Canonical Paths for MCMC: from Art to Science

Lingxiao Huang *   Pinyan Lu †   Chihao Zhang ‡

Abstract

Markov Chain Monte Carlo (MCMC) method is a widely used algorithm design scheme with many applications. To make efficient use of this method, the key step is to prove that the Markov chain is rapid mixing. Canonical paths is one of the two main tools to prove rapid mixing. However, there are much fewer success examples comparing to coupling, the other main tool. The main reason is that there is no systematic approach or general recipe to design canonical paths. Building up on a previous exploration by McQuillan [18], we develop a general theory to design canonical paths for MCMC: We reduce the task of designing canonical paths to solving a set of linear equations, which can be automatically done even by a machine.

Making use of this general approach, we obtain fully polynomial-time randomized approximation schemes (FPRAS) for counting the number of b-matching with $b \leq 7$ and b-edge-cover with $b \leq 2$. They are natural generalizations of matchings and edge covers for graphs. No polynomial time approximation was previously known for these problems.

1 Introduction

In statistics and computer science, Markov Chain Monte Carlo (MCMC) methods are a class of algorithms for sampling from a probability distribution based on constructing a Markov chain that has the desired distribution as its stationary (equilibrium) distribution. The state of the chain after a number of (random) steps is then used as a sample of the desired distribution. MCMC methods are primarily used for calculating approximations of multi-dimensional integrals, number of combinational objects, number of solutions for constraint satisfaction problems, partition function for statistic physics systems and so on [4, 6, 8, 7, 9, 10, 12, 13, 14, 19, 20, 21]. Typically, the support set of the distribution is exponentially large but we need the sampling algorithm to run in polynomial time. This requires that the Markov chain is rapidly mixing, namely, it is very close to the stationary distribution after polynomial number of steps.

Canonical path is one of the two main tools (the other one is coupling) to prove rapid mixing of the Markov chain. To make use of this tool, one need to design paths between each pair of states for the Markov chain and prove that the overall congestion at each link of the Markov chain is low. However, it is typically a very difficult task to come up with a low congestion routing especially for an exponentially large state graph of a Markov chain. Thus, the design of canonical paths for a given Markov chain remains a highly non-trivial artwork for masters. For the other main tool coupling, there are quite a few nice theories developed. One most important general approach is path coupling [3] which enables one to only analysis the local configuration of a single constraint rather than the global configuration. This is typically much easier to handle.

Due to the lack of general theory and approach, there are only very few notably successful examples of canonical path. One important example is the MCMC for sampling and counting matchings of a graph [11]. The states of the Markov chain is all matchings for a given input graph. The symmetric difference of two matchings of a graph is a disjoint union of paths and cycles. Then, the natural and success canonical path for matchings is "winding" the edges one by one just follow the natural order of these paths and cycles. Another important success example is the so called "sub-graph world" problem transformed from ferromagnetic Ising model [12]. For this problem, the symmetric difference of two configurations can be any graphs. But any graph has path-cycle decompositions, and their canonical paths simply do an arbitrary path-cycle decomposition and wind the edges following these paths and cycles. Since the constraint in each vertex for that problem is the simple parity function, they can prove that these canonical paths indeed have low congestion.

*Institute of Interdisciplinary Information Sciences, Tsinghua University. huanglingxiao1990@126.com. This author is supported in part by the National Basic Research Program of China Grant 2015CB358700, 2011CBA00300, 2011CBA00301, the National Natural Science Foundation of China Grant 61202009, 61033001, 61361136003.

†Microsoft Research. pinyanl@microsoft.com

‡Shanghai Jiao Tong University. chihao.zhang@gmail.com. This author is supported in part by the National Natural Science Foundation of China Grant 61261130589, 61472239.

In an unpublished manuscript [18], McQuillan proposed a beautiful generalization of this path-cycle decomposition idea called winding. In a high-level, one do not use a single fixed path-cycle decomposition but use a convex combination of exponentially many path-cycle decompositions and distribute the flow among these canonical paths. This idea itself alone is not new, such fractional canonical paths were used before, see for example [19]. The main contribution of [18] is a method to design such a convex combination by a local property for each constraint called windable. As long as each local constraint is windable, they can design the global path-cycle decompositions and thus canonical paths automatically. Therefore, this winding approach gives a systematic approach to design canonical paths for MCMC. This is similar to path coupling technique for coupling which enables us to only analysis the local constraint and configurations. However, to show that this windable property for the local constraints still require a construction for some mathematical objects. In their paper, they showed that the Not-All-Equal functions satisfies the properties by an explicit construction of these mathematical objects. It was not clear how to show whether a new constraint function satisfies this windable property or not.

In this paper, we give a characterization for the property of windable by a set of linear equations, which works both for unweighed and weighted constraints. Having that, the whole process of designing canonical paths becomes a routine of solving linear equations which can be automatically done by a machine. We also refine some definitions and presentation for the winding approach so that it is easier to understand and apply. We extend this approach to instances with edge weights as well.

It is very easy to verify that the matching constraint [11] and parity function [12] are indeed windable by our characterization. Moreover, with this powerful approach and characterization in hand, we design a number of new fully polynomial-time randomized approximation schemes (FPRAS) for approximate counting by simply verifying that the local constraint functions are windable by our new characterization theorem. Our first example is counting b-matchings, which is a natural generalization of matchings. A subset of edges for a graph is called a b-matching if every vertex is incident to at most b edges in the set. 1-matching is the conventional definition of matching for a graph. In particular, we obtain FPRAS for counting b-matchings with $b \leq 7$ for any graphs. Previously, FPRAS was only known for counting 1-matchings.

Another problem we resolve is a generalization of the edge cover problem. A subset of edges for a graph is called an edge cover if every vertex is incident to at least one edge in the set. Previously, MCMC based approximation algorithm for counting edge covers was only known for 3-regular graphs [2]. In fact, they also used canonical path to get rapid mixing and used path-cycle decomposition to construct canonical paths. Since they do not have a systematic approach but some ad-hoc construction and case-by-case analysis, they only succeeded for the very special 3-regular graphs. By our approach and characterization, we can show that there exist a convex combination of path-cycle decompositions which works for general graphs. Moreover, we generalize it to b-edge-cover by requiring that every vertex is incident to at least b edges in the set. We obtain FPRAS for counting b-edge-cover for $b \leq 2$. We note that FPTAS based on correlation decay technique for counting edge covers for general graphs was known [16, 17]. However, it seems that their technique have intrinsic difficulty for 2-edge-cover.

Interestingly, we can show that the constraint function of 8-matchings and 3-edge-cover are not windable by our characterization theorem. We do not know whether these transitions really corresponds to the boundaries of approximability or not. We leave these as interesting open questions.

The most interesting future direction is to design canonical paths for other Markov chains by this approach and thus get polynomial time approximation algorithms. Of course, we are not claiming that winding is the only way to design canonical paths. To develop other systematic approach for designing and analyzing canonical paths for MCMC is very interesting. We hope that our work can stimulate such kind research.

2 Preliminaries

Holant Problem. Let $G(V, E)$ be a graph. In this paper, we consider each edge $e = (u, v) \in E$ as two "half edges" e_u and e_v[1]. Let $\mathcal{E} \triangleq \{e_u, e_v \mid e = (u, v) \in E\}$ denote the set of all half edges. For every vertex $v \in V$, we use $\mathcal{E}(v)$ to denote the set of half edges incident to v.

An instance of a Holant problem is a tuple $\Lambda = \big(G(V, E), (f_v)_{v \in V}\big)$, where for every $v \in V$, $f_v : \{0, 1\}^{\mathcal{E}(v)} \to \mathbb{R}^+$ is a function, where $\mathbb{R}^+$ is the set of non-negative real numbers. For every assignment

[1] Here we consider "half edges" instead of 'edges' as usual, since our Markov chains work on these "half edges".

$\sigma \in \{0,1\}^{\mathcal{E}}$, we define the weight of σ as

$$w_\Lambda(\sigma) \triangleq \prod_{v \in V} f_v\left(\sigma \mid_{\mathcal{E}(v)}\right).$$

For every $\sigma \in \{0,1\}^{\mathcal{E}}$, we use $d(\sigma)$ to denote the number of edges $e = (u,v)$ such that $\sigma(e_u)$ and $\sigma(e_v)$ disagree, i.e., $d(\sigma) \triangleq |\{e = (u,v) \in E \mid \sigma(e_u) \neq \sigma(e_v)\}|$. For every $k \geq 0$, we denote $\Omega_k \triangleq \left\{\sigma \in \{0,1\}^{\mathcal{E}} \,\middle|\, d(\sigma) = k\right\}$ and $Z_k(\Lambda) \triangleq \sum_{\sigma \in \Omega_k} w_\Lambda(\sigma)$.

The set Ω_0 contains exactly all the assignments which are consistent at each edge. These are the ordinary assignments we usually studied and we call $Z(\Lambda) = Z_0(\Lambda)$ the *partition function* of Λ.

Symmetric Functions. A function $f : \{0,1\}^J \to \mathbb{R}^+$ is *symmetric*, if the value of the function only depends on the Hamming weight of its input. We use $|x| = \sum_{i \in J} x_i$ to denote the Hamming weight of x. Thus, for a symmetric function $f : \{0,1\}^J \to \mathbb{R}^+$ where $|J| = d$, we can write it as $f = [f_0, f_1, \ldots, f_d]$, where f_i is the value of f on inputs with Hamming weight i.

We define some special symmetric functions which will be used in this paper:

- $\mathbf{0}$ ($\mathbf{1}$): $f(x) = 0$ ($f(x) = 1$) for all $x \in \{0,1\}^J$.
- Even (Odd): $f(x) = 1$ if $|x|$ is even (odd). Otherwise, $f(x) = 0$.
- $= k$: $f(x) = 1$ if $|x| = k$. Otherwise, $f(x) = 0$.
- $\geq k$ ($\leq k$): $f(x) = 1$ if $|x| \geq k$ ($|x| \leq k$). Otherwise, $f(x) = 0$.
- $[a,b]$: $f(x) = 1$ if $a \leq |x| \leq b$. Otherwise, $f(x) = 0$.

When needed, we use a sub index to indicate the arity of a function. For example, Even_d and $(= k)_d$ is the Even and $= k$ function with arity d. If every function f_v is the function $(\leq 1)_{d_v}$, then the Holant problem $\Lambda = \left(G(V,E), (f_v)_{v \in V}\right)$ is the matching problem. Functions $\leq b$ are for b-matching problem and functions $\geq b$ are for b-edge-cover problem.

