Combinatorial Bernoulli Factories: Matchings, Flows, and Other Polytopes

Rad Niazadeh
University of Chicago Booth School
of Business
USA
rad.niazadeh@chicagobooth.edu

Renato Paes Leme
Google Research
USA
renatoppl@google.com

Jon Schneider
Google Research
USA
jschnei@google.com

ABSTRACT

A Bernoulli factory is an algorithmic procedure for exact sampling of certain random variables having only Bernoulli access to their parameters. Bernoulli access to a parameter $p \in [0, 1]$ means the algorithm does not know p , but has sample access to independent draws of a Bernoulli random variable with mean equal to p . In this paper, we study the problem of Bernoulli factories for polytopes: given Bernoulli access to a vector $x \in \mathcal{P}$ for a given polytope $\mathcal{P} \subset [0, 1]^n$, output a randomized vertex such that the expected value of the i -th coordinate is *exactly* equal to x_i . For example, for the special case of the perfect matching polytope, one is given Bernoulli access to the entries of a doubly stochastic matrix $[x_{ij}]$ and asked to sample a matching such that the probability of each edge (i, j) be present in the matching is exactly equal to x_{ij} .

We show that a polytope $\mathcal{P}$ admits a Bernoulli factory if and only if $\mathcal{P}$ is the intersection of $[0, 1]^n$ with an affine subspace. Our construction is based on an algebraic formulation of the problem, involving identifying a family of Bernstein polynomials (one per vertex) that satisfy a certain algebraic identity on $\mathcal{P}$. The main technical tool behind our construction is a connection between these polynomials and the geometry of zonotope tilings.

We apply these results to construct an explicit factory for the perfect matching polytope. The resulting factory is deeply connected to the combinatorial enumeration of arborescences and may be of independent interest. For the k -uniform matroid polytope, we recover a sampling procedure known in statistics as Sampford sampling.

CCS CONCEPTS

• **Mathematics of computing** → **Probability and statistics.**

KEYWORDS

exact sampling, Bernoulli factories

ACM Reference Format:

Rad Niazadeh, Renato Paes Leme, and Jon Schneider. 2021. Combinatorial Bernoulli Factories: Matchings, Flows, and Other Polytopes. In *Proceedings*

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

STOC '21, June 21–25, 2021, Virtual, Italy

© 2021 Association for Computing Machinery.

ACM ISBN 978-1-4503-8053-9/21/06.

<https://doi.org/10.1145/3406325.3451072>

of the 53rd Annual ACM SIGACT Symposium on Theory of Computing (STOC '21), June 21–25, 2021, Virtual, Italy. ACM, New York, NY, USA, 14 pages.
<https://doi.org/10.1145/3406325.3451072>

1 INTRODUCTION

Bernoulli factories are basic primitives used in statistics to generate exact samples of a random variable from independent samples of a related random variable. Bernoulli factory techniques have found their applications in settings as diverse as Bayesian mechanism design (Dughmi et al. 6; Cai et al. 4), quantum physics [5, 33], exact simulation of stochastic processes such as diffusion [3], Markov chain Monte Carlo (MCMC) methods [7], and exact Bayesian inference [9, 12]. In mechanism design they allow for black-box reductions for welfare maximization that exactly preserve the Bayesian incentive compatibility, which offers stronger game-theoretical guarantees than approximately incentive compatible reductions. In Bayesian inference and stochastic simulation, the exact sampling afforded by Bernoulli factories allows them to be used in iterative methods without errors compounding.

In this paper we study Bernoulli factories for general polytopes – with a particular focus on combinatorial settings. Before describing this (combinatorial) Bernoulli factory problem, it is useful to revisit the definition of the classic single-parameter version of the problem. The single parameter problem is typically phrased as generating new coins from old ones, where a coin here refers to a Bernoulli random variable. We are given access to a p -coin with unknown parameter p and asked to generate a sample of an $f(p)$ -coin for some known function $f : S \subseteq (0, 1) \rightarrow (0, 1)$. The algorithm does not know p , but has access to as many independent samples as it wants from a Bernoulli random variable with parameter p (the p -coin); the goal is to output 1 with probability $f(p)$.

