mathcal{B}_2}(x_{\mathcal{B}_2}), \ \forall x_{\mathcal{G}} \in \langle\langle h \rangle\rangle^n,$$

where, setting

$$T_k = \{i \mid i \in \langle n \rangle \text{ and } V_i \in \mathcal{V}_k\}, \forall k \in \langle 2 \rangle,$$

obviously (see our convention when we use two or more subgraphs or graphs again),

$$x_{\mathcal{G}} = (x_i)_{i \in \langle n \rangle}, \ x_{\mathcal{G}} \in \langle \langle h \rangle \rangle^n,$$

and

$$x_{\mathcal{B}_k} = (x_i)_{i \in T_k}, \ \forall k \in \langle 2 \rangle, \ x_{\mathcal{B}_k} \in \langle \langle h \rangle \rangle^{|T_k|} = \langle \langle h \rangle \rangle^{|\mathcal{V}_k|}, \ \forall k \in \langle 2 \rangle.$$

By Theorem 2.8 we have

$$\begin{aligned} Z_{\mathcal{G}} &= (h+1) \sum_{x_{\mathcal{G}} \in U^{\mathcal{G}}_{\left(x_{j_1}=0\right)}} \theta^{H_{\mathcal{G}}(x_{\mathcal{G}})} = (h+1) \sum_{\substack{x_{\mathcal{B}_1} \in U^{\mathcal{B}_1}_{\left(x_{j_1}=0\right)} \\ x_{\mathcal{B}_2} \in U^{\mathcal{B}_2}_{\left(x_{j_1}=0\right)}}} \theta^{H_{\mathcal{B}_1}(x_{\mathcal{B}_1}) + H_{\mathcal{B}_2}(x_{\mathcal{B}_2})} \\ &= (h+1) \sum_{x_{\mathcal{B}_1} \in U^{\mathcal{B}_1}_{\left(x_{j_1}=0\right)}} \theta^{H_{\mathcal{B}_1}(x_{\mathcal{B}_1})} \sum_{x_{\mathcal{B}_2} \in U^{\mathcal{B}_2}_{\left(x_{j_1}=0\right)}} \theta^{H_{\mathcal{B}_2}(x_{\mathcal{B}_2})} \\ &= (h+1) \sum_{x_{\mathcal{B}_1} \in U^{\mathcal{B}_1}_{\left(x_{j_1}=0\right)}} \theta^{H_{\mathcal{B}_1}(x_{\mathcal{B}_1})} \frac{1}{h+1} Z_{\mathcal{B}_2} \\ &= \frac{1}{h+1} Z_{\mathcal{B}_2} \left[ (h+1) \sum_{x_{\mathcal{B}_1} \in U^{\mathcal{B}_1}_{\left(x_{j_1}=0\right)}} \theta^{H_{\mathcal{B}_1}(x_{\mathcal{B}_1})} \right] = \frac{1}{h+1} Z_{\mathcal{B}_1} Z_{\mathcal{B}_2}. \end{aligned}$$

$m-1 \mapsto m$. By Theorem 3.1, $\exists i_1 \in \langle m \rangle$, $\exists k_1 \in \langle n \rangle$ such that $\mathcal{B}_{i_1}$ is an extremal block of $\mathcal{G}$ with cut-vertex $V_{k_1}$. It follows that $(\mathcal{B}_{i_1}, \mathcal{B}_{\neq i_1})$ is a 1-separation of $\mathcal{G}$ with cut-vertex $V_{k_1}$, where

$$\mathcal{B}_{\neq i_1} = \bigcup_{i \in \langle m \rangle, i \neq i_1} \mathcal{B}_i.$$

Consider, besides the Potts model on $\mathcal{B}_{i_1}$, the Potts model on $\mathcal{B}_{\neq i_1}$ ($\mathcal{B}_{i_1}$ and $\mathcal{B}_{\neq i_1}$ are subgraphs of $\mathcal{G}$). Finally, using the case (step) $m = 2$, we have

$$\begin{aligned} Z_{\mathcal{G}} &= \frac{1}{h+1} Z_{\mathcal{B}_{i_1}} Z_{\mathcal{B}_{\neq i_1}} = \frac{1}{h+1} Z_{\mathcal{B}_{i_1}} \left[ \frac{1}{(h+1)^{m-2}} \prod_{k \in \langle m \rangle, k \neq i_1} Z_{\mathcal{B}_k} \right] = \\ &= \frac{1}{(h+1)^{m-1}} Z_{\mathcal{B}_1} Z_{\mathcal{B}_2} \ldots Z_{\mathcal{B}_m}. \end{aligned}$$

□

Let $\mathcal{G}_1$ and $\mathcal{G}_2$ be two graphs. We write $\mathcal{G}_1 \cong \mathcal{G}_2$ if $\mathcal{G}_1$ and $\mathcal{G}_2$ are isomorphic — for isomorphic graphs and this notation, see, *e.g.*, [4, p. 40].

Further, we give a few examples for Theorem 3.2.

**Example 3.1.** Let $\mathcal{T} = (\mathcal{V}, \mathcal{E})$ be a tree with $\mathcal{V} = \{V_1, V_2, ..., V_n\}$ $(n \geq 2)$. Consider the Potts model on $\mathcal{T}$. If $n = 2$, by Theorem 2.8 we have

$$Z_{\mathcal{T}} = (h+1) \sum_{x_{\mathcal{T}} \in U^{\mathcal{T}}_{(x_1=0)}} \theta^{H_{\mathcal{T}}(x_{\mathcal{T}})} = (h+1)(1+h\theta).$$

Now, we consider that $n \geq 3$. In this case, $\mathcal{T}$ is a connected separable graph with $n-1$ blocks, $\mathcal{B}_1$, $\mathcal{B}_2$, ..., $\mathcal{B}_{n-1}$, $\mathcal{B}_1 \cong \mathcal{B}_2 \cong ... \cong \mathcal{B}_{n-1} \cong \mathcal{P}_2$, $\mathcal{P}_2$ is the path graph with 2 vertices. Consider the Potts model on $\mathcal{B}_k$, $\forall k \in \langle n-1 \rangle$. Consider the Potts model on $\mathcal{P}_2$. By Theorem 3.2 we have

$$\begin{aligned} Z_{\mathcal{T}} &= \frac{1}{(h+1)^{n-2}} (Z_{\mathcal{P}_2})^{n-1} = \\ &= \frac{1}{(h+1)^{n-2}} [(h+1)(1+h\theta)]^{n-1} = (h+1)(1+h\theta)^{n-1}. \end{aligned}$$

Therefore,

$$Z_{\mathcal{T}} = (h+1)(1+h\theta)^{n-1}, \forall n \geq 2.$$

This result was also obtained in [16], but by a different method. For $h=1$ and $\mathcal{T} = \mathcal{P}_n$ (for the 1-dimensional Ising model), $\mathcal{P}_n$ is the path graph with $n$ vertices, we have

$$Z_{\mathcal{P}_n} = 2(1+\theta)^{n-1}, \forall n \geq 2$$

(a known result, see, *e.g.*, [12, p. 36] or [16] — in [12], it is considered a different formula for energy, but each of the two formulas can be obtained from the other).

**Example 3.2.** Let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ be a connected separable graph ($|\mathcal{V}| \geq 3$). Let $\mathcal{G}_1 = (\mathcal{V}_1, \mathcal{E}_1)$ be a connected graph with $|\mathcal{V}_1| \geq 2$. Consider that the blocks of $\mathcal{G}$ are $\mathcal{B}_1$, $\mathcal{B}_2$, ..., $\mathcal{B}_m$ $(m \geq 2)$, and $\mathcal{B}_1 \cong \mathcal{B}_2 \cong ... \cong \mathcal{B}_m \cong \mathcal{G}_1$. Consider the Potts model on each of these graphs (the blocks are graphs). Then (by Theorem 3.2)

$$Z_{\mathcal{G}} = \frac{1}{(h+1)^{m-1}} (Z_{\mathcal{G}_1})^m.$$

If $h=1$ and $\mathcal{G}_1 = \mathcal{C}_k$, $\mathcal{C}_k$ is the cycle graph with $k$ vertices, we have

$$Z_{\mathcal{G}} = \frac{1}{2^{m-1}} \left((1-\theta)^k + (1+\theta)^k\right)^m$$

because

$$Z_{\mathcal{C}_k} = (1-\theta)^k + (1+\theta)^k$$

(for $Z_{\mathcal{C}_k}$, see, *e.g.*, [12, p. 35], [18], or, here, Theorem 4.3).
If $h=1$, $\mathcal{G}_1 = \mathcal{C}_3$, and the graph $\mathcal{G}$ has a cut-vertex only (therefore, all blocks $\mathcal{B}_1$, $\mathcal{B}_2$, ..., $\mathcal{B}_m$ are extremal), then $\mathcal{G} = \mathcal{F}_m$ (by definition), $\mathcal{F}_m$ is the friendship graph (with $m$ blocks, each block being isomorphic to $\mathcal{C}_3$), and

$$\begin{aligned} Z_{\mathcal{F}_m} &= \frac{1}{2^{m-1}} \left((1-\theta)^3 + (1+\theta)^3\right)^m = \\ &= \frac{1}{2^{m-1}} \left[2\left(1+3\theta^2\right)\right]^m = 2\left(1+3\theta^2\right)^m. \end{aligned}$$

If $h=1$ and $\mathcal{G}_1 = \mathcal{K}_l$, $\mathcal{K}_l$ is the complete graph with $l$ vertices, then, by Example 2.2,

$$Z_{\mathcal{G}} = \frac{1}{2^{m-1}} \left[2 \sum_{i \in \langle\langle l-1 \rangle\rangle} C_{l-1}^i \theta^{i(l-i)}\right]^m = 2\left[\sum_{i \in \langle\langle l-1 \rangle\rangle} C_{l-1}^i \theta^{i(l-i)}\right]^m.$$