We introduce a few operations for functions. For two functions f and g with same arity, we use $f \cdot g$ to denote the entry wise product of the two functions. For example:

- $[a,b]_d \cdot \mathsf{Even}_d$: $f(x) = 1$ if $a \leq |x| \leq b$ and $|x|$ is even. Otherwise, $f(x) = 0$.

For a function $f : \{0,1\}^J \to \mathbb{R}^+$ and an assignment $\pi \in \{0,1\}^I$ where $I \subseteq J$, we define the *pinning* of f by π as a function $G : \{0,1\}^{J \setminus I} \to \mathbb{R}^+$ such that for every $\sigma \in \{0,1\}^{J \setminus I}$, $G(\sigma) = f(\sigma \circ \pi)$ where $\sigma \circ \pi$ is the concatenation of σ and π. For symmetric functions in symmetric notation $[f_0, f_1, \ldots, f_d]$, a pinning gets a consecutive sub-sequence of $\{f_0, f_1, \ldots, f_d\}$. The complement of a function $\overline{F}$ takes a complement for each input entry before evaluation of the function. For symmetric function, it simple reverses the order as $[f_d, f_{d-1}, \ldots, f_0]$.

Windable Functions. In [18], a special family of functions called *windable functions* has been introduced:

Definition 1. *For any finite set J and any configuration $x \in \{0,1\}^J$, define $\mathcal{M}_x$ to be the set of partitions of $\{i \mid x_i = 1\}$ into pairs and at most one singleton. A function $F : \{0,1\}^J \to \mathbb{R}^+$ is **windable** if there exist values $B(x,y,M) \geq 0$ for all $x, y \in \{0,1\}^J$ and all $M \in \mathcal{M}_{x \oplus y}$ satisfying:*

1. *$F(x)F(y) = \sum_{M \in \mathcal{M}_{x \oplus y}} B(x,y,M)$ for all $x, y \in \{0,1\}^J$, and*
2. *$B(x,y,M) = B(x \oplus S, y \oplus S, M)$ for all $x, y \in \{0,1\}^J$ and all $S \in M \in \mathcal{M}_{x \oplus y}$.*

Here $x \oplus S$ denotes the vector obtained by changing x_i to $1 - x_i$ for the one or two elements i in S. [2]

Observation 2. If $|x|$ is even, each $M \in \mathcal{M}_x$ contains no singleton. Otherwise, if $|x|$ is odd, each $M \in \mathcal{M}_x$ contains exactly one singleton.

The following nice theorem was implicitly proved in [18].

Theorem 3. *There exists an FPRAS to compute the partition function $Z(\Lambda)$ for instances $\Lambda = (G(V,E), (f_v)_{v \in V})$ with $|V| = n$, if it holds that (1) the instance is* self-reducible *in the sense of [15]; (2) for every $v \in V$, the function f_v is windable; and (3) $\frac{Z_2(\Lambda)}{Z_0(\Lambda)} = n^{O(1)}$.*

The FPRAS is obtained by the MCMC method. The states of the Markov chain are all the assignments in $\Omega_0 \cup \Omega_2$, which contains all the consistent assignments (Ω_0) and nearly consistent assignments (Ω_2). The second condition ensures that the size of Ω_0 and $\Omega_0 \cup \Omega_2$ are polynomial related. To prove the rapid mixing of the Markov chain, the windable condition is used to construct canonical paths. Roughly

[2]Note that our definition seems different from [18], which defines $\mathcal{M}_x$ to be the set of partitions of $\{i \mid x_i = 1\}$ into pairs and singletons. While by the proof of Lemma 15 in [18], both two definitions are equivalent to $F_\oplus$ being even-windable. Thus, our definition is equivalent to [18] in fact.

speaking, by the pairings and singletons in the definition of windable, the graph is naturally decomposed into disjoint union of paths and cycles. Then the canonical path just winds the edges follow these paths and cycles. The formal definition and detail can be found in [18]. For the convenience of the readers, we also include a formal description for the Markov chain and canonical paths in appendix. To logically follow the results of this paper, all these are not needed except the statement of the above theorem.

3 Windability for Symmetric Functions

In this section, we obtain a characterization for all symmetric windable functions. Before that, we introduce one more definition which is also adapted from [18].

Definition 4. *A function $H : \{0,1\}^J \to \mathbb{R}^+$ has a **2-decomposition** if there are values $D(x, M) \geq 0$, where x ranges over $\{0,1\}^J$ and M ranges over partitions of J into pairs and at most one singleton, such that:*

1. *$H(x) = \sum_M D(x, M)$ for all x, where the sum is over partitions of J into pairs and at most one singleton, and*
2. *$D(x, M) = D(x \oplus S, M)$ for all x, M and all $S \in M$.*

Our definition for 2-decomposition is a generalization of [18], since we allow the length of J to be odd. By the new definition, we have the following lemma.

Lemma 5. *A function F is windable, if and only if for all pinnings G of F, the function $G \cdot \overline{G}$ has a 2-decomposition.*

Proof. If F is windable, for each $I \subseteq J$ and each $\mathbf{p} \in \{0,1\}^I$, define $D_{\mathbf{p}}(x, M) = B((x, \mathbf{p}), (\overline{x}, \mathbf{p}), M)$ for all $x \in \{0,1\}^{J \setminus I}$. By definition 1 and 4, we have that $D_{\mathbf{p}}$ is a 2-decomposition of $G \cdot \overline{G}$, where G is the pinning of F by $\mathbf{p}$.

For the backwards direction, for all $x, y \in \{0,1\}^J$, let $I = \{i \in J \mid x_i = y_i\}$ be the position where x and y agrees. Let $\mathbf{p} \in \{0,1\}^I$ be the restriction of x to I, which is the same as the restriction of y to I. Let x' be the restriction of x to $J \setminus I$. Define $B(x, y, M) = D_{\mathbf{p}}(x', M)$. Then by the definitions, it can be verified that B witnesses that F is windable. $\square$

We introduce matrices $\mathbf{A}_m$ for every integer $m \geq 1$, which will be used in our characterization theorem.

- If $m = 2n$ is even, then $\mathbf{A}_m = (a_{ij})_{\substack{0 \leq i \leq n \\ 0 \leq j \leq n}} \in \mathcal{Q}^{(n+1)\times(n+1)}$ where
$$a_{ij} = \begin{cases} \binom{i}{j}\binom{2n-i}{j} j!(i-j-1)!!(2n-i-j-1)!! \\ \qquad \text{if } i \equiv j \pmod 2; \\ 0 \qquad \text{otherwise.} \end{cases}$$
- If $m = 2n+1$ is odd, then $\mathbf{A}_m = (a_{ij})_{\substack{0 \leq i \leq n \\ 0 \leq j \leq n}} \in \mathcal{Q}^{(n+1)\times(n+1)}$ where
$$a_{ij} = \begin{cases} \binom{i}{j}\binom{2n+1-i}{j} j!(i-j-1)!!(2n+1-i-j)!! \\ \qquad \text{if } i \equiv j \pmod 2; \\ \binom{i}{j}\binom{2n+1-i}{j} j!(i-j)!!(2n-i-j) \\ \qquad \text{otherwise.} \end{cases}$$

The notation $n!!$ is the double factorial of n. For even n, $n!! = n \cdot (n-2) \cdots 2$; and for odd n $n!! = n \cdot (n-2) \cdots 1$. If $n = 0$ or $n = -1$, then $n!! = 1$ by convention. We note that $\mathbf{A}_m$ is a lower triangular matrix (which follows from the convention that $\binom{i}{j} = 0$ for $i < j$). The entry a_{ij} of $\mathbf{A}_m$ has following combinatorial interpretation: Consider we have m balls consisting of i different red balls and $m - i$ different blue balls. If $m = 2n$ is even, then a_{ij} is the number of ways to divide $2n$ balls into n pairs, such that the number of pairs with different colors is j. If $m = 2n+1$ is odd, then a_{ij} is the number of ways to divide $2n+1$ balls into n pairs and a singleton, such that the number of pairs with different colors is j.

Lemma 6. *Let $m \geq 1$ be an integer, $n = \lfloor \frac{m}{2} \rfloor$ and $H = [h_0, h_1, \ldots, h_m]$ be a symmetric function with $h_i = h_{m-i}$ for all $i = 0, 1, \cdots, n$. Let $\mathbf{h} = [h_0, h_1, \ldots, h_n]$ be a vector. Then H is 2-decomposible if and only if there exists an $\mathbf{x} \in \mathbb{R}^{n+1} \geq \mathbf{0}$ such that $\mathbf{A}_m \mathbf{x} = \mathbf{h}$.*

We note that we abuse the notation $\mathbf{h} = [h_0, h_1, \ldots, h_n]$ both as a symmetric function with arity n and a vector in $\mathbb{R}^{n+1}$ in the whole paper when meaning is clear from the context.

Proof. we first consider the case that $m = 2n$ is even. Let $\mathcal{M}$ denote the set of all partitions of $[m]$ into pairs. We define an equivalent relation $\sim$ between pairs (x, M) where $x \in \{0,1\}^m$ and $M \in \mathcal{M}$. Given a pair (x, M), let $k(x, M) \triangleq |\{(x_i, x_j) \in M \mid x_i \neq x_j\}|$, i.e., the number of pairs in M with different value. Then two pairs $(x, M) \sim (x', M')$ if $k(x, M) = k(x', M')$, namely M and M' contain the same number of pairs with different value. This relation induces equivalent classes $\{\Delta_k \mid k = 0, \ldots, n\}$ where each $\Delta_k = \{(x, M) \mid k(x, M) = k\}$.

We claim that the function H is 2-decomposible if and only if for every $0 \le k \le n$, there exists $D_k \ge 0$ such that for every $x \in \{0,1\}^m$, $H(x) = \sum_{M\in\mathcal{M}} D_{k(x,M)}$.

"If" direction is easy. Let $D(x, M) = D_{k(x,M)}$, then the first requirement is satisfied naturally. The second requirement is satisfied by the fact that $k(x, M) = k(x \oplus S, M)$ for any x, M and $S \in M$.

Thus we now assume H is 2-decomposible, i.e, for every $x \in \{0,1\}^m$ and $M \in \mathcal{M}$, there exists $D(x, M) \ge 0$ such that

1. $H(x) = \sum_{M\in\mathcal{M}} D(x, M)$, and

2. $D(x, M) = D(x \oplus S, M)$ for every $S \in M$.

We need to show that there exists $D_k \ge 0$ such that for every $x \in \{0,1\}^m$, $H(x) = \sum_{M\in\mathcal{M}} D_{k(x,M)}$.