For the function $f(p) = p^2$, for example, the algorithm can draw two samples from the p -coin and output 1 if both samples are 1, and outputs 0 otherwise. A less trivial example is the function $f(p) = e^{p-1}$. Rewriting this function as the probability generating function of a discrete Poisson random variable, i.e., $f(p) = \mathbb{E}_{X \sim \text{Poisson}(1)}[p^X]$, leads to the following algorithm: (i) sample $X \sim \text{Poisson}(1)$, (ii) draw X independent samples from the p -coin, and (iii) if all the samples are 1 output 1, otherwise output 0. Keane and O'Brien [1994] give necessary and sufficient conditions on function f for the existence of Bernoulli factories.

Before we proceed, we emphasize a crucial point: the Bernoulli factory problem asks for *exact* sampling, as opposed to (even very precise) approximate sampling. This property is essential the aforementioned applications in statistics, mechanism design and quantum mechanics, and is indeed the main motivation behind the study

of Bernoulli factories. Approximate sampling is much simpler; in general, one can build an estimator $\hat{p}$ from i.i.d. samples and then sample a Bernoulli r.v. with parameter $f(\hat{p})$. This, however, is *not* a Bernoulli factory.

Combinatorial Bernoulli Factories. In this paper we will be mostly concerned with sampling a combinatorial object (e.g., a matching or a flow) having black-box sample access to marginal probabilities, say the probability that an edge is present in the matching. Formally, we are given an n -dimensional polytope $\mathcal{P} \subseteq [0, 1]^n$ with vertices V . We are given n coins with unknown probabilities $x_1, \dots, x_n$ such that $x = (x_1, \dots, x_n) \in \mathcal{P} \cap (0, 1)^n$. A *Bernoulli factory* for $\mathcal{P}$ is a then a randomized procedure¹ for sampling a vertex $v \in V$ such that $\mathbb{E}[v] = x$.

It is typical in the Bernoulli factory literature (e.g., Keane and O'Brien 16, Nacu and Peres 22) to restrict the input coins to be non-deterministic, i.e., $x_i \in (0, 1)$. In some cases, though, it is possible to construct factories for all $x \in \mathcal{P}$ also allowing for $\{0, 1\}$ -coordinates (in other words, extending the factory to the boundary of $[0, 1]^n$). We call a such a factory a *strong Bernoulli factory*. We now ask the following question:

Under what conditions does a polytope $\mathcal{P} \subseteq [0, 1]^n$ admit a Bernoulli factory? If it admits one, how can one construct such a factory?

On the path to answer this question, it is useful to keep the following concrete examples in mind:

- **k -subsets** (also known as **k -uniform matroids**): n coins with unknown parameters $\{x_i\}_{i \in [n]}$ are given, such that $\sum_i x_i = k$ for some integer k . We are asked to sample a subset $S \subseteq [n]$ of k elements such that $\Pr[i \in S] = x_i$. This setting corresponds to the polytope

$$\mathcal{P} = \left\{ x \in [0, 1]^n \mid \sum_i x_i = k \right\},$$

which essentially is the k -uniform matroid polytope. The vertices correspond to the indicator vectors of subsets S of size k , i.e., bases of the k -uniform matroid.

- **Matchings**: Consider a complete bipartite graph with an independent x_{ij} -coin for each edge such that the parameters incident to every node sum to 1. We want to sample a perfect matching such that edge (i, j) is included with probability x_{ij} . This setting corresponds to the Birkhoff-von Neumann polytope,

$$\mathcal{P} = \left\{ x \in [0, 1]^{n \times n} \mid \sum_k x_{kj} = \sum_k x_{ik} = 1, \forall i, j \right\},$$

i.e., the set of doubly stochastic matrices. The vertices correspond to perfect matchings (or equivalently permutations over $[n]$) by the Birkhoff-von Neumann Theorem.

- **Flows**: Consider a directed graph (N, E) with a source s and a sink t and an x_{ij} -coin for each edge $(i, j) \in E$ such that for each node other than the source and the sink the sum of x_{ij} for incoming edges is the same as the sum of x_{ij} for

outgoing edges. Let the sum of outgoing edges of the source s is an integer k . We want to sample an integral (s, t) -flow of size k such that edge (i, j) is included with probability exactly x_{ij} .