If $h=1$, $\mathcal{G}_1 = \mathcal{K}_l$, and the graph $\mathcal{G}$ has a cut-vertex only, then $\mathcal{G}$=Wd$(l,m)$ (by definition), Wd$(l,m)$ is the windmill graph (with $m$ blocks, each block being isomorphic to $\mathcal{K}_l$), and

$$Z_{\mathrm{Wd}(l,m)} = 2\left[\sum_{i\in\langle\langle l-1\rangle\rangle} C_{l-1}^{i}\theta^{i(l-i)}\right]^{m}.$$

**Example 3.3.** Let $\mathcal{G}=(\mathcal{V},\mathcal{E})$ be a connected separable graph ($|\mathcal{V}|\geq 3$). Consider that the blocks of $\mathcal{G}$ are $\mathcal{B}_1$, $\mathcal{B}_2$, ..., $\mathcal{B}_m$ $(m\geq 2)$. Suppose that $\mathcal{B}_1\cong\mathcal{B}_2\cong ...\cong\mathcal{B}_u\cong\mathcal{P}_2$ and $\mathcal{B}_{u+1}\cong\mathcal{B}_{u+2}\cong ...\cong\mathcal{B}_m\cong\mathcal{C}_k$, where $u\in\langle m-1\rangle$. Consider the Potts model on each of these graphs. Then (by Theorem 3.2)

$$\begin{aligned} Z_{\mathcal{G}} &= \frac{1}{(h+1)^{m-1}}\left(Z_{\mathcal{P}_2}\right)^{u}\left(Z_{\mathcal{C}_k}\right)^{m-u} = \\ &= \frac{1}{(h+1)^{m-1}}\left[(h+1)(1+h\theta)\right]^{u}\left(Z_{\mathcal{C}_k}\right)^{m-u} = \frac{1}{(h+1)^{m-u-1}}(1+h\theta)^{u}\left(Z_{\mathcal{C}_k}\right)^{m-u} \end{aligned}$$

If $h=1$ and $k=3$, we have

$$Z_{\mathcal{G}} = \frac{1}{2^{m-u-1}}(1+\theta)^{u}\left[2\left(1+3\theta^{2}\right)\right]^{m-u} = 2(1+\theta)^{u}\left(1+3\theta^{2}\right)^{m-u}.$$

If $h=1$, $m=3$, $u=2$, the blocks $\mathcal{B}_1$ and $\mathcal{B}_2$ are extremal, and $k=3$, then $\mathcal{G}=$ the bull graph (by definition) and

$$Z_{\mathcal{G}} = 2(1+\theta)^{2}\left(1+3\theta^{2}\right).$$

**Remark 3.2.** The case of Theorem 3.2 when we have at least $m-1$ blocks isomorphic to $\mathcal{P}_2$ can be proved by induction by a different method, using Theorems 2.7 and 2.8. The proof is based on the fact that there exists an extremal block isomorphic to $\mathcal{P}_2$ both for the step $m=2$ and for the step $m-1\mapsto m$. We do the proof for the step $m=2$ only. Suppose that $\mathcal{B}_1\cong\mathcal{P}_2$ and has the vertices $V_1$ and $V_2$. Suppose that $V_2$ is the cut-vertex of 1-separation $(\mathcal{B}_1,\mathcal{B}_2)$. Consider the independent set $\{V_1\}$. By Theorems 2.7 and 2.8 we have

$$\begin{aligned} Z_{\mathcal{G}} &= (h+1)\sum_{(x_2,x_3,\ldots,x_n)_{\mathcal{G}}\in\{0\}\times\langle\langle h\rangle\rangle^{n-2}}\theta^{H_{\mathcal{G}}\left(y_{\mathcal{G}}^{(x_2,x_3,\ldots,x_n)_{\mathcal{G}}}\right)}\left(1+\sum_{w\in\langle h\rangle}\theta^{c^{\mathcal{G}}_{(w,0),1,x_2}}\right)= \\ &= (h+1)\sum_{x_{\mathcal{B}_2}\in U^{\mathcal{B}_2}_{(x_2=0)}}\theta^{H_{\mathcal{B}_2}\left(x_{\mathcal{B}_2}\right)}(1+h\theta) = \\ &= (1+h\theta)\left[(h+1)\sum_{x_{\mathcal{B}_2}\in U^{\mathcal{B}_2}_{(x_2=0)}}\theta^{H_{\mathcal{B}_2}\left(x_{\mathcal{B}_2}\right)}\right] = \frac{1}{h+1}Z_{\mathcal{P}_2}Z_{\mathcal{B}_2} = \frac{1}{h+1}Z_{\mathcal{B}_1}Z_{\mathcal{B}_2}. \end{aligned}$$

## 4. Potts model on graphs with a vertex of degree 2

In this section, under certain conditions, we give a formula for the normalization constant of Potts model on a graph with a vertex of degree 2. This formula leads to a recurrence relation for the normalization constant of Potts model on $\mathcal{C}_n$, the cycle graph with $n$ vertices ($n\geq 3$), and, further, we compute the normalization constant of Ising model on $\mathcal{C}_n$.

In this section, besides the Potts model on graphs, we must work with the Potts model on multigraphs — we work with nondirected finite multigraphs without loops. If $\mathcal{G}$ is a nonempty nondirected finite graph with loops or a nondirected finite multigraph with loops and $\mathcal{G}'$ is the graph or multigraph obtained from it by deleting/removing the loops, we set by convention

$$H_{\mathcal{G}}(x) = H_{\mathcal{G}'}(x), \ \forall x \in \langle\langle h \rangle\rangle^n,$$

supposing that $\mathcal{G}'$ is nonempty when it is a graph, where $n$ = the order of $\mathcal{G}$ (= the number of vertices of $\mathcal{G}$, see, *e.g.*, [4, p. 19]), $\langle\langle h \rangle\rangle$ = the set of colors of $\mathcal{G}$, ... So, we can work with $\mathcal{G}'$ instead of $\mathcal{G}$. The definition of Potts model on multigraphs (nondirected finite multigraphs without loops) is the same as that from Section 2 for the Potts model on graphs, with the difference that the edge set from there is replaced with an edge multiset. *E.g.*, considering the multigraph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, $\mathcal{V} = \{V_1, V_2, V_3\}$, $\mathcal{E} = \{[V_1, V_2], [V_1, V_2], [V_2, V_3], [V_3, V_1]\}$, and the Potts model on $\mathcal{G}$, we have, *e.g.*, $H(0,1,0) = 3$ ($x_1 = 0$ (0 is the color of $V_1$), $x_2 = 1$, $x_3 = 0$), not $H(0,1,0) = 2$, because $V_1$ and $V_2$ are joint by two edges, and $\pi_{(0,1,0)} = \frac{\theta^3}{Z}$.

It is easy to see that Theorem 2.8 can be extended for the Potts model on multigraphs. Moreover, Theorems 2.1, 2.2, and 2.4–2.7 can also be extended for the Potts model on multigraphs — good exercises for the reader! Moreover, Theorem 3.2 can also be extended for the Potts model on connected separable multigraphs —another good exercise for the reader!

In the next result, we introduce a new method, a "superposition" method, to compute the normalization constants for the Potts model, and which is based on Theorems 2.7 and 2.8, and the extension of Theorem 2.8 for the Potts model on multigraphs.

**Theorem 4.1.** *Let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ be a graph with $\mathcal{V} = \{V_1, V_2, ..., V_n\}$, where $|\mathcal{V}| = n \geq 4$ and $|\mathcal{E}| \geq 3$. Suppose that $degV_n = 2$ ($degV_n$ = the degree of $V_n$). Suppose that the vertices adjacent to $V_n$ are $V_{n-2}$ and $V_{n-1}$ ($V_{n-2}$ and $V_{n-1}$ are adjacent or not). Further, we construct a graph and a graph or multigraph. We delete the vertex $V_n$ and edges $[V_{n-2}, V_n]$ and $[V_{n-1}, V_n]$ (these edges are the incident edges with $V_n$),* and *obtain the graph, say, $\mathcal{G}_1$, $\mathcal{G}_1 = (\mathcal{V}_1, \mathcal{E}_1)$,*

$$\mathcal{V}_1 = \mathcal{V} - \{V_n\}, \ \mathcal{E}_1 = \mathcal{E} - \{[V_{n-2}, V_n], [V_{n-1}, V_n]\}.$$

*We then "superpose" $V_{n-2}$ on $V_{n-1}$ in $\mathcal{G}_1$, i.e., we remove (delete) $V_{n-1}$ from $\mathcal{V}_1$ and, in $\mathcal{E}_1$, each edge $[X, V_{n-1}]$ with $X \neq V_{n-2}$, if any, is replaced with the edge $[X, V_{n-2}]$, then, if $[V_{n-2}, V_{n-1}] \in \mathcal{E}_1$, this edge is removed from $\mathcal{E}_1$, and obtain the graph or multigraph (without loops), say, $\mathcal{G}_2$, $\mathcal{G}_2 = (\mathcal{V}_2, \mathcal{E}_2)$,*

$$\mathcal{V}_2 = \mathcal{V}_1 - \{V_{n-1}\} = \mathcal{V} - \{V_{n-1}, V_n\},$$

$$\mathcal{E}_2 = \{[X,Y] \mid [X,Y] \in \mathcal{E}_1 \text{ and } X, Y \neq V_{n-1}\} \cup \{[X, V_{n-2}] \mid [X, V_{n-1}] \in \mathcal{E}_1, \ X \neq V_{n-2}\},$$