Let $\sigma \in S_m$ be a permutation on $[m]$. For every $x \in \{0,1\}^n$, we use x_σ to denote $(x_{\sigma(1)}, \ldots, x_{\sigma(m)})$ and for every $M \in \mathcal{M}$, we use M_σ to denote the partition on $[m]$ that $(x_i, x_j) \in M \iff (x_{\sigma(i)}, x_{\sigma(j)}) \in M_\sigma$. It is easy to see that for every $0 \le k \le n$ and $\sigma \in S_m$, $(x, M) \in \Delta_k \iff (x_\sigma, M_\sigma) \in \Delta_k$.

For every $k \ge 0$, we fix some $(x^{(k)}, M^{(k)}) \in \Delta_k$ and define $D_k = \frac{1}{m!}\sum_{\sigma\in S_m} D(x^{(k)}_\sigma, M^{(k)}_\sigma)$. An important fact is that the value of D_k is an invariant for different choice of $(x^{(k)}, M^{(k)}) \in \Delta_k$. To see this, consider two pairs $(x, M), (x', M') \in \Delta_k$ where $x = (x_1, x_2, \ldots, x_m)$ and $x' = (x'_1, x'_2, \ldots, x'_m)$, we aim to show that

$$\sum_{\sigma\in S_m} D(x_\sigma, M_\sigma) = \sum_{\sigma\in S_m} D(x'_\sigma, M'_\sigma). \tag{3.1}$$

We can assume without lost of generality that no pair $S = (x_i, x_j) \in M$ with $x_i = x_j = 1$ and no pair $S' = (x'_i, x'_j) \in M'$ with $x'_i = x'_j = 1$. This is because for every $S \in M$, the mapping $g((x_\sigma, M_\sigma)) = ((x \oplus S)_\sigma, M_\sigma)$ is a bijection between $\{(x_\sigma, M_\sigma) \mid \sigma \in S_m\}$ and $\{((x \oplus S)_\sigma, M_\sigma) \mid \sigma \in S_m\}$, and moreover $D(x_\sigma, M_\sigma) = D((x \oplus S)_\sigma, M_\sigma)$. Thus for every $S = (x_i, x_j) \in M$ with $x_i = x_j = 1$, the identity (3.1) is equivalent if we replace x by $x \oplus S$. The same argument holds for x'.

Under this assumption, we have $\sum_{i=1}^n x_i = \sum_{i=1}^n x'_i$ and both pairs belong to Δ_k. This implies for some permutation $\pi \in S_m$, it holds that $(x_\pi, M_\pi) = (x', M')$ and justify (3.1).

It remains to verify that for every $x \in \{0,1\}^m$, $H(x) = \sum_{M\in\mathcal{M}} D_{k(x,M)}$. Since $H(\cdot)$ is symmetric, we have

$$\begin{aligned} H(x) &= \frac{1}{m!}\sum_{\sigma\in S_m} H(x_\sigma) = \frac{1}{m!}\sum_{\sigma\in S_m}\sum_{M\in\mathcal{M}} D(x_\sigma, M) \\ &= \frac{1}{m!}\sum_{M\in\mathcal{M}}\sum_{\sigma\in S_m} D(x_\sigma, M_\sigma) \\ &= \frac{1}{m!}\sum_{k=0}^{n}\sum_{M\in\mathcal{M}:(x,M)\in\Delta_k}\sum_{\sigma\in S_m} D(x_\sigma, M_\sigma). \end{aligned}$$

It then follows from our discussion in the last paragraph that

$$H(x) = \sum_{k=0}^{n}\sum_{M\in\mathcal{M}:(x,M)\in\Delta_k} D_k = \sum_{M\in\mathcal{M}} D_{k(x,M)}.$$

Therefore, the function H is 2-decomposible if and only if there exist $D_k \ge 0$ for every $k = 0, 1, \ldots, n$ such that for every $x = (x_1, x_2, \ldots, x_m) \in \{0,1\}^m$,

$$\begin{aligned} H(x) &= \sum_{M\in\mathcal{M}} D_{k(x,M)} = \sum_{k=0}^{n}\sum_{M\in\mathcal{M}:k(x,M)=k} D_k \\ &= \sum_{k=0}^{n} |\{M \in \mathcal{M} \mid k(x, M) = k\}|\, D_k. \end{aligned} \tag{3.2}$$

Since $H(\cdot)$ is a symmetric function, for every $x, x' \in \{0,1\}^m$ with same Hamming weight, identity (3.2) are the same. Moreover, the identity (3.2) for x with Hamming weight i is the same as the identity (3.2) for x with Hamming weight $m - i$. For $i = |x|$, the identity (3.2) becomes

$$h_i = \sum_{k=0}^{n} |\{M \in \mathcal{M} \mid k(x, M) = k\}|\, D_k = \sum_{k=0}^{n} a_{ik} D_k,$$

where the second equality uses the (combinatorial) definition of a_{ik}. Therefore, these $D_k \ge 0$ are the solution of the linear system $\mathbf{A}_m\mathbf{x} = \mathbf{h}$ defined in the statement of the lemma. This completes the proof for the case that m is even.

Then we consider the case that $m = 2n + 1$ is odd. Let $\mathcal{M}$ denote the set of all partitions of $[m]$ into pairs and a singleton. The proof is similar to the case that m is even, with some slight difference on verifying (3.1), as we have to deal with the singleton in each $M \in \mathcal{M}$. We define an equivalent relation $\sim$ as that $(x, M) \sim (x', M')$ if $k(x, M) = k(x', M')$. This definition is the same as the $m = 2n$ case as the singleton plays no role. For every $k = 0, \ldots, n$, we also define $\Delta_k = \{(x, M) \mid k(x, M) = k\}$ and claim the the function H is 2-decomposible if and only if for every $0 \le k \le n$, there exists $D_k \ge 0$ such that for every $x \in \{0,1\}^m$, $H(x) = \sum_{M\in\mathcal{M}} D_{k(x,M)}$. The proof for the claim is almost identical as the even case. When verifying (3.1), we can assume no pair $(x_i, x_j) \in M$ with $x_i = x_j = 1$ and that the singleton $(x_i) \in M$ satisfies $x_i = 0$ (and the same assumption for (x', M')), then the remaining argument can go through. $\square$

Our characterization of the windability of symmetric functions is summarized by following theorem:

Theorem 7. *Given a symmetric function $F : \{0,1\}^d \to \mathbb{R}^+$, F is windable if and only if for every pinning G of F with arity m, the function $H(x) = [h_0, h_1, \ldots, h_m] \triangleq G(x)G(\bar{x})$ satisfies the following condition: The linear equations $\mathbf{A}_m\mathbf{x} = \mathbf{h}$ has a nonnegative solution $\mathbf{x} \geq 0$, where $\mathbf{h} = [h_0, h_1, \ldots h_{\lfloor \frac{m}{2} \rfloor}]$.*

We note that there exists an unique solution for $\mathbf{A}_m\mathbf{x} = \mathbf{h}$ as $\mathbf{A}_m$ is a lower triangular matrix. So we only need to check that this solution is nonnegative or not.

3.1 Properties of $\mathbf{A}_m$ In this subsection, we obtain some properties of the matrix $\mathbf{A}_m$ which are useful to verify that the linear equations $\mathbf{A}_m \cdot \mathbf{x} = \mathbf{h}$ has a nonnegative solution or not.

First of all, for all $i = 0, 1, \cdots, \lfloor \frac{m}{2} \rfloor$ we have

$$\sum_{0 \leq j \leq i} a_{ij} = (2\lfloor \frac{m-1}{2} \rfloor + 1)!! = a_{00}.$$

This has a simple combinatorial explanation since the sum is the total number of partitions of m different objects into pairs and at most one singleton. This implies the following lemma.

Lemma 8. *Let $m \geq 1$ and $c \geq 0$, $\mathbf{A}_m\mathbf{x} = c \cdot \mathbf{1}$ has a nonnegative solution $\mathbf{x} = \frac{c}{a_{00}} \cdot \mathbf{1}$.*

In the case that $m = 2n$ is even, the matrix $\mathbf{A}_m$ has non-zero entries a_{ij} only if $i \equiv j \pmod 2$. Thus the existence of nonnegative solution for the linear equations $\mathbf{A}_m\mathbf{x} = \mathbf{h}$ is equivalent to the existence of nonnegative solutions for the two linear equations $\mathbf{A}_m\mathbf{x} = \mathbf{h}_0$ and $\mathbf{A}_m\mathbf{x} = \mathbf{h}_1$, where $\mathbf{h}_0$ (resp. $\mathbf{h}_1$) is obtained from $\mathbf{h}$ by setting $h_i = 0$ for all odd (resp. even) i. This fact implies the following corollary:

Corollary 9. *Let $H(x) = G(x)G(\bar{x})$ be a symmetric function with arity $m = 2n$. Define functions H_0, H_1 as $H_0 = H \cdot \mathsf{Even}$ and $H_1 = H \cdot \mathsf{Odd}$. Then H is 2-decomposible if and only if both H_0 and H_1 are 2-decomposible.*

Combined with Lemma 8, we directly have the following lemma.

Lemma 10. *If $m = 2n$ is even, $\mathbf{A}_m\mathbf{x} = \mathbf{h}$ has a nonnegative solution if $\mathbf{h} = \mathsf{Even}$ or Odd.*

The following lemma reveals an relation between $\mathbf{A}_{2n}$ and $\mathbf{A}_{2n-1}$.

Lemma 11. *Assume $n \geq 1$. Let $\mathbf{A}_{2n} = (a_{ij}) \in \mathbb{R}^{(n+1)\times(n+1)}$, and $\mathbf{A}_{2n-1} = (a'_{ij}) \in \mathbb{R}^{n \times n}$. If $0 \leq i \leq n$ and $i \equiv j \pmod 2$, we have the following equality:*

$$a_{ij} = a'_{i,j-1} + a'_{ij} = a'_{i-1,j-1} + a'_{i-1,j}.^3 \tag{3.3}$$

Moreover, given two vectors $\mathbf{h} \in \mathbb{R}^{(n+1)\times(n+1)}$ and $\mathbf{h}' = \mathbb{R}^{n \times n}$, we have the following two properties:

1. *If $\mathbf{h}$ is odd (all even entries of $\mathbf{h}$ are 0), and $h'_{2i} = h'_{2i+1} = h_{2i+1}$ satisfies for $0 \leq i \leq \lfloor n/2 \rfloor$.[4] Then $\mathbf{A}_{2n-1} \cdot \mathbf{x}' = \mathbf{h}'$ has a nonnegative solution if and only if $\mathbf{A}_{2n} \cdot \mathbf{x} = \mathbf{h}$ has a nonnegative solution.*

2. *If $\mathbf{h}$ is even (all odd entries of $\mathbf{h}$ are 0), and $h'_{2i-1} = h'_{2i} = h_{2i}$ satisfies for $0 \leq i \leq n/2$.[5] Then $\mathbf{A}_{2n-1} \cdot \mathbf{x}' = \mathbf{h}'$ has a nonnegative solution if and only if $\mathbf{A}_{2n} \cdot \mathbf{x} = \mathbf{h}$ has a nonnegative solution.*

Proof. We first prove Equality 3.3. In fact, it is not hard to verify it by definition. Here we give a combinatorial explanation. Recall that a_{ij} is the number of matchings in Δ_j when $\sum_{k \in [2n]} x_k = i$ $(0 \leq i \leq n)$. If $i \equiv j \pmod 2$ and $i < n$, there must exist an entry of value 0. Assume that $x_{2n} = 0$ without loss of generality. Then the matching among the remaining entries should be in either Δ_{j-1} or Δ_j, and $\sum_{k \in [2n-1]} x_k = i$. Thus, we have $a_{ij} = a'_{i,j-1} + a'_{ij}$. Similarly, if $i \equiv j \pmod 2$ and $i > 0$, there must exist an entry of value 1. We let $x_{2n} = 1$ without loss of generality. Then the matching among the remaining entries should be in either Δ_{j-1} or Δ_j, and $\sum_{k \in [2n-1]} x_k = i - 1$. In this case, we have that $a_{ij} = a'_{i-1,j-1} + a'_{i-1,j}$. Combine these two equalities, we prove Equality 3.3.