This setting corresponds to the flow polytope

$$\mathcal{P} = \left\{ x \in [0, 1]^E \mid \sum_{j \mid (i,j) \in E} x_{ij} = \sum_{j \mid (j,i) \in E} x_{ji}, i \neq s, t; \sum_{j \mid (s,j) \in E} x_{sj} = \sum_{j \mid (j,t) \in E} x_{jt} = k \right\}.$$

The vertices are integral (s, t) -flows. For $k = 1$ this means sampling a path from s to t .

Main Result and Techniques. We answer the above question by providing necessary and sufficient conditions to construct combinatorial Bernoulli factories. More formally, we show it is necessary and sufficient that $\mathcal{P}$ is of the form $\mathcal{H} \cap [0, 1]^n$, where $\mathcal{H} = \{x \in \mathbb{R}^n \mid Wx = b\}$ is an affine subspace, for the existence of a Bernoulli factory for $\mathcal{P}$. The result is constructive and allows us to obtain factories for k -subsets, matchings, flows, and all other polytopes of the mentioned form.

The necessary condition is simpler and follows from an argument in polyhedral combinatorics. We show that if the polytope $\mathcal{P}$ is not of the form $\mathcal{H} \cap [0, 1]^n$, there must exist two nearby points x_1 and x_2 in $\mathcal{P}$ and a vertex v such that x_2 must output v with non-negative probability while x_1 must output v with zero probability (see Figure 3). However, no algorithm can perfectly distinguish between x and x' with finitely many samples, so this is impossible.

The technically challenging part of this proof is to construct a factory for polytopes $\mathcal{P}$ of the form $\mathcal{H} \cap [0, 1]^n$. Interestingly, we convert what is originally a probability problem to an algebraic question about Bernstein polynomials, which we then solve with the aid of techniques from geometric combinatorics. The main pieces of this argument are as follows:

- **Race over Bernstein polynomials**: We give a recipe for constructing Bernoulli factories by associating with each vertex v of the polytope $\mathcal{P}$ a multivariate Bernstein polynomial $P_v(x)$ such that the polynomials satisfy $\sum_v P_v(x)(x - v) = 0$. This part of the proof follows from a combination of two ideas in the literature: (i) univariate Bernstein polynomials have been used many times to reason about Bernoulli factories in single-parameter settings, and (ii) the Bernoulli race construction of [6].
- **Generic and non-generic subspaces**: Each subspace $\mathcal{H}$ can be written in the form $Wx = b$ for an full-rank $k \times n$ matrix W . We say that a subspace is generic if the vertices of $\mathcal{H} \cap [0, 1]^n$ have exactly k coordinates in $(0, 1)^n$. For any fixed W the set of vectors b for which $Wx = b$ is non-generic has measure zero.

We first show how to construct strong Bernoulli factories for generic subspaces (see bullets below) and then obtain factories for non-generic subspaces as appropriately defined limits of generic factories. The non-generic construction is important since many polytopes of interest (k -subsets, matchings and flows) are non-generic.

¹For convenience, the Bernoulli factory algorithm is allowed to use external randomness, besides using the given coins. This is indeed without loss of generality, since it is shown by [32] that it is possible to sample any random variable with known probability using a p -coin with unknown $p \in (0, 1)$.

- *Polynomials from minors*: Given a generic affine subspace $\mathcal{H}$ of the form $Wx = b$ we can associate each vertex v of the polytope $\mathcal{P}$ to a subset S of size k of variables that are *basic* (in the terminology of the simplex method). The determinant of the subset of the $k \times k$ minor corresponding to the basic variables is then used to construct a Bernstein monomial associated with v .
- *Zonotope tilings*: Finally, we need to show that this construction satisfies the polynomial identity $\sum_v P_v(x)(x - v) = 0$ in the first bullet. This is done by associating each vertex of the polytope with a point in a geometric space. The polynomial identity is then proved by considering two distinct decompositions of this geometric space into zonotope tilings.

These ingredients lead to the following algorithm (Algorithm 1). In this algorithm, we use W_S to denote the $k \times k$ submatrix formed by the columns of W corresponding to indices in $S \subseteq [n]$. We also assume that the description of the affine subspace is such that $|\det W_S| \leq 1$ for each subset S of size k (this is without loss of generality, since one can always scale W and b to satisfy it).