*$\mathcal{E}_2$ is a set or a multiset, $\cup$ is the union of sets or multisets. Consider the Potts model on each of $\mathcal{G}$, $\mathcal{G}_1$, $\mathcal{G}_2$. Then*

$$Z_{\mathcal{G}} = \theta\left[2 + (h-1)\theta\right] Z_{\mathcal{G}_1} + (1-\theta)^2 Z_{\mathcal{G}_2}.$$

*Proof.* Consider the independent set $\{V_n\}$ of $\mathcal{G}$. By Theorem 2.7 we have

$$Z_{\mathcal{G}} = (h+1) \sum_{(x_1,x_2,\dots,x_{n-1})_{\mathcal{G}} \in \langle\langle h\rangle\rangle^{n-2}\times\{0\}} \left[\theta^{H_{\mathcal{G}}\left(y_{\mathcal{G}}^{(x_1,x_2,\dots,x_{n-1})_{\mathcal{G}}}\right)} \cdot \left(1+\sum_{w\in\langle h\rangle}\theta^{c^{\mathcal{G}}_{(w,0),n,x_{n-2},x_{n-1}}}\right)\right]$$

$$= (h+1)\left\{\sum_{(x_1,x_2,\dots,x_{n-1})_{\mathcal{G}} \in \langle\langle h\rangle\rangle^{n-3}\times\{0\}^2} \theta^{H_{\mathcal{G}}\left(y_{\mathcal{G}}^{(x_1,x_2,\dots,x_{n-1})_{\mathcal{G}}}\right)}\left(1+h\theta^2\right) + \right.$$

$$\left. + \sum_{(x_1,x_2,\dots,x_{n-1})_{\mathcal{G}} \in \langle\langle h\rangle\rangle^{n-3}\times\langle h\rangle\times\{0\}} \theta^{H_{\mathcal{G}}\left(y_{\mathcal{G}}^{(x_1,x_2,\dots,x_{n-1})_{\mathcal{G}}}\right)}\left[2+(h-1)\,\theta\right]\right\} =$$

(below, in the second term, the factor $\theta$ is due to the fact that $x_{n-2}\in\langle h\rangle$, $x_{n-1}=0$, $\left(y_{\mathcal{G}}^{(x_1,x_2,\dots,x_{n-1})_{\mathcal{G}}}\right)_n = 0$ (see the definition of $y^{\left(x_{j_1},x_{j_2},\dots,x_{j_q}\right)}$ in Theorem 2.6; $\left(y_{\mathcal{G}}^{(x_1,x_2,\dots,x_{n-1})_{\mathcal{G}}}\right)_n$ is the $n$th component of $y_{\mathcal{G}}^{(x_1,x_2,\dots,x_{n-1})_{\mathcal{G}}}$), and (the graph) $\mathcal{G}_1$ will be used instead of $\mathcal{G}$)

$$= (h+1)\left(1+h\theta^2\right) \sum_{(z_1,z_2,\dots,z_{n-1})_{\mathcal{G}_1} \in \langle\langle h\rangle\rangle^{n-3}\times\{0\}^2} \theta^{H_{\mathcal{G}_1}\left((z_1,z_2,\dots,z_{n-1})_{\mathcal{G}_1}\right)} +$$

$$+ (h+1)\left[2+(h-1)\,\theta\right]\theta \sum_{(z_1,z_2,\dots,z_{n-1})_{\mathcal{G}_1} \in \langle\langle h\rangle\rangle^{n-3}\times\langle h\rangle\times\{0\}} \theta^{H_{\mathcal{G}_1}\left((z_1,z_2,\dots,z_{n-1})_{\mathcal{G}_1}\right)} =$$

$$= (h+1)\left(1+h\theta^2\right) \sum_{z_{\mathcal{G}_1}\in U^{\mathcal{G}_1}_{(z_{n-2}=z_{n-1}=0)}} \theta^{H_{\mathcal{G}_1}(z_{\mathcal{G}_1})} +$$

$$+ (h+1)\left[2+(h-1)\,\theta\right]\theta \sum_{a\in\langle h\rangle} \sum_{z_{\mathcal{G}_1}\in U^{\mathcal{G}_1}_{(z_{n-2}=a,z_{n-1}=0)}} \theta^{H_{\mathcal{G}_1}(z_{\mathcal{G}_1})}.$$

By Theorem 2.8 we have

$$Z_{\mathcal{G}_1} = (h+1) \sum_{z_{\mathcal{G}_1}\in U^{\mathcal{G}_1}_{(z_{n-1}=0)}} \theta^{H_{\mathcal{G}_1}(z_{\mathcal{G}_1})} =$$

$$= (h+1)\left(\sum_{z_{\mathcal{G}_1}\in U^{\mathcal{G}_1}_{(z_{n-2}=z_{n-1}=0)}} \theta^{H_{\mathcal{G}_1}(z_{\mathcal{G}_1})} + \sum_{a\in\langle h\rangle} \sum_{z_{\mathcal{G}_1}\in U^{\mathcal{G}_1}_{(z_{n-2}=a,z_{n-1}=0)}} \theta^{H_{\mathcal{G}_1}(z_{\mathcal{G}_1})}\right),$$

so,

$$(h+1)\sum_{a\in\langle h\rangle} \sum_{z_{\mathcal{G}_1}\in U^{\mathcal{G}_1}_{(z_{n-2}=a,z_{n-1}=0)}} \theta^{H_{\mathcal{G}_1}(z_{\mathcal{G}_1})} = Z_{\mathcal{G}_1} - (h+1) \sum_{z_{\mathcal{G}_1}\in U^{\mathcal{G}_1}_{(z_{n-2}=z_{n-1}=0)}} \theta^{H_{\mathcal{G}_1}(z_{\mathcal{G}_1})}.$$

For the Potts model on $\mathcal{G}_2$, using Theorem 2.8 (when $\mathcal{G}_2$ is a graph) or its extension (when $\mathcal{G}_2$ is a multigraph), we have

$$Z_{\mathcal{G}_2} = (h+1) \sum_{z_{\mathcal{G}_2} \in U^{\mathcal{G}_2}_{(z_{n-2}=0)}} \theta^{H_{\mathcal{G}_2}(z_{\mathcal{G}_2})} = (h+1) \sum_{z_{\mathcal{G}_1} \in U^{\mathcal{G}_1}_{(z_{n-2}=z_{n-1}=0)}} \theta^{H_{\mathcal{G}_1}(z_{\mathcal{G}_1})}.$$

Finally, we have

$$\begin{aligned} Z_{\mathcal{G}} &= \left(1+h\theta^2\right) Z_{\mathcal{G}_2} + \theta\left[2+(h-1)\theta\right]\left(Z_{\mathcal{G}_1} - Z_{\mathcal{G}_2}\right) = \\ &= \theta\left[2+(h-1)\theta\right] Z_{\mathcal{G}_1} + (1-\theta)^2 Z_{\mathcal{G}_2}. \end{aligned}$$

□

It is easy to see — a good exercise for the reader! — that Theorem 4.1 can be generalized — if $\mathcal{G}$ is either a graph with a vertex of degree 2 or a multigraph whose underlying graph (see, *e.g.*, [4, p. 30] for this graph) has a vertex of degree 2, and we then construct a graph or multigraph, $\mathcal{G}_1$, and we then construct a graph or multigraph, $\mathcal{G}_2$, ... (for the completion, see Theorem 4.1, see also its proof — some things will be similar to those from Theorem 4.1), we obtain a generalization of Theorem 4.1.

Below we give an application of the above result. This application is for the Potts model on $\mathcal{C}_n$ (the cycle graph with $n$ vertices). One reason to study this model is the following: the Potts model on $\mathcal{C}_n$ can be seen as a 1-dimensional Potts model with cyclic boundary condition as the Ising model on $\mathcal{C}_n$ is seen, see, *e.g.*, [12, pp. 31–32], as a 1-dimensional Ising model with cyclic boundary condition. For another reason, see the next section (Theorem 5.2, Remark 5.1, ...).

**Theorem 4.2.** *Consider the Potts model on $C_n$, $\forall n \geq 3$. Then*

$$Z_{\mathcal{C}_{n+1}} = (h+1)\,\theta\left[2+(h-1)\theta\right](1+h\theta)^{n-1} + (1-\theta)^2 Z_{\mathcal{C}_{n-1}}, \forall n \geq 4.$$

*Proof.* Let $n \geq 4$. By Theorem 4.1, taking $\mathcal{G} = \mathcal{C}_{n+1}$, we have $\mathcal{G}_1 = \mathcal{P}_n$ (the path graph with $n$ vertices) and $\mathcal{G}_2 = \mathcal{C}_{n-1}$, and, further,

$$Z_{\mathcal{C}_{n+1}} = \theta\left[2+(h-1)\theta\right] Z_{\mathcal{P}_n} + (1-\theta)^2 Z_{\mathcal{C}_{n-1}} =$$

(see Example 3.1 for $Z_{\mathcal{P}_n}$)

$$= (h+1)\,\theta\left[2+(h-1)\theta\right](1+h\theta)^{n-1} + (1-\theta)^2 Z_{\mathcal{C}_{n-1}}.$$

□

Using the recurrence relation from Theorem 4.2, below we compute the normalization constant for the Ising model on $\mathcal{C}_n$. For other two computation methods for this constant, see [12, pp. 31–35] (in this book, for this constant, it is given an equivalent formula to the formula from our article) and [18].