If $\mathbf{h}$ is odd, suppose $\mathbf{x}$ is the solution for the linear equations $\mathbf{A}_{2n} \cdot \mathbf{x} = \mathbf{h}$. Observe that $\mathbf{x}$ is also odd by the definition of $\mathbf{A}_{2n}$. Let $x'_{2i} = x'_{2i+1} = x_{2i+1}$ for $0 \leq i \leq \lfloor n/2 \rfloor$. We show that this $\mathbf{x}'$ is exactly the solution of $\mathbf{A}_{2n-1} \cdot \mathbf{x}' = \mathbf{h}'$. Then by the construction of $\mathbf{x}'$, we know that $\mathbf{x}$ is nonnegative if and only if $\mathbf{x}'$ is nonnegative, which completes the proof. Consider the $(2i)$th row $(0 \leq i \leq \lfloor \frac{n-1}{2} \rfloor)$ and $(2i+1)$th row

[3] If $i = 0$, the equality is $a_{00} = a'_{00}$. If $i = n$, the equality is $a_{nj} = a'_{n-1,j-1} + a'_{n-1,j}$.

[4] Since h'_{n+1} does not exist, if n is even and $i = \lfloor n/2 \rfloor$, the condition is $h'_n = h_{n+1}$.

[5] If $i = 0$, the condition is $h'_0 = h_0$.

of $\mathbf{A}_{2n-1}$ ($0 \le i \le \lfloor \frac{n-2}{2} \rfloor$), we have the following equalities which shows that $\mathbf{A}_{2n-1} \cdot \mathbf{x}' = \mathbf{h}'$.

$$\begin{aligned}\sum_{0\le j\le 2i} a'_{2i,j}x'_j &= \sum_{0\le j\le i} a'_{2i,2j}x'_{2j} + a'_{2i,2j+1}x'_{2j+1}\\ &= \sum_{0\le j\le i} a'_{2i,2j}x_{2j+1} + a'_{2i,2j+1}x_{2j+1}\\ &= \sum_{0\le j\le i} (a'_{2i,2j} + a'_{2i,2j+1})x_{2j+1}\\ &= \sum_{0\le j\le i} a_{2i+1,2j+1}x_{2j+1} = h_{2j+1} = h'_{2j},\end{aligned}$$

$$\begin{aligned}&\sum_{0\le j\le 2i+1} a'_{2i+1,j}x'_j\\ &= \sum_{0\le j\le i} a'_{2i+1,2j}x'_{2j} + a'_{2i+1,2j+1}x'_{2j+1}\\ &= \sum_{0\le j\le i} a'_{2i+1,2j}x_{2j+1} + a'_{2i+1,2j+1}x_{2j+1}\\ &= \sum_{0\le j\le i} (a'_{2i+1,2j} + a'_{2i+1,2j+1})x_{2j+1}\\ &= \sum_{0\le j\le i} a_{2i+1,2j+1}x_{2j+1} = h_{2j+1} = h'_{2j+1}.\end{aligned}$$

If $\mathbf{h}$ is even, suppose $\mathbf{x}$ is the solution for the linear equations $\mathbf{A}_{2n} \cdot \mathbf{x} = \mathbf{h}$. Observe that $\mathbf{x}$ is also even. Let $x'_{2i-1} = x'_{2i} = x_{2i}$ for $0 \le i \le \lfloor n/2 \rfloor$. By the same argument as above, this $\mathbf{x}'$ is exactly the solution of $\mathbf{A}_{2n-1} \cdot \mathbf{x}' = \mathbf{h}'$. So we prove the whole lemma. $\square$

4 Counting b-Edge-Covers

In this section, we obtain **FPRAS** for counting b-edge-cover for $b \le 2$ as an application of our characterization. By Theorem 3, we need to prove that the function $\ge b$ is windable for $b \le 2$, and bound the ratio of Z_2/Z_0.

Lemma 12. *If $b \le 2$, the weight functions $\ge b$ are windable.*

Lemma 13. *For any counting b-edge-cover instance, we have that $Z_2/Z_0 \le 4n^2$, where n is the number of edges.*

We first prove Lemma 12. Consider the pinning function G of $\ge b$. Since $b \le 2$, G might be $\mathbf{1}_m$, $(\ge 1)_m$ or $(\ge 2)_m$. Let $H(x) = [h_0, h_1, \dots, h_m] \triangleq G(x)G(\bar{x})$, and let $\mathbf{h} = [h_0, h_1, \dots h_{\lfloor \frac{m}{2} \rfloor}]$. By the definition, we know that $\mathbf{h}$ can only be $\mathbf{1}_{\lfloor \frac{m}{2} \rfloor}$, $(\ge 1)_{\lfloor \frac{m}{2} \rfloor}$ or $(\ge 2)_{\lfloor \frac{m}{2} \rfloor}$. Then by Theorem 7, we need to show that $\mathbf{A}_m\mathbf{x} = \mathbf{h}$ always has a nonnegative solution. Thus, we only need to prove the following lemma.

Lemma 14. *If $m \ge 1$ and $b \le 2$, $\mathbf{A}_m\mathbf{x} = (\ge b)$ has a nonnegative solution.*

Proof. If $b = 0$, $\mathbf{h} = \mathbf{1}_n$ has been proved in Lemma 8. We assume $b = 1, 2$ in the following. We consider two different cases: m is even and m is odd.

The first case is that $m = 2n$ is even. If $b = 1$, $\mathbf{h} = (\ge 1)_n$. By Corollary 9, we only need to prove that both $\mathbf{A}_m\mathbf{x} = \mathbf{h}_0$ and $\mathbf{A}_m\mathbf{x} = \mathbf{h}_1$ have nonnegative solutions. Observe that $\mathbf{h}_1 = \mathsf{Odd}_n$. By Lemma 10, we only need to consider $\mathbf{h}_0 = (\ge 2)_n \cdot \mathsf{Even}_n$. Let $x_{2j} = \left(1 - (-1)^j \frac{(2j-1)!!}{\prod_{i=1}^{j}(2n-2i)}\right) \frac{1}{(2n-1)!!}$ if $0 \le j \le \lfloor \frac{n}{2} \rfloor$ and $x_j = 0$ if otherwise. Note that $x_0 = 1 - \frac{(-1)!!}{1} = 0$. If $j > 0$, the numerator $(2j-1)!!$ is no larger than the denominator $\prod_{i=1}^{j}(2n-2i)$ because we have $2j - 1 \le n - 1 \le 2n - 2$. So $x_{2j} \ge 0$ always holds. Thus, we prove that $\mathbf{x}$ is a nonnegative vector.

The remaining task is to show $\mathbf{x}$ is the solution. We note that $x_0 = 0$ and thus the first equation is satisfied. For i is odd, it is easy to see that $\sum_{0\le j\le i} a_{ij}x_j = 0$ since we have $a_{ij} = 0$ for even j and $x_j = 0$ for odd j. In the following, we only need to verify that $\sum_{0\le j\le i} a_{ij}x_j = 1$ for even $i = 2k \ge 2$. For these, we have

$$\begin{aligned}&\sum_{0\le j\le k} a_{2k,2j}x_{2j}\\ &= \sum_{0\le j\le k} a_{2k,2j}\left(\frac{1}{(2n-1)!!} + x_{2j} - \frac{1}{(2n-1)!!}\right)\\ &\overset{(\heartsuit)}{=} 1 + \sum_{0\le j\le k} a_{2k,2j}\left(x_{2j} - \frac{1}{(2n-1)!!}\right)\\ &= 1 + \sum_{0\le j\le k} \binom{2k}{2j}\binom{2n-2k}{2j}(2j)!(2k-2j-1)!!\\ &\qquad \cdot (2n-2k-2j-1)!!\left(x_{2j} - \frac{1}{(2n-1)!!}\right)\\ &= 1 - \sum_{0\le j\le k} (-1)^j \frac{(2k)!(2n-2k)!(n-j-1)!}{2(k-j)!(n-k-j)!j!(2n-1)!}\\ &= 1 - \frac{(2k)!(2n-2k)!}{2(2n-1)!} \sum_{0\le j\le k} (-1)^j \frac{(n-j-1)!}{(k-j)!(n-k-j)!j!}\\ &\overset{(\diamond)}{=} 1 - \frac{(2k)!(2n-2k)!}{2(2n-1)!} \sum_{0\le j\le k} \frac{(-1)^j\binom{k}{j}\binom{n-j}{k}}{n-j}\\ &= 1 - 0 = 1,\end{aligned}$$

where $(\heartsuit)$ is because the sum of entries in each row of $\mathbf{A}_m$ is $(2n-1)!!$, which equals to the total number of partitions. The equality $(\diamond)$ uses the fact $\sum_{0\le j\le k} \frac{(-1)^j\binom{k}{j}\binom{n-j}{k}}{n-j} = 0$, which is by the following technical Lemma.

Lemma 15. $\sum_{j=0}^{m} \frac{(-1)^j \binom{m}{j}\binom{n-j}{m}}{n-j} = 0.$

Proof. Consider $f(x) = \sum_{j=0}^{m} \frac{(-1)^j \binom{m}{j}\binom{n-j}{m} x^j}{n-j}$. It is not hard to see that

$$f(x) = \frac{\binom{n}{m} {}_2F_1(-m, m-n; 1-n; x)}{n},$$

where ${}_2F_1(a, b; c; z) = \sum_{i\geq 0} \frac{(a)_i (b)_i}{(c)_i} \cdot \frac{z^i}{i!}$. Here $(a)_i = \prod_{j=0}^{i-1}(a+j)$.