Algorithm 1 BERNOULLI FACTORY FOR GENERIC SUBSPACES

Pick a vertex $v \in V$ uniformly at random.
 Let $S = \{i \in [n] \mid v_i \in (0, 1)\}$.
 For each $i \notin S$, sample an x_i -coin. If the sample is not equal to v_i , restart.
 For each $i \in S$ sample two x_i -coins. If the samples are 1 and 0, proceed. Otherwise, restart.
 With probability $|\det W_S|$ output v . With remaining probability, restart.

Consequence for k -subset. For the k -subset problem, we recover the sampling procedure in statistics known as Sampford sampling [26]. Our result shows that this particular procedure can indeed be implemented as a Bernoulli factory. Since the k -subset polytope is non-generic it is obtained via the limit of generic polytopes. While the factories we construct for generic polytopes are strong factories, the limit of these factories diverges on the boundary of $[0, 1]^n$ – hence, we obtain a Bernoulli factory in the limit but not a strong factory. Indeed this is also a feature of Sampford’s original sampling process – it also requires all probabilities to be strictly in $(0, 1)$.

One may ask whether it is possible to extend Sampford sampling to also allow for deterministic variables. We prove a impossibility result showing that can exist no strong Bernoulli factory for k -subset with exponential tails (i.e., the probability of requiring more than t coins is at most c^t for some c). In particular, this implies that it is impossible to construct a strong Bernoulli factory for k -subset by running a Bernoulli race over Bernstein polynomials. This result can be found in Section 7 of the online version of this paper.

Consequence for matching. For the matching problem, our Bernoulli factory has a particularly nice combinatorial structure. This structure is somewhat surprising since it goes through combinatorial constructions that do not seem to be related to sampling perfect matchings at first glance. In particular, we show our Bernstein polynomials can be alternatively obtained by enumerating

particular monomials, one for each rooted arborescence in the complete graph K_n . The final argument to show the desired polynomial identity relies on counting arborescences using variants of Kirchhoff’s matrix-tree theorem [17, 31] and additional combinatorial arguments related to trees and cycle covers.

Here is the sampling procedure for matchings (Algorithm 2). Recall that we have an $n \times n$ doubly-stochastic matrix $[x_{ij}]$ with all entries in $(0, 1)$. For each (i, j) we have access to independent samples of an x_{ij} -coin. Our goal is to sample a perfect matching such that each edge (i, j) is included in the matching with probability exactly equal to x_{ij} .

Algorithm 2 BERNOULLI FACTORY FOR MATCHING

Pick uniformly at random a permutation π over $[n]$.
 For each $i \in [n]$ sample the $x_{i\pi(i)}$ -coin. If any sample is 0, restart.
 Pick uniformly at random a spanning tree of the complete graph K_n .
 Let T be the set of edges (i, j) of the tree oriented toward vertex 1.
 For each edge $(i, j) \in T$ sample the $x_{i\pi(j)}$ -coin. If any sample is 0, restart.
 Output the matching $\{(i, \pi(i))\}_{i \in [n]}$.

Paper Organization. In Section 2 we provide a formal definition of a Bernoulli factory as a decision tree and describe how it can be constructed via Bernstein polynomials. In Section 3 we give a self-contained presentation of a factory for matching via a combinatorial construction. In Section 4 we give necessary conditions on $\mathcal{P}$ for the existence of Bernoulli factories. We show in the following two sections that those conditions are also sufficient. In Section 5, we construct strong factories for generic subspaces via the geometry of zonotope tilings. In Section 6, we describe how to obtain factories for non-generic subspaces as limits of factories for generic ones. Some proofs and results are deferred to the online version of this paper² due to space constraints.

Further Related Work. Beyond the work of Keane and O’Brien [1994], several other papers have studied different constructions and fast Bernoulli factory algorithms for functions $f : (0, 1) \rightarrow (0, 1)$. [22] give the necessary and sufficient conditions for the existence of fast Bernoulli factories. An alternative algorithm for general analytic functions is proposed in [18]. A fast Bernoulli factory for rational functions is proposed in [21]. More recently, [20] show how to construct a more practical Bernoulli factory for rational functions using coupling from the past [25]. Both of these results extend to the “dice enterprise problem”, where the goal is exact simulation of a multivariate rational mapping $f : \Delta^m \rightarrow \Delta^m$ between probability simplices. Faster Bernoulli factories for linear functions are studied in [13, 14]. [19] studies near-optimal Bernoulli factories for power series and [11] study a particular class of Bernstein polynomials. Finally, extending Bernoulli factory algorithms to quantum settings is studied in [23].