**Theorem 4.3.** *Consider the Ising model on $C_n$, $\forall n \geq 3$. Then*

$$Z_{\mathcal{C}_n} = (1-\theta)^n + (1+\theta)^n, \ \forall n \geq 3.$$

*Proof.* Induction on $n$.
$n = 3$. By Theorem 2.8 we have $(h = 1)$

$$Z_{\mathcal{C}_3} = 2\left(1+3\theta^2\right).$$

(Theorem 2.7 can also be used.) Since

$$(1-\theta)^3 + (1+\theta)^3 = 2\left(1+3\theta^2\right),$$

we have

$$Z_{\mathcal{C}_3} = (1-\theta)^3 + (1+\theta)^3 .$$

$n = 4$. Similar to the case $n = 3$.
$n-1 \mapsto n+1$. By Theorem 4.2 we have

$$\begin{aligned} Z_{\mathcal{C}_{n+1}} &= 4\theta\,(\theta+1)^{n-1} + (1-\theta)^2\left[(1-\theta)^{n-1} + (1+\theta)^{n-1}\right] = \\ &= 4\theta\,(\theta+1)^{n-1} + (1-\theta)^{n+1} + (1-\theta)^2\,(1+\theta)^{n-1} = \\ &= (1-\theta)^{n+1} + \left(4\theta + 1 - 2\theta + \theta^2\right)(1+\theta)^{n-1} = \\ &= (1-\theta)^{n+1} + \left(1 + 2\theta + \theta^2\right)(1+\theta)^{n-1} = \\ &= (1-\theta)^{n+1} + (1+\theta)^2\,(1+\theta)^{n-1} = (1-\theta)^{n+1} + (1+\theta)^{n+1} . \end{aligned}$$

□

## 5. Bounds

Simple expressions, closed-form expressions for the normalization constant of Potts model in concrete cases are possible — we think so — in a small number of such cases (see, *e.g.*, Examples 3.1, 3.2, and 3.3). Such expressions are also possible in the limit in some cases. We do not hope more. So, for this constant, we must find approximations and lower and upper bounds — if possible, good and very good approximations, good and very good lower and upper bounds. In this section, we present two ways to obtain bounds for the normalization constant of Potts model. One of these ways is for lower bounds, and is based on Theorem 2.7, while the other is for lower and upper bounds, and is based on connected separable spanning subgraphs and Theorem 3.2. We will illustrate these two ways — moreover, for one of the illustrative examples, two upper bounds for the free energy per site are given, one of them being in the limit.

Each term of the sum from Theorem 2.7 is a lower bound for the normalization constant of Potts model. Several such terms by summing up lead to a better lower bound for this constant — the larger the number of terms is, the better the lower bound is. We give just one result here — computing a "big" term of the sum from Theorem 2.7, below it is given a lower bound for the normalization constant of Potts model.

**Theorem 5.1.** *Under the same conditions as in Theorem 2.7 we have*

$$Z \geq (h+1) \prod_{l \in \langle p \rangle} \left(1 + h\theta^{\deg V_{i_l}}\right),$$

*where* $\deg V_{i_l}$ *is the degree of* $V_{i_l}$, $\forall l \in \langle p \rangle$ .

*Proof.* By Theorem 2.7 we have

$$\begin{aligned} Z &\geq (h+1)\,\theta^{H\left(y^{(0,0,\ldots,0)}\right)} \prod_{l\in\langle p\rangle}\left(1+\sum_{w\in\langle h\rangle}\theta^{c_{(w,0),i_l,0,0,\ldots,0}}\right) = \\ &= (h+1)\prod_{l\in\langle p\rangle}\left(1+h\theta^{\deg V_{i_l}}\right). \end{aligned}$$

□

Let $\mathcal{G}=(\mathcal{V},\mathcal{E})$ be a graph. Let $J_1, J_2 \subseteq \mathbb{R}$ be two nonempty finite sets. Suppose that

$$\mathcal{V}=\left\{V_{(z_1,z_2)} \mid (z_1,z_2)\in J_1\times J_2\right\}.$$

Let $(a,b)\in\mathbb{R}^2$. Consider the graph $\mathcal{G}+(a,b)=(\mathcal{V}+(a,b),\mathcal{E}+(a,b))$, where

$$\mathcal{V}+(a,b)=\left\{V_{(z_1+a,z_2+b)} \;\middle|\; V_{(z_1,z_2)}\in\mathcal{V}\right\},$$

$$\mathcal{E}+(a,b)=\left\{\left[V_{(u_1+a,u_2+b)},V_{(z_1+a,z_2+b)}\right] \;\middle|\; \left[V_{(u_1,u_2)},V_{(z_1,z_2)}\right]\in\mathcal{E}\right\}.$$

We call the graph $\mathcal{G}+(a,b)$ the $(a,b)$*-translated graph of* $\mathcal{G}$.

Consider the 2-dimensional grid graph $\mathcal{G}_{n_1,n_2}=(\mathcal{V}_{n_1,n_2},\mathcal{E}_{n_1,n_2})$ of dimensions $n_1$ and $n_2$, where $n_1$, $n_2\geq 1$ ($n_1$, $n_2\in\mathbb{N}$), $n_1n_2\geq 2$,

$$\mathcal{V}_{n_1,n_2}=\left\{V_{(i_1,i_2)} \mid (i_1,i_2)\in\langle n_1\rangle\times\langle n_2\rangle\right\},$$

and

$$\begin{aligned}\mathcal{E}_{n_1,n_2}=\{\left[V_{(i_1,i_2)},V_{(j_1,j_2)}\right] \mid\ & (i_1,i_2),(j_1,j_2)\in\langle n_1\rangle\times\langle n_2\rangle \text{ and}\\ & \text{either } j_1=i_1 \text{ and } j_2-i_2=1 \text{ or } j_1-i_1=1 \text{ and } j_2=i_2\}.\end{aligned}$$

Further, we construct a connected separable spanning subgraph of $\mathcal{G}_{n_1,n_2}$ when $n_1=n_2=6k$, where $k\geq 1$. The blocks of this subgraph are isomorphic to $\mathcal{C}_4$ or $\mathcal{P}_2$, and are constructed as follows.

$$\mathcal{B}_1=(\mathcal{V}_1,\mathcal{E}_1)\cong\mathcal{C}_4,$$

where

$$\mathcal{V}_1=\left\{V_{(1,1)},V_{(2,1)},V_{(2,2)},V_{(1,2)}\right\}$$

and

$$\mathcal{E}_1=\left\{\left[V_{(1,1)},V_{(2,1)}\right],\left[V_{(2,1)},V_{(2,2)}\right],\left[V_{(2,2)},V_{(1,2)}\right],\left[V_{(1,2)},V_{(1,1)}\right]\right\},$$

$$\mathcal{B}_2=\mathcal{B}_1+(1,1),$$

$$\mathcal{B}_3=\mathcal{B}_2+(1,-1)=\mathcal{B}_1+(2,0),$$

$$\mathcal{B}_4=\mathcal{B}_3+(1,1),$$

$$\mathcal{B}_5=\mathcal{B}_4+(1,-1),$$

$$\vdots$$

$$\mathcal{B}_{6k-2}=\mathcal{B}_{6k-3}+(1,1),$$

$$\mathcal{B}_{6k-1}=\mathcal{B}_{6k-2}+(1,-1).$$

Using the blocks $\mathcal{B}_1$, $\mathcal{B}_2$, ..., $\mathcal{B}_{6k-1}$, we construct (consider) the blocks

$$\mathcal{B}_{t(6k-1)+1}=\mathcal{B}_1+(0,3t),$$

$$\mathcal{B}_{t(6k-1)+2}=\mathcal{B}_2+(0,3t),$$

$$\vdots$$

$$\mathcal{B}_{t(6k-1)+6k-1} = \mathcal{B}_{6k-1} + (0, 3t)\,,$$

$\forall t,\ 1 \leq t \leq 2k-1$ $(1 + 3t \leq 6k-2 \Longrightarrow t \leq 2k-1)$.

All the above blocks ($\mathcal{B}_1$, $\mathcal{B}_2$, ..., $\mathcal{B}_{6k-1}$, ...) are isomorphic to $\mathcal{C}_4$, and the number of them is (equal to) $2k\,(6k-1)$ because

$$6k - 1 + (6k-1) \max_{1 \leq t \leq 2k-1} t = 6k - 1 + (6k-1)(2k-1) = 2k\,(6k-1)\,.$$

Now, we construct the blocks which are isomorphic to $\mathcal{P}_2$. The subgraph

$$\left(\left\{V_{(1,u)}, V_{(1,u+1)}\right\},\ \left\{\left[V_{(1,u)}, V_{(1,u+1)}\right]\right\}\right),$$

where $2 \leq u \leq 6k-1$, is considered to be a block if $\left[V_{(1,u)}, V_{(1,u+1)}\right]$ is not an edge of any above block which is isomorphic to $\mathcal{C}_4$.

For each $u \in \{2, 3, 5, 6, 8, 9, ..., 6k-4, 6k-3, 6k-1\}$, we obtain such a block. For $u = 3$, we have $V_{(1,u+1)} = V_{(1,4)} \in \mathcal{B}_{1\cdot(6k-1)+1}$ (here, $t = 1$); for $u = 6$, we have $V_{(1,u+1)} = V_{(1,7)} \in \mathcal{B}_{2(6k-1)+1}$ (here, $t = 2$); ...; for $u = 6k-3$, we have $V_{(1,u+1)} = V_{(1,6k-2)} \in \mathcal{B}_{(2k-1)(6k-1)+1}$ (here, $t = 2k-1$). So, the number of these blocks is $4k-1$ (because

$$2 \max_{1 \leq t \leq 2k-1} t + 1 = 2\,(2k-1) + 1 = 4k - 2 + 1 = 4k - 1).$$

The subgraph

$$\mathcal{B}' = \left(\left\{V_{(6k,2)}, V_{(6k,3)}\right\},\ \left\{\left[V_{(6k,2)}, V_{(6k,3)}\right]\right\}\right)$$

is considered to be a block. The subgraphs $\mathcal{B}' + (0, 3t)\,,\ 1 \leq t \leq 2k-1$ $(2 + 3t \leq 6k - 1 \Longrightarrow t \leq 2k-1)$ are also considered blocks. The number of these blocks, $\mathcal{B}'$ and $\mathcal{B}' + (0, 3t)\,,\ 1 \leq t \leq 2k-1$, is $2k$. We finished the construction of blocks which are isomorphic to $\mathcal{P}_2$. The number of these blocks is $4k - 1 + 2k$, *i.e.*, $6k - 1$.