By Equality 15.3.3 in [1], ${}_2F_1(-m, m-n; 1-n; x) = (1-x) \cdot {}_2F_1(1+m-n, 1-m; 1-n; x)$. Let $x = 1$, we prove the lemma. □

If $b = 2$, $\mathbf{h} = (\geq 2)_n$. We still consider the linear equations $\mathbf{A}_m\mathbf{x} = \mathbf{h}_0$ and $\mathbf{A}_m\mathbf{x} = \mathbf{h}_1$. Note that $\mathbf{h}_0 = (\geq 2)_n \cdot \mathsf{Even}_n$ which has been proved in the last case. So we focus on $\mathbf{h}_1$ which equals to $(\geq 3)_n \cdot \mathsf{Odd}_n$. Let $x_{2j+1} = (1-(-1)^j \frac{(2j+1)!!}{\prod_{i=2}^{j+1}(2n-2i)}) \frac{1}{(2n-1)!!}$ $(0 \leq j \leq \frac{n-1}{2})$. Otherwise let $x_j = 0$. Then we show the correctness.

If $j = 0$, note that $x_1 = 0$. If $j = 1$, it is not hard to see that $x_3 > 0$. If $j > 1$, we have $n \geq 5$. Observe that the numerator $(2j+1)!!$ is no larger than the denominator $\prod_{i=2}^{j+1}(2n-2i)$ since we have $2j+1 \leq 2n-4$. So $x_{2j+1} \geq 0$ always holds. Thus, we prove that $\mathbf{x}$ is a non-negative vector.

The remaining task is to show that $\mathbf{x}$ is exactly the solution. We note that $x_1 = 0$ and thus the second equation is satisfied. For i is even, it is easy to see that $\sum_{0\leq j\leq i} a_{ij}x_j = 0$ since we have $a_{ij} = 0$ for odd j and $x_j = 0$ for even j. In the following, we only need to consider the $(2k+1)$th rows $(0 \leq k \leq \lfloor \frac{n-1}{2} \rfloor)$. In fact, we have the following equalities.

$$\begin{aligned}
&\sum_{0\leq j\leq k} a_{2k+1,2j+1}x_{2j+1} \\
&= \sum_{0\leq j\leq k} a_{2k+1,2j+1}\left(\frac{1}{(2n-1)!!} + x_{2j+1} - \frac{1}{(2n-1)!!}\right) \\
&\overset{(\heartsuit)}{=} 1 + \sum_{0\leq j\leq k} a_{2k+1,2j+1}\left(x_{2j+1} - \frac{1}{(2n-1)!!}\right) \\
&= 1 + \sum_{0\leq j\leq k} \binom{2k+1}{2j+1}\binom{2n-2k-1}{2j+1} \\
&\qquad \cdot (2j+1)!(2k-2j-1)!!(2n-2k-2j-3)!! \\
&\qquad \cdot \left(x_{2j+1} - \frac{1}{(2n-1)!!}\right) \\
&= 1 - \frac{(2k+1)!(2n-2k-1)!(n-1)}{(2n-1)!} \\
&\qquad \cdot \sum_{0\leq j\leq k} (-1)^j \frac{(n-j-2)!}{(k-j)!(n-k-j-1)!j!} \\
&\overset{(\diamondsuit)}{=} 1 - \frac{(2k+1)!(2n-2k-1)!(n-1)}{(2n-1)!} \\
&\qquad \cdot \sum_{0\leq j\leq k} \frac{(-1)^j\binom{k}{j}\binom{n-1-j}{k}}{n-1-j} \\
&= 1 - 0 = 1.
\end{aligned}$$

where $(\heartsuit)$ is because the sum of entries in each row of $\mathbf{A}_m$ is $(2n-1)!!$ which equals to the total number of partitions, and $(\diamondsuit)$ uses Lemma 15.

If $m = 2n-1$ is odd. We want to show that $\mathbf{A}_{2n-1} \cdot \mathbf{x} = (\geq b)_{n-1}$ has a nonnegative solution for $b \leq 2$. If $b = 1$, by Lemma 11, we only need to show that $\mathbf{A}_{2n} \cdot \mathbf{x} = (\geq 1)_n \cdot \mathsf{Even}_n = (\geq 2)_n \cdot \mathsf{Even}_n$ has a non-negative solution, which has been proved in the first case. Finally, if $b = 2$, by Lemma 11, we only need to prove that $\mathbf{A}_{2n} \cdot \mathbf{x} = (\geq 2)_n \cdot \mathsf{Odd}$ has a nonnegative solution. Note that $(\geq 2)_n \cdot \mathsf{Odd} = (\geq 3)_n \cdot \mathsf{Odd}_n$. By the first case, we finish the proof.

Thus, we prove that $\mathbf{A}_m\mathbf{x} = (\geq b)$ always has a nonnegative solution if $b \leq 2$. □

The second part is to prove Lemma 13.

Proof. We construct a mapping from Ω_2 to Ω_0 to bound Z_2/Z_0. For any satisfying assignment $x \in \{0,1\}^{2n}$ in Ω_2, assume that i, j are the two half edges which violates the equality constraint on edges, and $x_i = x_j = 0$ (the corresponding other two half edges are assigned 1). Let y be the assignment obtained by x flipping on ith and jth entries. Note that $y \in \Omega_0$ is also a satisfying assignment by the definition of b-edge-cover.

On the other hand, from a satisfying assignment $y \in \Omega_0$, we can construct at most $4n^2$ satisfying assignments $x \in \Omega_2$ by flipping on two half edges. So we map at most $4n^2$ satisfying assignments $x \in \Omega_2$ to y. Thus, we have $Z_2/Z_0 \le 4n^2$ by this mapping. □

Combining Lemma 12 and 13, we have the following theorem.

Theorem 16. *There is an **FPRAS** for counting b-edge-cover problems if $b \le 2$.*

5 Counting b-Matchings

In this section, we provide another application for counting b-matchings.

Theorem 17. *There is an **FPRAS** for counting b-matching problems if $b \le 7$.*

Similarly, by Theorem 3, we only need to prove the following two lemmas.

Lemma 18. *If $b \le 7$, the weight functions $\le b$ are windable.*

Lemma 19. *For any counting b-matching instance, we have that $Z_2/Z_0 \le 4n^2$, where n is the number of edges.*

For preparation, we show the following lemma first.

Lemma 20. *Let $n = \lfloor \frac{m}{2} \rfloor$. Then $\mathbf{A}_m \mathbf{x} = (= n)_n$ has a nonnegative solution.*

Proof. Since the RHS only has one non-zero entry at the last row, it is easy to see that $x_n = \frac{1}{a_{nn}}$ and $x_i = 0$ for $i = 0, 1, \cdots, n-1$ is a non-negative solution. □

Now we are ready to prove Lemma 18.

Proof. (Lemma 18) Consider the pinning function G of $\le b$. We have that $G = (\le k)_m$, where $k \le 7$. Recall that we define $H(x) = [h_0, h_1, \ldots, h_m] \triangleq G(x)G(\bar{x})$. Then we have $H = [m-k, k]_m$. To make H non-trivial, we need $k \le m \le 2k$. Let $\mathbf{h} = [h_0, h_1, \ldots h_{\lfloor \frac{m}{2} \rfloor}]$, then $h = (\ge m-k)_{\lfloor \frac{m}{2} \rfloor}$. If $m \le k+2$, then $\mathbf{h} = (\ge l)$ with $l \le 2$ which has been proved by Lemma 14. By Lemma 20, the cases that $m = 2k$ and $m = 2k-1$ are also correct. So we only need to consider the cases that $k+3 \le n \le 2k-2$ and $k \le 7$. We enumerate all of them in the following
Case $k = 5, m = 8$. $\mathbf{x} = (0, 0, 0, \frac{1}{60}, \frac{1}{24})$ is the non-negative solution.
Case $k = 6, m = 9$. $\mathbf{x} = (0, 0, 0, \frac{1}{360}, \frac{1}{360})$ is the non-negative solution.
Case $k = 6, m = 10$. $\mathbf{x} = (0, 0, 0, 0, \frac{1}{360}, \frac{1}{120})$ is the non-negative solution.
Case $k = 7, m = 10$. $\mathbf{x} = (0, 0, 0, \frac{1}{630}, \frac{1}{360}, \frac{1}{2520})$ is the non-negative solution.
Case $k = 7, m = 11$. $\mathbf{x} = (0, 0, 0, 0, \frac{1}{2520}, \frac{1}{2520})$ is the non-negative solution.
Case $k = 7, m = 12$. $\mathbf{x} = (0, 0, 0, 0, 0, \frac{1}{2520}, \frac{1}{720})$ is the non-negative solution.
□

The remaining task is to prove Lemma 19.

Proof. (Lemma 19) The argument is almost the same as Lemma 13 except that from a satisfying assignment $x \in \Omega_2$, we map it to a satisfying assignment $y \in \Omega_0$ by deleting two half edges, instead of adding two half edges. Again, we construct a mapping from Ω_2 to Ω_0, and show that $Z_2/Z_0 \le 4n^2$. □

Remark: Our FPRAS for both b-matchings and b-edge-covers can be extended to instances with edge weights. On the other hand, the results cannot be extended to counting 8-matchings or 3-edge-covers since these constraint functions are not windable. These facts are also showed by our characterization theorem and we present them in the following two sections.

6 Edge Weighted b-Edge-Covers and b-Matchings

In this section, we consider the version that each edge $e \in E$ has a nonnegative weight w_e. We want to show that both counting weighted b-edge-cover and b-matching problems have an **FPRAS**.

Given a graph $G = (V, E)$. The trick is to add a constraint on each edge. For each edge e, we separate it into two edges e^0 and e^1. Between e^0 and e^1, we add a new constraint $(1, 0, w_e)$. Now we construct a new graph $G' = (V \cup E, E^0 \cup E^1)$. It is easy to see that the partition function for this new Holant instance is exactly the partition function for the edge weighted counting problem.

We first prove the constraint for each edge is windable.

Lemma 21. *If $a \ge 0$, the function $(1, 0, a)$ is windable.*

Proof. For all pinnings G of this function, we can observe that $G\overline{G}$ is either $\mathbf{0}$ or $c \cdot \mathbf{1}$, where c is some nonnegative constant. By Lemma 5 and 8, we prove the lemma . □

Compared to the unweighted version, we have $|E|$ more constraints on edges. Note that the half edges are between vertex constraints and edge constraints. In other words, each edge $e \in E$ is partitioned into four half edges. It only needs to show that Z_2/Z_0 is still bounded. We first consider the weighted b-edge-cover problems.