In addition to those mentioned earlier, Bernoulli factory techniques have recently found other applications in different corners of

²<https://arxiv.org/abs/2011.03865>

computer science and statistics. They have been successfully applied to exact simulations of diffusions [3], designing exact simulation methods using MCMCs for Bayesian inference [8, 12], designing particle filters [27], and designing blackbox reductions in Bayesian mechanism design [4, 6].

Indirectly related to us is the line of work on efficient approximate sampling from particular family of distributions (e.g., maximum entropy) over combinatorial polytopes (e.g., matching and matroid polytopes), satisfying a given vector of marginals. For example, see [1, 29, 30]. Our work diverges from this literature by the fact that a Bernoulli factory algorithm has only Bernoulli access to the marginal vector, and that it should satisfy the marginals exactly. Also, some aspects of the Bernoulli factory problem resemble the exact simulation of MCMCs in different contexts [2, 15, 24].

2 PRELIMINARIES

We start by formally defining a general Bernoulli factory that captures both the standard single-parameter Bernoulli factory and our generalization to the Bernoulli factory for polytopes. We then show how to use particular polynomials to construct those.

2.1 General Bernoulli Factories

Below we define a general Bernoulli factory outputting elements in a set V using a decision tree. We note that this definition is not tied to any particular function f .

Definition 2.1 (Bernoulli factory). A Bernoulli factory $\mathcal{F}$ with output in V is represented by a (possibly infinite) rooted binary tree $\mathcal{T}$. Each node in $\mathcal{T}$ has either 2 children (in which case it is a *non-leaf*) or 0 children (in which case it is a *leaf*). Each non-leaf w is labeled with one of the n random variables $\{x_1, x_2, \dots, x_n\}$ or with a constant $c \in (0, 1)$. When executing the protocol, we either flip the x_i -coin in the label or a c -coin with known probability c . The edges from a non-leaf w to its two children are labelled 0 and 1, corresponding to the output of the coin flip at w . Each leaf node ℓ is labelled with some $v \in V$, representing the output of our Bernoulli factory upon reaching this leaf node.

To execute the factory $\mathcal{F}$, we start at the root node and repeatedly follow the following procedure. If we are at a non-leaf node w , we flip the coin given by w 's label, receive a result $r \in \{0, 1\}$, and follow the edge with label r to one of w 's children. If we are at a leaf node ℓ , we simply output the label of ℓ .

We let $\mathcal{F}(x) \in V \cup \{\emptyset\}$ denote the random variable corresponding to the output of the Bernoulli factory $\mathcal{F}$ on input $x \in [0, 1]^n$ (if $\mathcal{F}(x)$ does not terminate, we write $\mathcal{F}(x) = \emptyset$). We say a factory $\mathcal{F}$ *terminates almost surely (a.s.)* on a domain $S \subseteq [0, 1]^n$ if $\Pr[\mathcal{F}(x) = \emptyset] = 0$ for all $x \in S$. Moreover, if the tree $\mathcal{T}$ corresponding to a Bernoulli factory $\mathcal{F}$ is finite, we say that $\mathcal{F}$ is a *finite Bernoulli factory*.

Definition 2.2 (One-bit Bernoulli factory). We say that $\mathcal{F}$ outputting in $\{0, 1\}$ is a *one-bit Bernoulli factory* for the function $f : S \rightarrow [0, 1]$ on S if (i) $\mathcal{F}$ terminates a.s. on S , and (ii) $\Pr[\mathcal{F}(x) = 1] = f(x)$, $\forall x \in S$.