For the next result, we need to compute (to know) $Z_{\mathcal{P}_2}$ and $Z_{\mathcal{C}_4}$. By Theorem 2.8,

$$Z_{\mathcal{P}_2} = (h+1)\,(1 + h\theta)\,.$$

To compute $Z_{\mathcal{C}_4}$, consider $\mathcal{C}_4 = (\mathcal{V}, \mathcal{E})\,,$

$$\mathcal{V} = \{V_1, V_2, V_3, V_4\} \text{ and } \mathcal{E} = \{[V_1, V_2]\,, [V_2, V_3]\,, [V_3, V_4]\,, [V_4, V_1]\}\,.$$

By Theorem 2.7 for the maximum independent set $\{V_2, V_4\}$ we have

$$Z_{\mathcal{C}_4} = (h+1) \sum_{(x_1,x_3)_{\mathcal{C}_4} \in \{0\} \times \langle\langle h \rangle\rangle} \left[\theta^{H_{\mathcal{C}_4}\left(y_{\mathcal{C}_4}^{(x_1,x_3)_{\mathcal{C}_4}}\right)} \cdot \right.$$

$$\left. \cdot \left(1 + \sum_{w \in \langle h \rangle} \theta^{c^{\mathcal{C}_4}_{(w,0),2,x_1,x_3}}\right) \left(1 + \sum_{w \in \langle h \rangle} \theta^{c^{\mathcal{C}_4}_{(w,0),4,x_1,x_3}}\right)\right].$$

Since

$$H_{\mathcal{C}_4}\left(y_{\mathcal{C}_4}^{(x_1,x_3)_{\mathcal{C}_4}}\right) = \begin{cases} 0 & \text{if } x_1 = x_3 = 0, \\ 2 & \text{if } x_1 = 0,\ x_3 \in \langle h \rangle \end{cases}$$

and

$$c^{\mathcal{C}_4}_{(w,0),2,x_1,x_3} = c^{\mathcal{C}_4}_{(w,0),4,x_1,x_3} = \begin{cases} 2 & \text{if } x_1 = x_3 = 0, \\ 0 & \text{if } x_1 = 0,\ x_3 \in \langle h \rangle\,,\ x_3 = w, \\ 1 & \text{if } x_1 = 0,\ x_3 \in \langle h \rangle\,,\ x_3 \neq w, \end{cases}$$

$\forall w \in \langle h \rangle$, it follows that

$$Z_{\mathcal{C}_4} = (h+1)\left\{\left(1+h\theta^2\right)^2 + h\theta^2\left[2+(h-1)\theta\right]^2\right\}.$$

**Theorem 5.2.** *Let $n = 6k$, where $k \geq 1$. Consider the Potts model on (the grid graph) $G_{n,n}$.*
*(i) If $0 < \theta < 1$, then*

$$Z_{\mathcal{G}_{n,n}} \leq (h+1)\left\{\left(1+h\theta^2\right)^2 + h\theta^2\left[2+(h-1)\theta\right]^2\right\}^{\frac{n}{3}(n-1)}(1+h\theta)^{n-1}.$$

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

$$Z_{\mathcal{G}_{n,n}} \geq (h+1)\left\{\left(1+h\theta^2\right)^2 + h\theta^2\left[2+(h-1)\theta\right]^2\right\}^{\frac{n}{3}(n-1)}(1+h\theta)^{n-1}.$$

*Proof.* Denote by $\mathcal{G}$ the connected separable spanning subgraph of $\mathcal{G}_{n,n}$ constructed above. It is easy to prove that (a similar case is in the proof of Theorem 4.1 in [16])

$$H_{\mathcal{G}_{n,n}}(x) \geq H_{\mathcal{G}}(x), \ \forall x \in \langle\langle h\rangle\rangle^{n^2}$$

($x = x_{\mathcal{G}_{n,n}} = x_{\mathcal{G}}$; $x_{\mathcal{G}_{n,n}} = x_{\mathcal{G}}$ because $\mathcal{G}_{n,n}$ and $\mathcal{G}$ have the same vertices, $n^2$ vertices). By Theorem 3.2 we have

$$\begin{aligned} Z_{\mathcal{G}} &= \frac{1}{(h+1)^{(2k+1)(6k-1)-1}}\left(Z_{\mathcal{C}_4}\right)^{2k(6k-1)}\left(Z_{\mathcal{P}_2}\right)^{6k-1} = \\ &= (h+1)\left\{\left(1+h\theta^2\right)^2 + h\theta^2\left[2+(h-1)\theta\right]^2\right\}^{\frac{n}{3}(n-1)}(1+h\theta)^{n-1}. \end{aligned}$$

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

$$\begin{aligned} Z_{\mathcal{G}_{n,n}} &= \sum_{x\in\langle\langle h\rangle\rangle^{n^2}} \theta^{H_{\mathcal{G}_{n,n}}(x)} \leq \sum_{x\in\langle\langle h\rangle\rangle^{n^2}} \theta^{H_{\mathcal{G}}(x)} = Z_{\mathcal{G}} = \\ &= (h+1)\left\{\left(1+h\theta^2\right)^2 + h\theta^2\left[2+(h-1)\theta\right]^2\right\}^{\frac{n}{3}(n-1)}(1+h\theta)^{n-1}. \end{aligned}$$

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

$$\begin{aligned} Z_{\mathcal{G}_{n,n}} &= \sum_{x\in\langle\langle h\rangle\rangle^{n^2}} \theta^{H_{\mathcal{G}_{n,n}}(x)} \geq \sum_{x\in\langle\langle h\rangle\rangle^{n^2}} \theta^{H_{\mathcal{G}}(x)} = Z_{\mathcal{G}} = \\ &= (h+1)\left\{\left(1+h\theta^2\right)^2 + h\theta^2\left[2+(h-1)\theta\right]^2\right\}^{\frac{n}{3}(n-1)}(1+h\theta)^{n-1}. \end{aligned}$$

□

**Remark 5.1.** (a) In Theorem 5.2, $n = 6k$. The other cases, $n = 6k+1$, $n = 6k+2$, ..., $n = 6k+5$, can also be studied.
(b) The spanning trees of connected separable spanning subgraph $\mathcal{G}$ of $\mathcal{G}_{n,n}$ from the proof of Theorem 5.2 are also spanning trees of $\mathcal{G}_{n,n}$. Consider such a spanning tree, say, $\mathcal{T}$. By Theorem 4.1 from [16] or proceeding as in the proof of Theorem 5.2 we have

$$Z_{\mathcal{G}} \leq Z_{\mathcal{T}} = (h+1)(1+h\theta)^{n-1} \text{ if } 0 < \theta < 1$$

and

$$Z_{\mathcal{G}} \geq Z_{\mathcal{T}} = (h+1)(1+h\theta)^{n-1} \text{ if } \theta \geq 1.$$

Further, we have (see the proof of Theorem 5.2)

$$Z_{\mathcal{G}_{n,n}} \le Z_{\mathcal{G}} \le Z_{\mathcal{T}} \text{ if } 0 < \theta < 1$$

and

$$Z_{\mathcal{G}_{n,n}} \ge Z_{\mathcal{G}} \ge Z_{\mathcal{T}} \text{ if } \theta \ge 1.$$

Therefore, in Theorem 5.2, we obtained bounds for $Z_{\mathcal{G}_{n,n}}$ better than those from Theorem 4.1 in [16].

(c) If we know the normalization constant for the Potts model on a given graph, $\mathcal{G}' = (\mathcal{V}', \mathcal{E}')$, or bounds — if possible, good and very good bounds for it —, we can compute other things on this model (see, *e.g.*, [16]; see, *e.g.*, also [3, p. 6]). Consider, *e.g.*, the free energy per site, $f_{n'}^{\mathcal{G}'}$,

$$f_{n'}^{\mathcal{G}'} = \frac{\ln Z_{\mathcal{G}'}}{n'}$$

(see, *e.g.*, [16]), where $n' = |\mathcal{V}'|$. Further, we consider $\mathcal{G}_{n,n}$, $\mathcal{G}$, and $\mathcal{T}$ from (b). Consider that $0 < \theta < 1$; the case when $\theta \ge 1$ is left to the reader. By (b) we have $(n = 6k)$

$$f_{n^2}^{\mathcal{G}_{n,n}} = \frac{\ln Z_{\mathcal{G}_{n,n}}}{n^2} \le f_{n^2}^{\mathcal{G}} = \frac{\ln Z_{\mathcal{G}}}{n^2} \le f_{n^2}^{\mathcal{T}} = \frac{\ln Z_{\mathcal{T}}}{n^2}.$$

It follows that

$$\limsup_{n\to\infty} f_{n^2}^{\mathcal{G}_{n,n}} \le \limsup_{n\to\infty} f_{n^2}^{\mathcal{G}} \le \limsup_{n\to\infty} f_{n^2}^{\mathcal{T}}$$

$(n \to \infty \Longrightarrow n^2 \to \infty)$. From Theorem 4.12(iii) in [16], we have

$$\limsup_{n\to\infty} f_{n^2}^{\mathcal{T}} = \lim_{n\to\infty} f_{n^2}^{\mathcal{T}} = \ln(h\theta + 1);$$