Lemma 22. *For any counting b-edge-cover instance where $b \le 2$, we have that $Z_2/Z_0 \le \frac{16n^2}{\min w_e^2}$.*[6] *Here, n is the number of edges.*

Proof. Similar to Lemma 13, we construct a mapping from Ω_2 to Ω_0. Since the half edges are different, the rules for the mapping are also different.

Consider a satisfying assignment $x \in \{0,1\}^{2n}$ in Ω_2, exactly two pair of half edges disagree with each other. We call them 'bad' pairs. For an edge e, we partition it into four different half edges. If there exists a 'bad' pair of half edges on e, there might be exactly one, two or three half edges of value 1. We call this edge a 'bad' edge. Note that there are at most two such 'bad' edges. Assume they are e_1 and e_2. Let y be the assignment obtained by x fixing all half edges to be 1 on e_1 and e_2. Note that $y \in \Omega_0$ is also a satisfying assignment by the definition of b-edge-cover. Moreover, $F(x)/F(y) \le \frac{1}{\min_e w_e^2}$.

On the other hand, from a satisfying assignment $y \in \Omega_0$, we can construct at most $16n^2$ satisfying configurations $x \in \Omega_2$ by flipping on two random half edges. Note that for each such x, we also have $F(x)/F(y) \le \frac{1}{\min_e w_e^2}$. Moreover, we map at most $16n^2$ satisfying configurations $x \in \Omega_2$ to y. Thus, by this mapping, we have that $Z_2/Z_0 \le \frac{16n^2}{\min_e w_e^2}$. □

Theorem 23. *There is an **FPRAS** for counting weighted b-edge-cover problems if $b \le 2$.*

For counting b-matching problems, we have similar results.

Lemma 24. *For any counting weighted b-matching instance where $b \le 7$, we have that $Z_2/Z_0 \le 16n^2 \max_e w_e^2$.*[7] *Here, n is the number of edges.*

Proof. The proof is very similar to Lemma 13, except that from a satisfying assignment $x \in \Omega_2$, we map x to an assignment $y \in \Omega_0$ by fixing all half edges to be 1 instead of 0 on "bad" edges. Another difference is that we have $F(x)/F(y) \le \max_e w_e^2$. □

Combined with Theorem 3, we have the following theorem.

Theorem 25. *There is an **FPRAS** for counting weighted b-matching problems if $b \le 7$.*

Remark: Observe the weight function $H = (1, 0, w_e)$. Note that the even entries of H is a geometric sequence. In general, we have the following lemma.

Lemma 26. *For a symmetric function $H : \{0,1\}^J \to \mathbb{R}^+$, if both the even and the odd subsequences are geometric sequences, then H is a windable function.*

Proof. We still focus on showing that for each pinning $G : \{0,1\}^m \to \mathbb{R}^+$ of H, $\mathbf{A}_m\mathbf{x} = \mathbf{h}$ has a nonnegative solution by Theorem 7, where $\mathbf{h}$ is the prefix of $G \cdot \overline{G}$.

If m is odd, by the property of geometric sequences, we observe that $h = c \cdot \mathbf{1}$ $(c > 0)$. If $m = 2n$ is even, by Corollary 9, we only need to show that both $\mathbf{A}_m\mathbf{x} = \mathbf{h}_0$ and $\mathbf{A}_m\mathbf{x} = \mathbf{h}_1$ have a nonnegative solution. By the property of geometric sequences, it is not hard to see that $h_0 = c_1 \cdot \mathsf{Even}_n$ and $h_1 = c_2 \cdot \mathsf{Odd}_n$ $(c_1, c_2 > 0)$. By Lemma 8 and 10, we prove that H is windable. □

By Lemma 26, we can show that FPRAS exists for this class of symmetric functions similar to B-matching. Note that $[1, \mu, 1, \mu, \ldots]$ is a special case, which has a well-known FPRAS in [12]. So we give an FPRAS for a more general class of counting problems.

7 Unwindable Functions

In this section, we give some examples of unwindable functions, which shows that our approach cannot be directly extended to 3-edge-cover and 8-matching problems.

Lemma 27. *If $b \ge 3$ and $|J| \ge b + 8$, the weight functions $(\ge b)_J$ are not windable.*

Proof. If $b \ge 3$ and $|J| \ge b + 8$, there must be a pinning G by $\mathbf{p}$, where $G = (\ge 3)_{11}$. By Theorem 7, we only need to show that $\mathbf{A}_{11} \cdot \mathbf{x} = (\ge 3)_6$ has nonpositive solution. In fact, we know that $\mathbf{x} = (0, 0, 0, \frac{1}{5040}, \frac{1}{5040}, -\frac{1}{10080})$ by calculation. □

Lemma 27 shows that why our technique can not work for arbitrary b-edge-covers. By this lemma, we can conclude the following corollary which shows that why winding technique does not work for arbitrary b-matchings.

Corollary 28. *If $b \ge 8$ and $|J| \ge b + 3$, the weight functions $(\le b)_J$ are not windable.*

Proof. For a weight function $F = (\le b)_J$, let $F' = (\ge |J| - b)_J$. Consider a pinning G of F by $\mathbf{p}$. We construct another pinning G' of F' by $\overline{\mathbf{p}}$. Note that

[6] We assume $\min_e w_e$ is a constant. This assumption is reasonable. Since if $\min_e w_e$ is exponentially small, counting weighted b-edge-cover problem can be as hard as minimal edge-cover problem.

[7] In this paper, we assume $\max_e w_e$ is a constant. This assumption is reasonable. Since if $\max_e w_e$ is exponentially large, counting weighted b-matching problem can be as hard as counting perfect matching.

for any x, we have that $G(x) = G'(\overline{x}) = \overline{G'}(x)$. Then $G \cdot \overline{G}$ is exactly the same as $G'\overline{G'}$. So F is windable if and only if F' is windable.
Note that $|J| - b \geq 3$ and $|J| \geq |J| - b + 8$. By Lemma 27, we prove the corollary. □

References

[1] Milton Abramowitz, Irene A Stegun, et al. Handbook of math. functions. *US Gov't. Printing Off., Washington, DC*, 1965.

[2] Ivona Bezáková and William Rummler. Sampling edge covers in 3-regular graphs. In *Mathematical Foundations of Computer Science 2009*, pages 137–148. Springer, 2009.

[3] R. Bubley and M. Dyer. Path coupling: A technique for proving rapid mixing in Markov chains. In *Proceedings of the 38th Annual Symposium on Foundations of Computer Science*, pages 223–231. IEEE, 1997.

[4] Mary Cryan, Martin Dyer, Leslie Ann Goldberg, Mark Jerrum, and Russell Martin. Rapidly mixing markov chains for sampling contingency tables with a constant number of rows. *SIAM Journal on Computing*, 36(1):247–278, 2006.

[5] Persi Diaconis and Daniel Stroock. Geometric bounds for eigenvalues of markov chains. *The Annals of Applied Probability*, pages 36–61, 1991.

[6] Martin Dyer, Alan Frieze, and Ravi Kannan. A random polynomial-time algorithm for approximating the volume of convex bodies. *Journal of the ACM (JACM)*, 38(1):1–17, 1991.

[7] Martin Dyer and Catherine Greenhill. On Markov chains for independent sets. *Journal of Algorithms*, 35(1):17–49, 2000.

[8] Martin Dyer, Mark Jerrum, and Eric Vigoda. Rapidly mixing Markov chains for dismantleable constraint graphs. In *Randomization and Approximation Techniques in Computer Science*, pages 68–77. Springer, 2002.

[9] Leslie Ann Goldberg and Mark Jerrum. A polynomial-time algorithm for estimating the partition function of the ferromagnetic Ising model on a regular matroid. In *Proceedings of the 39th International Colloquium Conference on Automata, Languages and Programming*, pages 521–532, 2011.

[10] Mark Jerrum. A very simple algorithm for estimating the number of k-colorings of a low-degree graph. *Random Structures & Algorithms*, 7(2):157–166, 1995.

[11] Mark Jerrum and Alistair Sinclair. Approximating the permanent. *SIAM journal on computing*, 18(6):1149–1178, 1989.

[12] Mark Jerrum and Alistair Sinclair. Polynomial-time approximation algorithms for the Ising model. *SIAM Journal on computing*, 22(5):1087–1116, 1993.

[13] Mark Jerrum and Alistair Sinclair. The Markov chain Monte Carlo method: an approach to approximate counting and integration. *Approximation algorithms for NP-hard problems*, pages 482–520, 1996.

[14] Mark Jerrum, Alistair Sinclair, and Eric Vigoda. A polynomial-time approximation algorithm for the permanent of a matrix with nonnegative entries. *Journal of the ACM*, 51:671–697, July 2004.

[15] Mark Jerrum, Leslie Valiant, and Vijay Vazirani. Random generation of combinatorial structures from a uniform distribution. *Theoretical Computer Science*, 43:169–188, July 1986.

[16] Chengyu Lin, Jingcheng Liu, and Pinyan Lu. A simple FPTAS for counting edge covers. In *Proceedings of the 25th Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 341–348, 2014.

[17] Jingcheng Liu, Pinyan Lu, and Chihao Zhang. FPTAS for counting weighted edge covers. In *Proceedings of the 22nd Annual European Symposium on Algorithms*, pages 654–665, 2014.

[18] Colin McQuillan. Approximating holant problems by winding. *arXiv preprint arXiv:1301.2880*, 2013.

[19] Ben Morris and Alistair Sinclair. Random walks on truncated cubes and sampling 0-1 knapsack solutions. In *Proceedings of the 40th IEEE Symposium on Foundations of Computer Science*, pages 230–240. IEEE, 2002.

[20] Alistair Sinclair. Improved bounds for mixing rates of markov chains and multicommodity flow. *Combinatorics, probability and Computing*, 1(04):351–370, 1992.

[21] Eric Vigoda. Improved bounds for sampling coloring. In *Proceedings of the 37th IEEE Symposium on Foundations of Computer Science*, pages 51–59. IEEE, 1999.

Appendix

To be self-contained and for the convenience of readers, we include a formal proof for Theorem 3 in this appendix. These proofs are essentially adapted from [18].

We first construct a Markov chain to sample from $\Omega_0 \cup \Omega_2$.