2.2 Bernstein Polynomials

Fix a Bernoulli factory $\mathcal{F}$, and consider a leaf node ℓ of $\mathcal{F}$. Let $\Pr[\mathcal{F}(x) \rightarrow \ell]$ denote the probability that $\mathcal{F}(x)$ terminates at leaf ℓ . By multiplying out the probabilities of each transition along the path from the root of $\mathcal{F}$ to ℓ , we can write $\Pr[\mathcal{F}(x) \rightarrow \ell]$ in the form

$$\Pr[\mathcal{F}(x) \rightarrow \ell] = c \prod_{i=1}^n x_i^{a_i} (1 - x_i)^{b_i} \quad (1)$$

for some $c \in [0, 1]$ and non-negative integers a_i, b_i (for example, b_i is the number of times variable x_i appears on the path to v where we take the edge labelled 0). The expression on the right-hand side of Equation (1) is known as a *Bernstein monomial*.

Definition 2.3 (Bernstein polynomial). A Bernstein monomial in n variables is a polynomial of the form: $M(x) = \prod_{i=1}^n x_i^{a_i} (1 - x_i)^{b_i}$ for $a_i, b_i \in \mathbb{Z}^{\geq 0}$. We will say that $a_i + b_i$ is the degree with respect to variable i of this monomial and denote it $\deg_i(M)$. A *Bernstein polynomial* in n variables is a positive combination of finitely many Bernstein monomials: $P(x) = \sum_{j=1}^k c_j M_j(x)$ for Bernstein monomials $M_j(x)$ and coefficients $c_j \in \mathbb{R}^+$.

Note that we can write $\Pr[\mathcal{F}(x) = v]$ as the sum of $\Pr[\mathcal{F}(x) \rightarrow \ell]$ over all leaves ℓ with label v . This means we can write $\Pr[\mathcal{F}(x) = v]$ as a weighted series of Bernstein monomials; in particular, if $\mathcal{F}(x)$ is a finite Bernoulli factory, then $\Pr[\mathcal{F}(x) = v]$ is a Bernstein polynomial in x . One fact that will prove particularly useful is a partial converse to this: given any Bernstein polynomial $P(x)$, it is always possible to construct a Bernoulli factory for a suitably normalized version of $P(x)$.

LEMMA 2.4. Let $P(x) = \sum_{j=1}^k c_j M_j(x)$ be a Bernstein polynomial in n variables, and let $C = \sum_{j=1}^k c_j$. Then there exists a finite one-bit Bernoulli factory for $P(x)/C$.

PROOF. Consider the following Bernoulli factory $\mathcal{F}$. We first sample a monomial $M_j(x)$ with probability c_j/C (by using external randomness). Now, if $M_j(x) = \prod_{i=1}^n x_i^{a_i} (1 - x_i)^{b_i}$, flip each coin i a total of $\deg_i(M)$ times. If for each coin i , the first a_i flips returned 1 and the next b_i flips returned 0, output 1 for the overall factory $\mathcal{F}$. Otherwise, output 0.

Conditioned on sampling monomial M_j , we return 1 with probability $\prod_{i=1}^n x_i^{a_i} (1 - x_i)^{b_i} = M_j(x)$. Since we sample monomial M_j with probability c_j/C , the total probability $\mathcal{F}(x) = 1$ equals $\Pr[\mathcal{F}(x) = 1] = \sum_j \frac{c_j}{C} M_j(x) = \frac{P(x)}{C}$, as desired. ■

We now describe a method for constructing a factory outputting in V from a collection of one-bit Bernoulli factories for each element $v \in V$. This method is known as a *Bernoulli race* and it was introduced in [6]. We summarize its properties in the following theorem.

THEOREM 2.5 (BERNOULLI RACE). Fix a domain $S \subseteq [0, 1]^n$. For each $v \in V$, let $\mathcal{F}_v$ be a one-bit Bernoulli factory implementing a function $f_v : S \rightarrow [0, 1]$. If $\sum_{v \in V} f_v(x) > 0$, $\forall x \in S$, then there exists a Bernoulli factory $\mathcal{G}$ that terminates a.s. on S and outputs $v \in V$ with probability $f_v(x) / \sum_{v' \in V} f_{v'}(x)$.

PROOF. Consider the following procedure for $\mathcal{G}(x)$:

- (i) Sample a v uniformly at random from V .
- (ii) Run the factory $F_v(x)$. If the factory returns 1, output v . Otherwise, return to step (i).

We claim that this procedure terminates a.s. on S and outputs vertex v with probability $f_v(x)/\sum_{v'} f_{v'}(x)$. To see this, first note that each iteration of this procedure terminates with probability $(\sum_{v'} f_{v'}(x))/|V| > 0$. Since there is a positive chance of terminating each round, and since each individual factory F_v terminates a.s. on S , this procedure will terminate a.s. on S .