$\lim\limits_{n\to\infty} f_{n^2}^{\mathcal{T}}$ is the limit free energy per site, see [16], of (or for the) Potts model on $\mathcal{T}$. $\limsup\limits_{n\to\infty} f_{n^2}^{\mathcal{G}}$ can be computed;

$$\limsup_{n\to\infty} f_{n^2}^{\mathcal{G}} = \limsup_{n\to\infty} \frac{\ln Z_{\mathcal{G}}}{n^2} =$$

$$= \limsup_{n\to\infty} \frac{\ln (h+1) \left\{ \left(1 + h\theta^2\right)^2 + h\theta^2 \left[2 + (h-1)\theta\right]^2 \right\}^{\frac{n}{3}(n-1)} (1 + h\theta)^{n-1}}{n^2} =$$

$$= \lim_{n\to\infty} \frac{\ln(h+1)}{n^2} + \lim_{n\to\infty} \frac{\ln \left\{ \left(1 + h\theta^2\right)^2 + h\theta^2 \left[2 + (h-1)\theta\right]^2 \right\}^{\frac{n}{3}(n-1)}}{n^2} +$$

$$+ \lim_{n\to\infty} \frac{\ln(1 + h\theta)^{n-1}}{n^2} = \lim_{n\to\infty} \frac{\frac{n^2}{3} - \frac{n}{3}}{n^2} \ln \left\{ \left(1 + h\theta^2\right)^2 + h\theta^2 \left[2 + (h-1)\theta\right]^2 \right\} +$$

$$+ \lim_{n\to\infty} \frac{n-1}{n^2} \ln(1 + h\theta) = \frac{1}{3} \ln \left\{ \left(1 + h\theta^2\right)^2 + h\theta^2 \left[2 + (h-1)\theta\right]^2 \right\}$$

($\lim\limits_{n\to\infty} f_{n^2}^{\mathcal{G}}$ exists; $\limsup\limits_{n\to\infty} f_{n^2}^{\mathcal{G}} = \lim\limits_{n\to\infty} f_{n^2}^{\mathcal{G}}$). Finally, we have

$$\limsup_{n\to\infty} f_{n^2}^{\mathcal{G}_{n,n}} \le \frac{1}{3} \ln \left\{ \left(1 + h\theta^2\right)^2 + h\theta^2 \left[2 + (h-1)\theta\right]^2 \right\} \le \ln(h\theta + 1).$$

Therefore, we obtained a bound (an upper bound) for $\limsup\limits_{n\to\infty} f_{n^2}^{\mathcal{G}_{n,n}}$ better than $\ln(h\theta + 1)$ (recall that we considered the case when $0 < \theta < 1$ only; the bound

$\ln(h\theta+1)$ also appears in Theorem 4.12(i) in [16]; Theorem 4.12(i) in [16] can be generalized replacing "lim" with "lim sup"). If we know or can prove — mathematical proof, not (physical or not) arguments or postulates —that $\lim_{n\to\infty} f_{n^2}^{\mathcal{G}_{n,n}}$ exists — for the Potts (not Ising) model on $\mathcal{G}_{n,n}$ —, then, in the above inequality and other places, $\limsup_{n\to\infty} f_{n^2}^{\mathcal{G}_{n,n}}$ can be replaced with $\lim_{n\to\infty} f_{n^2}^{\mathcal{G}_{n,n}}$. We call $\limsup_{n\to\infty} f_{n^2}^{\mathcal{G}_{n,n}}$ the *superior limit free energy per site of* (or *for the*) *Potts model on* $\mathcal{G}_{n,n}$.

Recall that the bounds given in this section are, first of all, illustrative — if they are useful or not, this is another story. This subject can much be developed. The reader, if he/she wants, can try, *e.g.*, to give bounds for the normalization constant of Potts model on the 3-dimensional grid graph.

## 6. Sampling

In this section, we give a method for sampling from $\langle\langle h\rangle\rangle^n$ according to the Potts model and some comments for it. For the Potts model on $\mathcal{G}_{n_1,n_2,...,n_d}$, the $d$-dimensional grid graph, $d\geq 1$, $n_1$, $n_2$, ..., $n_d\geq 1$, $n_1n_2...n_d\geq 2$, we obtain an exact sampling method for half or half+1 vertices.

Consider the Potts model $\pi=(\pi_x)_{x\in\langle\langle h\rangle\rangle^n}$ on the graph $\mathcal{G}=(\mathcal{V},\mathcal{E})$, $\mathcal{V}=\{V_1,V_2,...$ $...,V_n\}$ $(n\geq 2)$, from Section 2. Recall that the graph $\mathcal{G}$ has no isolated vertices (see Section 2 again). In fact, on sampling, it is sufficient to consider only this case because in the case when the graph $\mathcal{G}$ has isolated vertices we can proceed as follows: each isolated vertex is colored with the color $i$ with the probability $\frac{1}{h+1}$, $\forall i\in\langle\langle h\rangle\rangle$. Recall that $(x_1,x_2,...,x_n)$, $x_1$, $x_2$, ..., $x_n\in\langle\langle h\rangle\rangle$, are the configurations of (graph) $\mathcal{G}$; $x\in\langle\langle h\rangle\rangle^n$, $\langle\langle h\rangle\rangle^n$ is the set of configurations.

Let $I=\{V_{i_1},V_{i_2},...,V_{i_p}\}$ be an independent set of $\mathcal{G}$, $p\geq 1$ $(I\neq\emptyset$; $p\geq 1\Longrightarrow I\neq\emptyset)$, better, a maximal independent set of $\mathcal{G}$, best, a maximum independent set of $\mathcal{G}$. Consider $I^c=\{V_{j_1},V_{j_2},...,V_{j_q}\}$, the complement of $I$, where $p+q=n=|\mathcal{V}|$ $(q\geq 1)$. The Potts model on $\mathcal{G}$ is a $\Delta$-wavy probability distribution, where

$$\Delta=\left(S_{\left(x_{j_1},x_{j_2},...,x_{j_q}\right)}\right)_{\left(x_{j_1},x_{j_2},...,x_{j_q}\right)\in\{0\}\times\langle\langle h\rangle\rangle^{q-1}},$$

see Theorem 2.5. Using the sampling method for the $\Delta$-wavy probability distributions from Section 1, we obtain the following sampling method for the Potts model on $\mathcal{G}$ — this method can also be used exactly or approximately.

*Step* 1. Sample from

$$\Delta=\left(S_{\left(x_{j_1},x_{j_2},...,x_{j_q}\right)}\right)_{\left(x_{j_1},x_{j_2},...,x_{j_q}\right)\in\{0\}\times\langle\langle h\rangle\rangle^{q-1}}$$

$(\Delta\in\mathrm{Par}(\langle\langle h\rangle\rangle^n))$ according to the probability distribution

$$\tau=\left(\tau_{S_{\left(x_{j_1},x_{j_2},...,x_{j_q}\right)}}\right)_{\left(x_{j_1},x_{j_2},...,x_{j_q}\right)\in\{0\}\times\langle\langle h\rangle\rangle^{q-1}},$$

where, see Step 1 from Section 1 and Theorem 2.6,

$$\tau_{S_{\left(x_{j_1},x_{j_2},\ldots,x_{j_q}\right)}}=P\left(S_{\left(x_{j_1},x_{j_2},\ldots,x_{j_q}\right)}\right)=$$
$$=\frac{h+1}{Z}\theta^{H\left(y^{\left(x_{j_1},x_{j_2},\ldots,x_{j_q}\right)}\right)}\prod_{l\in\langle p\rangle}\left(1+\sum_{w\in\langle h\rangle}\theta^{c_{(w,0),i_l,x_{w_1},x_{w_2},\ldots,x_{w_{s(i_l)}}}}\right),$$

$\forall\left(x_{j_1},x_{j_2},...,x_{j_q}\right)\in\{0\}\times\langle\langle h\rangle\rangle^{q-1}$, where... — for the completion, see Theorem 2.6.

Suppose that the result of sampling is, say, $S_{\left(b_{j_1},b_{j_2},\ldots,b_{j_q}\right)}$, where $\left(b_{j_1},b_{j_2},...,b_{j_q}\right)\in\{0\}\times\langle\langle h\rangle\rangle^{q-1}$.

*Step* 2. Sample from $S_{\left(b_{j_1},b_{j_2},\ldots,b_{j_q}\right)}$ according to the probability distribution

$$\pi|_{S_{\left(b_{j_1},b_{j_2},\ldots,b_{j_q}\right)}}=\left(\frac{\pi_x}{\tau_{S_{\left(b_{j_1},b_{j_2},\ldots,b_{j_q}\right)}}}\right)_{x\in S_{\left(b_{j_1},b_{j_2},\ldots,b_{j_q}\right)}}=$$
$$=\left(\frac{\pi_x}{P\left(S_{\left(b_{j_1},b_{j_2},\ldots,b_{j_q}\right)}\right)}\right)_{x\in S_{\left(b_{j_1},b_{j_2},\ldots,b_{j_q}\right)}}.$$

Suppose that the result of sampling is, say, $z$, where $z\in\langle\langle h\rangle\rangle^n$, $z_{j_1}=b_{j_1}$, $z_{j_2}=b_{j_2}$, ..., $z_{j_q}=b_{j_q}$.

$z$ is the result of sampling from $\langle\langle h\rangle\rangle^n$ according to the Potts model $\pi$ (see the sampling method from Section 1 again).

Step 1 is a challenging problem because $|\Delta|=(h+1)^{q-1}$, $q=n-p$ ($p=|I|$, so, we need an independent set as large as possible), and the components of $\tau$ are not too simple ($Z$ is known or not, ...). We could obtain good results at Step 1 using Theorem 1.4 if $\tau$ is a nontrivial wavy probability distribution or, more generally, if $\tau$ is a nontrivial $\Gamma$-wavy probability distribution, $|\Gamma|$ being sufficiently small. At first glance, the worst case for Step 1 is when $\mathcal{G}\cong\mathcal{K}_n$ ($\mathcal{K}_n=$ the complete graph) because $\{V_i\}$ is a maximum independent set, so, $p=1$, $\forall i\in\langle n\rangle$.