Let $\Lambda = (G(V, E), (f_v)_{v \in V})$ be an instance with $|V| = n$ and every f_v is windable. Let $\mathcal{E}$ be the set of half edges in G. The state space of the chain is $\Omega = \Omega_0 \cup \Omega_2$. For every two configuration $\sigma, \pi \in \Omega$, the transition probability $P'(\sigma, \pi)$ is defined as

$$P'(\sigma,\pi) = \begin{cases} \frac{2}{n^2} \min\left(1, \frac{w_\Lambda(\pi)}{w_\Lambda(\sigma)}\right), \\ \qquad \text{if } d(\sigma,\pi) = 2; \\ 1 - \frac{2}{n^2} \sum_{\rho: d(\sigma,\rho)=2} \min\left(1, \frac{w_\Lambda(\rho)}{w_\Lambda(\sigma)}\right), \\ \qquad \text{if } \sigma = \pi; \\ 0, \quad \text{otherwise}, \end{cases}$$

where $d(\sigma, \pi)$ denote the Hamming distance between

σ and π.

Our Markov chain is the lazy version of above, i.e., for every two configurations $\sigma, \pi \in \Omega$, define $P(\sigma,\pi) = \frac{1+P'(\sigma,\pi)}{2}$ if $\sigma = \pi$ and $P(\sigma,\pi) = \frac{P'(\sigma,\pi)}{2}$ if $\sigma \neq \pi$[8].

For every $\sigma \in \Omega$, we denote $\mu_\Lambda(\sigma) \triangleq \frac{w_\Lambda(\sigma)}{Z_0+Z_2}$ and for every set $S \subseteq \Omega$, we denote $\mu_\Lambda(S) \triangleq \sum_{\sigma\in S}\mu_\Lambda(\sigma)$.

The following rapid mixing result for above chain was established in [18]. For self-reducible instances, it is standard to obtain **FPRAS** from this rapidly mixing Markov chain [15].

Lemma 29. *For all $\sigma \in \Omega$ and all non-negative integers t, we have*

$$\left\|P^t(\sigma,\cdot) - \mu_\Lambda\right\|_{TV} \leq \frac{1}{2}\left(\mu_\Lambda(\sigma)\right)^{-\frac{1}{2}} \exp\left(-t\cdot\mu_\Lambda(\Omega_0)^2/n^4\right).$$

The remaining part of this section is devoted to prove Lemma 29.

A Congestion and Canonical Paths

Let $\mathcal{G}(\Omega,\mathcal{E})$ be the transition graph of our Markov chain where for every pair of configurations $\sigma,\pi\in\Omega$, $(\sigma,\pi)\in\mathcal{E}$ if and only if $P(\sigma,\pi)>0$.

A *flow-path* γ is a directed path in $\mathcal{G}$ equipped with a weight $\mathrm{wt}(\gamma)$. Canonical paths Γ from $X\subseteq\Omega$ to $Y\subseteq\Omega$ is a set of flow-paths satisfying

$$\sum_{\substack{\text{paths } \gamma\in\Gamma \\ \text{from } x \text{ to } y}} \mathrm{wt}(\gamma) = \pi(x)\pi(y) \quad \text{for all } x\in X \text{ and } y\in Y.$$

The *congestion* of Γ is defined as

$$\rho(\Gamma) \triangleq \max_{(\sigma,\pi)\in\mathcal{E}} \frac{1}{\pi(\sigma)P(\sigma,\pi)} \sum_{\gamma\in\Gamma \text{ s.t. } (\sigma,\pi)\in\gamma} \mathrm{wt}(\gamma).$$

The following lemma was established in [5] and [20]:

Lemma 30. *For every canonical paths Γ from Ω to Ω, every $\sigma\in\Omega$ and every nonegative t, it holds that*

$$\left\|P^t(\sigma,\cdot) - \mu_\Lambda(\cdot)\right|_{TV} \leq \frac{1}{2}\left(\mu_\Lambda(\sigma)\right)^{-\frac{1}{2}} \exp\left(-\frac{t}{n\rho(\Gamma)}\right).$$

Thus it remains to construct a flow-path Γ such that $\rho(\Gamma) \leq \frac{n^3}{\mu_\Lambda(\Omega_0)^2}$.

B The Construction of Canonical Paths

In this section, we describe the construction of canonical paths.

[8]Note that the chain defined here is slightly different with the one used in [18]

Flow from Ω_0 to Ω. Let $\sigma\in\Omega_0$ and $\pi\in\Omega_2$ be two configurations and $z=\sigma\oplus\pi$. Consider a tuple $\left(M_v\in\mathcal{M}_{z|_{\mathcal{E}(v)}}\right)_{v\in V}$, define T as the set of singletons in $\bigcup_{v\in V}M_v$, i.e., $T\triangleq\{S\in M_v \mid v\in V \text{ and } S \text{ is a singleton}\}$. We fix a partition of T into pairs (note that $|T|$ is even by the definition of Ω_0 and Ω_2) and denote the partition as M'. Define $M\triangleq\bigcup_{v\in V}M_v\cup M'\in\mathcal{M}_z$, we call M the partition induced by $\left(M_v\in\mathcal{M}_{z|_{\mathcal{E}(v)}}\right)_{v\in V}$.

Then for every tuple $\left(M_v\in\mathcal{M}_{z|_{\mathcal{E}(v)}}\right)_{v\in V}$, we define a canonical path $\gamma_{\sigma,\pi,M}$ as follows, where $M\in\mathcal{M}_z$ is the partition induced by the tuple: We first construct a graph $G_{M,z}=(V_z,E_M)$ where

- $V_z=\{e_v\in\mathcal{E}\mid z(e_v)=1\}$;
- $E_M=M\cup\left\{\{e_u,e_v\}\in V_z^2 \;\middle|\; \{u,v\}\in E\right\}$.

Since both $\sigma,\pi\in\Omega$, which implies $G_{M,z}$ is a graph consisting of disjoint cycles and a path. We recursively choose an order of edges $\{e_1,e_2,\ldots,e_m\}$ in E_M as follows:

- If there is a unique path $P=(e_1,e_2,\ldots,e_k)$, then start from e_1 and choose edges along the path in the same order. After this is done, remove P.
- If there is no path, choose a cycle $C=(e_1,e_2,\ldots,e_k,e_1)$ such that $\{e_1,e_2\}\in M$. Then start from e_1 and choose edges along the cycle. After this is done, remove C.

This order induces an order of pairs in M. We denote it by $\{S_1,S_2,\ldots,S_t\}$ where each $S_k\in M$ is a pair of half edges.

For every $k=0,1,2,\ldots,t$, let $E_k\triangleq\bigcup_{i=1}^k S_k$. We then construct a flow-path $\gamma_{\sigma,\pi,M}$ in Ω as

$$\sigma=\sigma\oplus E_0\to\sigma\oplus E_1\to\cdots\to\sigma\oplus E_t=\pi,$$

and equip the path with weight

$$\mathrm{wt}(\gamma_{\sigma,\pi,M})=\prod_{v\in V}B_v(\sigma|_{\mathcal{E}(v)},\pi|_{\mathcal{E}(v)},M_v)/(Z_0+Z_2)^2,$$

where for every $v\in V$, $B_v(\cdot,\cdot,\cdot)$ is the set of values witnessing f_v is windable.

Then for every $\sigma \in \Omega_0$ and $\pi \in \Omega$, it holds that

$$\begin{aligned}
&\sum_{M \in \mathcal{M}_z} \mathrm{wt}\left(\gamma_{\sigma,\pi,M}\right) \\
&= \frac{1}{(Z_0+Z_2)^2} \sum_{\left\{M_v \in \mathcal{M}_{z\cap\mathcal{E}(v)}\right\}_{v\in V}} \prod_{v\in V} B_v(\sigma|_{\mathcal{E}(v)}, \pi|_{\mathcal{E}_v}, M_v) \\
&= \frac{1}{(Z_0+Z_2)^2} \cdot \prod_{v\in V} \sum_{M_v \in \mathcal{M}_{z\cap\mathcal{E}(v)}} B_v(\sigma|_{\mathcal{E}(v)}, \pi|_{\mathcal{E}(v)}, M_v) \\
&\overset{(\heartsuit)}{=} \frac{1}{(Z_0+Z_2)^2} \cdot \prod_{v\in V} f_v(\sigma|_{\mathcal{E}(v)}) f_v(\pi|_{\mathcal{E}(v)}) \\
&= \mu_\Lambda(\sigma)\mu_\Lambda(\pi),
\end{aligned}$$

where $(\heartsuit)$ is due to the definition of windability. We denote Γ_0 the canonical paths constructed above.

Flow from Ω to Ω. For every $\sigma, \pi \in \Omega$, for every $\rho \in \Omega_0$, every $M_1 \in \mathcal{M}_{\sigma\oplus\rho}$, every $M_2 \in \mathcal{M}_{\rho\oplus\pi}$, we construct a path $\gamma_{\sigma,\pi,\rho,M_1,M_2}$ which is the concatenation of γ_{σ,ρ,M_1} and γ_{ρ,π,M_2} (since the transition graph of our Markov chain is undirected, we can safely reverse paths in Γ_0). The weight of $\gamma_{\sigma,\pi,\rho,M_1,M_2}$ is $\frac{\mathrm{wt}\left(\gamma_{\sigma,\rho,M_1}\right)\mathrm{wt}\left(\gamma_{\rho,\pi,M_2}\right)}{\mu_\Lambda(\rho)\mu_\Lambda(\Omega_0)}$. The flow is legal since

$$\begin{aligned}
&\sum_{\rho\in\Omega_0} \sum_{M_1 \in \mathcal{M}_{\sigma\oplus\rho}} \sum_{M_2 \in \mathcal{M}_{\rho\oplus\pi}} \mathrm{wt}\left(\gamma_{\sigma,\pi,\rho,M_1,M_2}\right) \\
&= \sum_{\rho\in\Omega_0} \sum_{M_1 \in \mathcal{M}_{\sigma\oplus\rho}} \sum_{M_2 \in \mathcal{M}_{\rho\oplus\pi}} \frac{\mathrm{wt}\left(\gamma_{\sigma,\rho,M_1}\right)\mathrm{wt}\left(\gamma_{\rho,\pi,M_2}\right)}{\mu_\Lambda(\rho)\mu_\Lambda(\Omega_0)} \\
&= \sum_{\rho\in\Omega_0} \frac{\mu_\Lambda(\sigma)\mu_\Lambda(\rho)\mu_\Lambda(\pi)}{\mu_\Lambda(\Omega_0)} \\
&= \mu_\Lambda(\sigma)\mu_\Lambda(\pi).
\end{aligned}$$

C Analysis

In this section, we bound the congestion of the canonical paths constructed in the previous section.