Now, note that we can write

$$\Pr[\mathcal{G}(x) = v] = \frac{f_v(x)}{V} + \left(1 - \frac{\sum_{v'} f_{v'}(x)}{V}\right) \Pr[\mathcal{G}(x) = v], \quad (2)$$

since there is a $\frac{f_v(x)}{V}$ chance we output v in any given round, and a $1 - \frac{\sum_{v'} f_{v'}(x)}{V}$ chance we restart the procedure. Rearranging Equation (2), we have $\Pr[\mathcal{G}(x) = v] = f_v(x)/\sum_{v'} f_{v'}(x)$, as desired. ■

In our applications, we will specifically want to take Bernoulli races over finite Bernoulli factories implementing Bernstein polynomials.

COROLLARY 2.6 (BERNOULLI RACE OVER BERNSTEIN POLYNOMIALS). *For each $v \in V$, let $P_v(x)$ be a Bernstein polynomial in n variables. Fix a domain $S \subseteq [0, 1]^n$. If $\sum_v P_v(x) > 0$ for all $x \in S$, then there exists a Bernoulli factory which terminates a.s. on S and outputs v with probability $P_v(x)/\sum_{v'} P_{v'}(x)$.*

PROOF. By Lemma 2.4, for each $v \in V$, there exists a $C_v \geq 1$ such that for any $C \geq C_v$, there exists a finite Bernoulli factory for $P_v(x)/C$ over S . Choose $C = \max_v C_v$, and run a Bernoulli race over factories implementing $P_v(x)/C$. By Theorem 2.5, such a race will output v with probability $P_v(x)/\sum_{v'} P_{v'}(x)$, as desired. ■

2.3 Combinatorial Factories

Finally, we return to the main focus of this paper. Recall that we wish to, given x_i -coins corresponding to the coordinates of a point x within some polytope $\mathcal{P} \subseteq [0, 1]^n$, output a vertex v of $\mathcal{P}$ so that $\mathbb{E}[v] = x$.

Definition 2.7 (Bernoulli factory for a polytope $\mathcal{P}$). Let $\mathcal{P} \subseteq [0, 1]^n$ be a polytope contained in the unit hypercube, and let $\tilde{\mathcal{P}} = \mathcal{P} \cap (0, 1)^n$. Let V denote the set of vertices of $\mathcal{P}$. A Bernoulli factory for $\mathcal{P}$ is a factory $\mathcal{F}$ outputting in V such that

$$\mathbb{E}[\mathcal{F}(x)] = x, \forall x \in \tilde{\mathcal{P}}.$$

If the factory terminates almost surely for all $x \in \mathcal{P}$ (as opposed to just $x \in \tilde{\mathcal{P}}$), we say it is a *strong Bernoulli factory* for $\mathcal{P}$.

Our main tool for constructing polytope factories will be to assign a Bernstein polynomial to each vertex and run a Bernoulli race over such polynomials (see Corollary 2.6).

THEOREM 2.8. *In the setting of Definition 2.7, if $P_v(x)$ is a non-zero Bernstein polynomial in n variables for each $v \in V$ satisfying the following vector equality:*

$$\sum_{v \in V} P_v(x)(v - x) = 0, \forall x \in \mathcal{P}. \quad (3)$$

then running a Bernoulli race over the polynomials $P_v(x)$ (per Corollary 2.6) results in a Bernoulli factory for $\mathcal{P}$. Moreover, if

$$\sum_{v \in V} P_v(x) > 0, \forall x \in \mathcal{P} \quad (4)$$

it results in a strong Bernoulli factory for $\mathcal{P}$.

PROOF. Since non-zero Bernstein polynomials are strictly positive on $(0, 1)^n$ this automatically guarantees a.s. termination on $\tilde{\mathcal{P}}$. To check that $\mathbb{E}[\mathcal{F}(x)] = x$, it is sufficient to re-arrange Equation (3) as follows: $\sum_{v \in V} P_v(x)v = \sum_{v \in V} P_v(x)x$. Dividing by $\sum_{v \in V} P_v(x)$ we obtain exactly $\mathbb{E}[\mathcal{F}(x)] = x$. ■

3 A FACTORY FOR MATCHING

In this section, we construct a Bernoulli factory for the Birkhoff-von Neumann perfect matching polytope using a race over Bernstein polynomials, as described in Corollary 2.6. In later sections we will see how to systematically construct such factories for general polytopes; for now we will simply demonstrate the factory through its corresponding polynomials and prove that it works.