As to Step 2, if we use the Gibbs sampler in a generalized sense from Theorem 1.4, this chain attains its stationarity at time 1, so, we have an exact sampling method having, see Theorem 2.4, $p+1$ steps ($p+1$ substeps of Step 2) — $2\leq p+1\leq n$; $p+1=2$ when $\mathcal{G}\cong\mathcal{K}_n$, and a maximum independent set of this graph is considered; $p+1=n$ when $\mathcal{G}$ is the star graph (with $n$ vertices), and the maximum independent set of this graph is considered.

For Step 2, we consider the case when $\mathcal{G}=\mathcal{G}_{n_1,n_2,\ldots,n_d}$. $\mathcal{G}_{n_1,n_2,\ldots,n_d}$ is the $d$-dimensional grid graph of dimensions $n_1$, $n_2$, ..., $n_d$,

$$\mathcal{G}_{n_1,n_2,\ldots,n_d}=\left(\mathcal{V}_{n_1,n_2,\ldots,n_d},\mathcal{E}_{n_1,n_2,\ldots,n_d}\right),$$

where $d\geq1$, $n_1$, $n_2$, ..., $n_d\geq1$, $n_1n_2...n_d\geq2$,

$$\mathcal{V}_{n_1,n_2,\ldots,n_d}=\left\{V_{(i_1,i_2,\ldots,i_d)}\mid\ (i_1,i_2,...,i_d)\in\langle n_1\rangle\times\langle n_2\rangle\times...\times\langle n_d\rangle\right\},$$

and

$$\mathcal{E}_{n_1,n_2,\ldots,n_d}=$$
$$=\left\{\left[V_{(i_1,i_2,\ldots,i_d)},V_{(j_1,j_2,\ldots,j_d)}\right]\mid\ (i_1,i_2,...,i_d),(j_1,j_2,...,j_d)\in\langle n_1\rangle\times\langle n_2\rangle\times...\times\langle n_d\rangle\right.$$

$$\text{and } \exists! k \in \langle d \rangle \text{ such that } j_k - i_k = 1 \text{ and } j_u = i_u,\ \forall u \in \langle d \rangle - \{k\}\}$$

($\exists!$ = there exists a unique).

$\mathcal{G}_{n_1,n_2,...,n_d}$ is a bipartite graph, its bipartition is $(X, Y)$, where

$$X = \begin{cases} \left\{V_{(i_1,i_2,...,i_d)} \,\middle|\, V_{(i_1,i_2,...,i_d)} \in \mathcal{V}_{n_1,n_2,...,n_d} \text{ and } i_1 + i_2 + ... + i_d \text{ is even}\right\} \\ \quad \text{if } d \text{ is even,} \\ \left\{V_{(i_1,i_2,...,i_d)} \,\middle|\, V_{(i_1,i_2,...,i_d)} \in \mathcal{V}_{n_1,n_2,...,n_d} \text{ and } i_1 + i_2 + ... + i_d \text{ is odd}\right\} \\ \quad \text{if } d \text{ is odd,} \end{cases}$$

and $Y = X^c$ (any edge of $\mathcal{G}_{n_1,n_2,...,n_d}$ has one end in $X$ and one end in $Y$, so, $(X, Y)$ is a bipartition of $\mathcal{G}_{n_1,n_2,...,n_d}$ — this is unique because $\mathcal{G}_{n_1,n_2,...,n_d}$ is a connected graph).

Some of the above things on $\mathcal{G}_{n_1,n_2,...,n_d}$ and the next result, on $\mathcal{G}_{n_1,n_2,...,n_d}$ too, could be known to the reader, but this fact is not too important, it is important the fact that we need them.

**Theorem 6.1.** (i) $V_{(1,1,...,1)} \in X$.
*(ii)* $|X| = |Y|$ *if* $n_1 n_2 ... n_d$ *is even.*
(iii) $|X| = |Y| + 1$ *if* $n_1 n_2 ... n_d$ *is odd.*

*Proof.* (i) Obvious.
(ii) Suppose that $n_1 n_2 ... n_d$ is even. Then $\exists k \in \langle d \rangle$ such that $n_k$ is even. Set

$$A_k = \left\{V_{(i_1,i_2,...,i_d)} \,\middle|\, V_{(i_1,i_2,...,i_d)} \in \mathcal{V}_{n_1,n_2,...,n_d} \text{ and } i_s = 1,\ \forall s \in \langle d \rangle - \{k\}\right\}.$$

Therefore,

$$A_k = \left\{V_{(1,1,...,1)}\ (i_k = 1), V_{(1,1,...,1,2,1,...,1)}\ (i_k = 2), ..., V_{(1,1,...,1,n_k,1,...,1)}\ (i_k = n_k)\right\}.$$

The set $A_k$ has $n_k$ elements (vertices), $\frac{n_k}{2}$ of them belong to $X$ and the other $\frac{n_k}{2}$ belong to $Y$ ($V_{(1,1,...,1)} \in X$ (see (i)), $V_{(1,1,...,1,2,1,...,1)} \in Y$, ..., $V_{(1,1,...,1,n_k,1,...,1)} \in Y$ (because $n_k$ is even)).
Further, we denote elements of $\langle d \rangle - \{k\}$ by $k_1, k_2, ..., k_{d-1}$. Suppose that $k_1 < k_2 < ... < k_{d-1}$ ($k_1 = 1$ if $k \neq 1$, $k_1 = 2$ if $k = 1$, etc.). Set

$$A_{k,k_1,k_2,...,k_t} = \left\{V_{(i_1,i_2,...,i_d)} \,\middle|\, V_{(i_1,i_2,...,i_d)} \in \mathcal{V}_{n_1,n_2,...,n_d} \right.$$
$$\left. \text{and } i_s = 1,\ \forall s \in \langle d \rangle - \{k, k_1, k_2, ..., k_t\}\right\},$$

$\forall t \in \langle d - 1 \rangle$. The condition "$i_s = 1, \forall s \in \langle d \rangle - \{k, k_1, k_2, ..., k_t\}$" from the definition of $A_{k,k_1,k_2,...,k_t}$ vanishes when $t = d - 1$, so, $A_{k,k_1,k_2,...,k_{d-1}} = \mathcal{V}_{n_1,n_2,...,n_d}$. It follows that

$$A_{k,k_1,k_2,...,k_t} = \bigcup_{b \in \langle n_{k_t} \rangle} \left\{V_{(i_1,i_2,...,i_d)} \,\middle|\, V_{(i_1,i_2,...,i_d)} \in \mathcal{V}_{n_1,n_2,...,n_d},\ i_{k_t} = b, \right.$$
$$\left. \text{and } i_s = 1,\ \forall s \in \langle d \rangle - \{k, k_1, k_2, ..., k_t\}\right\} =$$
$$= A_{k,k_1,k_2,...,k_{t-1}} \cup \bigcup_{b \in \langle n_{k_t} \rangle - \{1\}} \left\{V_{(i_1,i_2,...,i_d)} \,\middle|\, V_{(i_1,i_2,...,i_d)} \in \mathcal{V}_{n_1,n_2,...,n_d},\ i_{k_t} = b, \right.$$
$$\left. \text{and } i_s = 1,\ \forall s \in \langle d \rangle - \{k, k_1, k_2, ..., k_t\}\right\},$$

$\forall t \in \langle d - 1 \rangle$ ($A_{k,k_1,k_2,...,k_t} = A_{k,k_1,k_2,...,k_{t-1}}$ if $n_{k_t} = 1$ $(t \in \langle d - 1 \rangle)$) and

$$\left|A_{k,k_1,k_2,...,k_{t-1}}\right| = \left|\left\{V_{(i_1,i_2,...,i_d)} \,\middle|\, V_{(i_1,i_2,...,i_d)} \in \mathcal{V}_{n_1,n_2,...,n_d},\ i_{k_t} = b, \right.\right.$$
$$\left.\left. \text{and } i_s = 1,\ \forall s \in \langle d \rangle - \{k, k_1, k_2, ..., k_t\}\right\}\right|,$$

$\forall t \in \langle d-1 \rangle$ , $\forall b \in \langle n_{k_t} \rangle - \{1\}$ ($k_1$, $k_2$, ..., $k_{t-1}$ vanish when $t = 1$).
The set $A_{k,k_1}$ has $n_k n_{k_1}$ elements ($n_k n_{k_1}$ is even), $\frac{n_k n_{k_1}}{2}$ of them belong to $X$ and the other $\frac{n_k n_{k_1}}{2}$ belong to $Y$ because $A_{k,k_1} = A_k \cup ...$ and $|A_k| = ...$ (see above), and $A_k$ has $\frac{n_k}{2}$ elements belonging to $X$ and $\frac{n_k}{2}$ elements belonging to $Y$.
Proceeding in this way for $A_{k,k_1,k_2}$, for $A_{k,k_1,k_2,k_3}$, ..., for $A_{k,k_1,k_2,...,k_{d-1}}$, we obtain that the set $\mathcal{V}_{n_1,n_2,...,n_d}$ ($A_{k,k_1,k_2,...,k_{d-1}} = \mathcal{V}_{n_1,n_2,...,n_d}$) has $n_k n_{k_1} ... n_{k_{d-1}}$ elements, $\frac{n_k n_{k_1} ... n_{k_{d-1}}}{2}$ of them belong to $X$ and the other $\frac{n_k n_{k_1} ... n_{k_{d-1}}}{2}$ belong to $Y$. Therefore,