Lemma 31. *Let $\Lambda = (G(V,E), (f_v)_{v\in V})$ be an instance with $|V| = n$ and every f_v is windable, then $Z_0 Z_4 \le Z_2 Z_2$.*

Proof. Note that

$$\begin{aligned}
Z_0 Z_4 &= \sum_{\substack{\sigma\in\Omega_0\\ \pi\in\Omega_4}} w_\Lambda(\sigma) w_\Lambda(\pi) \\
&= \sum_{\substack{\sigma\in\Omega_0\\ \pi\in\Omega_4}} \prod_{v\in V} f_v(\sigma|_{\mathcal{E}(v)}) f_v(\pi|_{\mathcal{E}(v)}) \\
&= \sum_{\substack{\sigma\in\Omega_0\\ \pi\in\Omega_4}} \prod_{v\in V} \sum_{M_v \in \mathcal{M}_{z|_{\mathcal{E}(v)}}} B_v(\sigma|_{\mathcal{E}(v)}, \pi|_{\mathcal{E}(v)}, M_v) \\
&= \sum_{\substack{\sigma\in\Omega_0\\ \pi\in\Omega_4}} \sum_{\left\{M_v \in \mathcal{M}_{z|_{\mathcal{E}(v)}}\right\}_{v\in V}} \prod_{v\in V} B_v(\sigma|_{\mathcal{E}(v)}, \pi|_{\mathcal{E}(v)}, M_v),
\end{aligned}$$

where in the last two lines $z = \sigma \oplus \pi$ and $B_v(\cdot,\cdot,\cdot)$ is the family of values witnessing the windability of f_v.

Fix $(\sigma,\pi) \in \Omega_0 \times \Omega_4$ and $\left\{M_v \in \mathcal{M}_{z|_{\mathcal{E}(v)}}\right\}_{v\in V}$ where $z = \sigma \oplus \pi$. Let M be the set of pairs in $\bigcup_{v\in V} M_v$. Define a graph $G_{M,z} = (V_z, E_M)$ where

- $V_z = \{e_v \in \mathcal{E} \mid z(e_v) = 1\}$;
- $E_M = M \cup \left\{\{e_u, e_v\} \in V_z^2 \;\middle|\; \{u,v\} \in E\right\}$.

Since $(\sigma,\pi) \in \Omega_0 \times \Omega_4$, $G_{M,z}$ consists of two disjoint paths and many disjoint cycles. Let P be one of the path, then by the definition of the windability, it holds that

$$\begin{aligned}
&\prod_{v\in V} B_v(\sigma|_{\mathcal{E}(v)}, \pi|_{\mathcal{E}(v)}, M_v) \\
&= \prod_{v\in V} B_v((\sigma\oplus P)|_{\mathcal{E}(v)}, (\pi\oplus P)|_{\mathcal{E}(v)}, M_v),
\end{aligned}$$

where we use $\sigma \oplus P$ to denote the configurations obtained from σ by flipping the value on vertices in P.

This finishes the proof by noting that $(\sigma\oplus P, \pi\oplus P) \in \Omega_2 \times \Omega_2$ and the mapping $(\sigma,\pi) \to (\sigma\oplus P, \pi\oplus P)$ is injective. □

Lemma 32. *Let Γ_0 be the canonical paths from Ω_0 to Ω constructed above, then $\rho(\Gamma_0) \le \frac{n^3}{\mu_\Lambda(\Omega_0)}$.*

Proof. The congestion of Γ_0 is

$$\rho(\Gamma_0) = \max_{(\sigma,\pi)} \frac{1}{\mu_\Lambda(\sigma)P(\sigma,\pi)} \sum_{\gamma\in\Gamma_0 \text{ with } (\sigma,\pi)\in\gamma} \mathrm{wt}(\gamma).$$

By the definition of the Markov chain, it holds that

$\mu_\Lambda(\sigma)P(\sigma,\pi) = \frac{1}{n^2}\min\left(\mu_\Lambda(\sigma),\mu_\Lambda(\pi)\right)$, thus

$$\begin{aligned}
\rho(\Gamma_0) &\le \max_{\pi\in\Omega} \frac{n^2}{\mu_\Lambda(\pi)} \sum_{\gamma\in\Gamma_0 \text{ with } \pi\in\gamma} \mathrm{wt}\left(\gamma\right) \\
&\le \max_{\pi\in\Omega} \frac{n^2}{\mu_\Lambda(\pi)} \sum_{\substack{\sigma_1\in\Omega_0\\ \sigma_2\in\Omega}} \sum_{\substack{\left(M_v\in\mathcal{M}_{z|_{\mathcal{E}(v)}}\right)_{v\in V}\\ \text{with } \pi\in\gamma_{\sigma_1,\sigma_2,M}}} \mathrm{wt}\left(\gamma_{\sigma_1,\sigma_2,M}\right) \\
&\quad \left(M \text{ is induced by } \left(M_v\right)_{v\in V}, z=\sigma_1\oplus\sigma_2\right) \\
&= \max_{\pi\in\Omega} \frac{n^2}{w_\Lambda(\pi)(Z_0+Z_2)} \sum_{\substack{\sigma_1\in\Omega_0\\ \sigma_2\in\Omega}} \sum_{\substack{\left(M_v\in\mathcal{M}_{z|_{\mathcal{E}(v)}}\right)_{v\in V}\\ \text{with } \pi\in\gamma_{\sigma_1,\sigma_2,M}}} \\
&\quad \prod_{v\in V} B_v(\sigma_1|_{\mathcal{E}(v)},\sigma_2|_{\mathcal{E}(v)},M_v) \\
&= \max_{\pi\in\Omega} \frac{n^2}{w_\Lambda(\pi)(Z_0+Z_2)} \sum_{\substack{\sigma_1\in\Omega_0\\ \sigma_2\in\Omega}} \sum_{\substack{\left(M_v\in\mathcal{M}_{z|_{\mathcal{E}(v)}}\right)_{v\in V}\\ \text{with } \pi\in\gamma_{\sigma_1,\sigma_2,M}}} \\
&\quad \prod_{v\in V} B_v(\pi|_{\mathcal{E}(v)},(\pi\oplus\sigma_1\oplus\sigma_2)|_{\mathcal{E}(v)},M_v) \\
&\le \max_{\pi\in\Omega} \frac{n^2}{w_\Lambda(\pi)(Z_0+Z_2)} \sum_{\sigma_1\in\Omega_0} \sum_{w\in\Omega_0\cup\Omega_2} \sum_{\substack{\left(M_v\in\mathcal{M}_{z|_{\mathcal{E}(v)}}\right)_{v\in V}\\ \text{with } \pi\in\gamma_{\sigma_1,\sigma_2,M}}} \\
&\quad \prod_{v\in V} B_v(\pi|_{\mathcal{E}(v)},(\pi\oplus w)|_{\mathcal{E}(v)},M_v) \\
&\quad \text{where } \left(w\triangleq\sigma_1\oplus\sigma_2\right) \\
&\le \max_{\pi\in\Omega} \frac{n^3}{w_\Lambda(\pi)(Z_0+Z_2)} \\
&\quad \cdot \sum_{w\in\Omega_0\cup\Omega_2} \prod_{v\in V} f_v(\pi|_{\mathcal{E}(v)}) f_v((\pi\oplus w)|_{\mathcal{E}(v)}) \\
&\le n^2\cdot\frac{Z_0+Z_2+Z_4}{Z_0+Z_2} \\
&\le \frac{n^3}{\mu_\Lambda(\Omega_0)},
\end{aligned}$$

where the last inequality is due to Lemma 31. □

Lemma 33. *Let Γ be the canonical paths from Ω to Ω constructed above, then $\rho(\Gamma)\le\frac{n^3}{\mu_\Lambda(\Omega_0)^2}$.*

Proof. The congestion of Γ is

$$\rho(\Gamma) = \max_{(\sigma,\pi)} \frac{\mu_\Lambda(\sigma)P(\sigma,\pi)}{\sum_{\gamma\in\Gamma \text{ with } (\sigma,\pi\in\gamma)} \mathrm{wt}\left(\gamma\right)}.$$

By the definition of Γ, each $\gamma\in\Gamma$ is the concatenation of two paths in Γ_0. Denote $\mathbf{1}_A$ the indicator function of the event A, we have

$$\begin{aligned}
\rho(\Gamma) &= \max_{(\sigma,\pi)} \frac{1}{\mu_\Lambda(\sigma)P(\sigma,\pi)} \sum_{\substack{x,z\in\Omega\\ y\in\Omega_0}} \sum_{M_1\in\mathcal{M}_{x\oplus y}} \sum_{M_2\in\mathcal{M}_{z\oplus y}} \\
&\quad \mathbf{1}_{\substack{(\sigma,\pi)\in\gamma_{x,y,M_1}\\ \vee(\sigma,\pi)\in\gamma_{z,y,M_2}}} \cdot \frac{\mathrm{wt}\left(\gamma_{x,y,M_1}\right)\mathrm{wt}\left(\gamma_{z,y,M_2}\right)}{\mu_\Lambda(y)\mu_\Lambda(\Omega_0)} \\
&= \max_{(\sigma,\pi)} \frac{1}{\mu_\Lambda(\sigma)P(\sigma,\pi)} \sum_{\substack{x,z\in\Omega\\ y\in\Omega_0}} \sum_{\substack{M_1\in\mathcal{M}_{x\oplus y}\\ \text{with } (\sigma,\pi)\in\gamma_{x,y,M_1}}} \\
&\quad \sum_{M_2\in\mathcal{M}_{y\oplus z}} \frac{\mathrm{wt}\left(\gamma_{x,y,M_1}\right)\mathrm{wt}\left(\gamma_{z,y,M_2}\right)}{\mu_\Lambda(y)\mu_\Lambda(\Omega_0)} \\
&= \max_{(\sigma,\pi)} \frac{1}{\mu_\Lambda(\sigma)P(\sigma,\pi)} \sum_{\substack{x,z\in\Omega\\ y\in\Omega_0}} \sum_{\substack{M_1\in\mathcal{M}_{x\oplus y}\\ \text{with } (\sigma,\pi)\in\gamma_{x,y,M_1}}} \\
&\quad \frac{\mathrm{wt}\left(\gamma_{x,y,M_1}\right)\mu_\Lambda(z)}{\mu_\Lambda(\Omega_0)} \\
&= \max_{(\sigma,\pi)} \frac{1}{\mu_\Lambda(\sigma)P(\sigma,\pi)} \sum_{\substack{x\in\Omega\\ y\in\Omega_0}} \sum_{\substack{M_1\in\mathcal{M}_{x\oplus y}\\ \text{with } (\sigma,\pi)\in\gamma_{x,y,M_1}}} \\
&\quad \frac{\mathrm{wt}\left(\gamma_{x,y,M_1}\right)}{\mu_\Lambda(\Omega_0)} \\
&= \frac{\rho(\Gamma_0)}{\mu_\Lambda(\Omega_0)} \\
&\le \frac{n^3}{\mu_\Lambda(\Omega_0)^2}
\end{aligned}$$

□