$$|X| = |Y| = \frac{n_1 n_2 ... n_d}{2}.$$

(iii) Suppose that $n_1 n_2 ... n_d$ is odd. In this case, $n_1$, $n_2$, ..., $n_d$ are odd numbers. We can use the above sets $A_k$ and $A_{k,k_1,k_2,...,k_t}$, $t \in \langle d-1 \rangle$ , with only the difference that $k$ is chosen from $\langle d \rangle$ by us (here, $n_k$ is odd). Further, we use the above sets and take $k = 1$.
The set $A_1$ has $n_1$ elements, $\left\lfloor \frac{n_1}{2} \right\rfloor + 1$ of them belong to $X$ and the other $\left\lfloor \frac{n_1}{2} \right\rfloor$ belong to $Y$ ($V_{(1,1,...,1)} \in A_1 \cap X$; $\lfloor x \rfloor = \max \{ z \mid z \in \mathbb{Z} \text{ and } z \le x \}$ , $\forall x \in \mathbb{R}$). It follows that the set $A_{1,2}$ has $n_1 n_2$ elements, $\left\lfloor \frac{n_1 n_2}{2} \right\rfloor + 1$ of them belong to $X$ and the other $\left\lfloor \frac{n_1 n_2}{2} \right\rfloor$ belong to $Y$. Proceeding in this way for $A_{1,2,3}$, for $A_{1,2,3,4}$, ..., for $A_{1,2,...,d}$, we obtain that $\mathcal{V}_{n_1,n_2,...,n_d}$ ($A_{1,2,...,d} = \mathcal{V}_{n_1,n_2,...,n_d}$) has $n_1 n_2 ... n_d$ elements, $\left\lfloor \frac{n_1 n_2 ... n_d}{2} \right\rfloor + 1$ of them belong to $X$ and the other $\left\lfloor \frac{n_1 n_2 ... n_d}{2} \right\rfloor$ belong to $Y$. Therefore,

$$|X| = \left\lfloor \frac{n_1 n_2 ... n_d}{2} \right\rfloor + 1 > |Y| = \left\lfloor \frac{n_1 n_2 ... n_d}{2} \right\rfloor .$$

□

Since $|X| \ge |Y|$ (by Theorem 6.1), we take $I = X$, and have

$$p = |I| = |X| = \begin{cases} \frac{n_1 n_2 ... n_d}{2} & \text{if } n_1 n_2 ... n_d \text{ is even,} \\ \left\lfloor \frac{n_1 n_2 ... n_d}{2} \right\rfloor + 1 & \text{if } n_1 n_2 ... n_d \text{ is odd.} \end{cases}$$

Therefore, for Step 2, using the Gibbs sampler in a generalized sense from Theorem 1.4, we have an exact sampling method for the Potts model on $\mathcal{G}_{n_1,n_2,...,n_d}$ having, see Theorem 2.4, $\frac{n_1 n_2 ... n_d}{2} + 1$ steps (substeps of Step 2) if $n_1 n_2 ... n_d$ is even and $\left\lfloor \frac{n_1 n_2 ... n_d}{2} \right\rfloor + 2$ steps if $n_1 n_2 ... n_d$ is odd — an exact sampling method for half or half+1 vertices of the grid graph.

It remains to find, if any, a fast exact sampling method for the Potts model on $\mathcal{G}_{n_1,n_2,...,n_d}$. For the Potts model on $\mathcal{G}_{n_1}$, we have a fast exact sampling method having $n_1$ steps — $n_1$ vertices, $n_1$ steps, one-to-one correspondence — leaving the inversion of a bijective function aside, see [16].

At present we know that our sampling method for the Potts model with Steps 1 and 2 is — using Theorem 1.4, ... — fast and exact in some cases, such as, when the graph is $\mathcal{K}_{2,n}$, $n \ge 2$ (the more general case when the graph is $\mathcal{K}_{m,n}$, the complete bipartite graph, $1 \le m \le n$, could be analyzed), because, in this case, we have good things both for Step 1 and for Step 2; for Step 1, $P\left(S_{(0,j)}\right) = P\left(S_{(0,k)}\right)$ , $\forall j, k \in \langle h \rangle$ (hint: use bijective functions and $\oplus$; due these equations, $\tau$ is a nice probability distribution, it is an almost uniform probability distribution, so, the computation of probabilities $P\left(S_{(0,i)}\right)$ , $i \in \langle h \rangle$ , is not necessary (see, *e.g.*, [19, Comment 4] for the almost uniform probability distributions...)), $P\left(S_{(0,0)}\right)$ can easy be computed, $P\left(S_{(0,0)}\right) + hP\left(S_{(0,1)}\right) = 1$ (we can use this equation if we want to compute the probabilities $P\left(S_{(0,i)}\right)$, $i \in \langle h \rangle$), $Z$ is computed in Example 2.1 (for $\mathcal{K}_{2,n-2}$), and we can take $\Gamma \preceq \Gamma' = \left(\left\{S_{(0,0)}\right\}, \left\{S_{(0,1)}, S_{(0,2)}, ..., S_{(0,h)}\right\}\right)$ such that $\tau$ be a $\Gamma$-wavy probability distribution — and we can use the sampling method from Section 1 (not from this section) for the $\Gamma$-wavy probability distribution $\tau$ —, we can take $\Gamma = \Gamma'$ when $h$ is not too large, considering, in this latter case, that $\tau$ is a trivial wavy probability distribution on (the subset) $\left\{S_{(0,1)}, S_{(0,2)}, ..., S_{(0,h)}\right\}$ ($\tau$ is a trivial wavy probability distribution on $\left\{S_{(0,0)}\right\}$ (by convention)) while, as to Step 2, using the Gibbs sampler in a generalized sense (Theorems 1.4 and 2.4), our method has $n+1$ steps ($n+1$ substeps of Step 2; $n \longmapsto n+1$ is a very good polynomial function (in $n$)). For the case when the graph is $\mathcal{K}_{1,n}$, see [16] — the fast exact sampling method from there, which is for the Potts model on $\mathcal{K}_{1,n-1}$ (we worked with the star graph with $n$ vertices in [16]), is, in fact, the sampling method from here for the Potts model on $\mathcal{K}_{1,n-1}$ using Theorem 1.4, ...

Note that, mathematically speaking — the technology is not taken into account —, our sampling method for the Potts model depends on $\theta$, $h$, and (the graph) $\mathcal{G}$. For the case when the graph is $\mathcal{K}_{m,n}$, it depends on $\theta$, $h$, $m$, and $n$ — the smaller $h$, $m$, and $n$ are, the faster our sampling method is; as to $\theta$, $\theta^{mn}$ is the quantity with the greatest exponent we need for Step 1. Note, moreover, that $\mathcal{K}_{m,n}$ has no cycles when $m = 1$ ($n \geq 1$) while it has cycles when $2 \leq m \leq n$ ($\mathcal{K}_{2,2} = \mathcal{G}_{2,2} = \mathcal{C}_4$), so, we can have fast exact sampling both when the graphs have no cycles and when they have cycles. The graphs which have no cycles are the trees ($\mathcal{G}_{n_1}$ is a tree, $\forall n_1 \geq 2$ ) and their generalizations, the forests, see [16] for fast exact sampling (and other things) for the Potts model on these graphs.

*What is the fastest exact sampling method we can have (obtain) for the Potts model, in particular, for the Potts model on* $\mathcal{G}_{n_1,n_2,...,n_d}$?

## References


[1] A.S. Asratian, T.M.J. Denley, and R. Häggkvist, *Bipartite Graphs and Their Applications*, Cambridge University Press, Cambridge, 1998.

[2] D. Avis, A. Hertz, and O. Marcotte (Eds.), *Graph Theory and Combinatorial Optimization*, Springer, New York, 2005.

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

[4] A. Benjamin, G. Chartrand, and P. Zhang, *The Fascinating World of Graph Theory*, Princeton University Press, Princeton, 2015.

[5] J.A. Bondy and U.S.R. Murty, *Graph Theory with Applications*, North-Holland, New York, 1979.

[6] G. Chartrand and P. Zhang, *Chromatic Graph Theory*, CRC Press, Boca Raton, 2008.

[7] J.A. De Loera, X. Goaoc, F. Meunier, and N.H. Mustafa, The discrete yet ubiquitous theorems of Carathéodory, Helly, Sperner, Tucker, and Tverberg, *Bull. Amer. Math. Soc. (N.S.)* **56** (2019), 415−511.

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

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

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

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

[12] B.M. McCoy and T.T. Wu, *The Two-Dimensional Ising Model*, 2nd Edition, Dover, Mineola, N.Y., 2014.
[13] U. Păun, $G_{\Delta_1,\Delta_2}$ in action, *Rev. Roumaine Math. Pures Appl.* **55** (2010), 387−406.
[14] U. Păun, A hybrid Metropolis-Hastings chain, *Rev. Roumaine Math. Pures Appl.* **56** (2011), 207−228.
[15] U. Păun, $G$ method in action: from exact sampling to approximate one, *Rev. Roumaine Math. Pures Appl.* **62** (2017), 413−452.
[16] U. Păun, $G$ method in action: normalization constant, important probabilities, and fast exact sampling for Potts model on trees, *Rev. Roumaine Math. Pures Appl.* **65** (2020), 103−130.
[17] U. Păun, A Gibbs sampler in a generalized sense, *An. Univ. Craiova Ser. Mat. Inform.* **43** (2016), 62−71.
[18] U. Păun, A Gibbs sampler in a generalized sense, II, *An. Univ. Craiova Ser. Mat. Inform.* **45** (2018), 103−121.
[19] U. Păun, Ewens distribution on $S_n$ is a wavy probability distribution with respect to $n$ partitions, *An. Univ. Craiova Ser. Mat. Inform.* **47** (2020), 1−24.
[20] R.B. Potts, Some generalized order-disorder transitions, *Proc. Cambridge Philos. Soc.* **48** (1952), 106−109.
[21] E. Seneta, *Non-negative Matrices and Markov Chains*, 2nd Edition, Springer-Verlag, Berlin, 1981; revised printing, 2006.
[22] W.T. Tutte, *Graph Theory*, Cambridge University Press, Cambridge, 2001.
[23] 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 5, Romania
*E-mail address*: `paun@csm.ro`