Throughout this section, let $\mathcal{B}_n \subseteq [0, 1]^{n \times n}$ denote the n^{th} Birkhoff-von Neumann polytope. This polytope contains all doubly stochastic n -by- n matrices. By the Birkhoff-von Neumann theorem, $\mathcal{B}_n$ has $n!$ vertices, each corresponding to one of the n -by- n permutation matrices (e.g., see [28]). Each permutation π can in turn be thought of as a perfect matching in the complete bipartite graph $K_{n,n}$. Let S_n be the set of permutations of $[n]$. We identify $\pi \in S_n$ with the n -by- n permutation matrix $[I_{i,j}]_{n \times n} \in \{0, 1\}^{n \times n}$, where $I_{i,j} = \mathbb{I}\{j = \pi(i)\}$. We will abuse notation and use π to denote both a permutation and its corresponding matrix.

Overview. Recall from Theorem 2.8 that we can specify a factory for the polytope $\mathcal{B}_n$ by specifying a non-zero Bernstein polynomial $P_\pi(x)$ in n^2 variables for each vertex $\pi \in S_n$ of $\mathcal{B}_n$, satisfying Equation (3) for $x \in \mathcal{B}_n$. To specify these polynomials, we identify their monomials with certain directed graphs. An *arborescence rooted at r* is a directed graph T where for any vertex v of T , there is exactly one directed path from v to r (in other words, it is a directed spanning tree where all edges are oriented towards r). Let $\mathcal{T}_r(n)$ be the set of arborescences rooted at r with n labelled vertices $1, \dots, n$. Each element of $\mathcal{T}_r(n)$ is a collection of directed edges. Fix an arbitrary root $r \in [n]$. Then consider the following polynomials:

$$\forall \pi \in S_n : P_\pi(x) = \prod_{i=1}^n x_{i,\pi(i)} \sum_{T \in \mathcal{T}_r(n)} \prod_{(u,v) \in T} x_{u,\pi(v)}. \quad (5)$$

Figure 1: Arborescences in $\mathcal{T}_1(3)$ corresponding to the monomials in P_e .

Example 3.1. There are 3 arborescences rooted at 1 on 3 vertices: $\{(2, 1), (3, 1)\}$, $\{(3, 1), (2, 3)\}$, and $\{(2, 1), (3, 2)\}$. For the identity permutation ϵ , we thus have (see Figure 1):

$$P_\epsilon(x) = x_{1,1}x_{2,2}x_{3,3} (x_{2,1}x_{3,1} + x_{3,1}x_{2,3} + x_{2,1}x_{3,2}).$$

Replacing ϵ with a different permutation π corresponds to applying π to the second subscript of each variable. For example, for the permutation $\pi = (2, 3, 1)$, we have:

$$P_\pi(x) = x_{1,2}x_{2,3}x_{3,1} (x_{2,2}x_{3,2} + x_{3,2}x_{2,1} + x_{2,2}x_{3,3}).$$

The polynomials $\{P_\pi(x)\}_{\pi \in S_n}$ defined above are clearly non-zero Bernstein polynomials, as required by Theorem 2.8. To show that $\mathcal{B}_n$ admits a Bernoulli factory, it only remains to show that the vector equality in Equation (3) holds for all $x \in \mathcal{B}_n$. Note that in Equation (5), we choose an arbitrary vertex as the root of our arborescences in order to identify the polynomials $\{P_\pi(x)\}_{\pi \in S_n}$. Interestingly – as we formally show in Proposition 3.2 – the right hand side of Equation (5) is the same for *any* choice of root as long as we restrict our attention to points $x \in \mathcal{B}_n$. This property is indeed related to the fact that each polynomial $P_\pi(x)$ can be written as the product of a symmetric term and minor of a particular weighted directed graph Laplacian. More interestingly – as we formally show in Proposition 3.5 – this property, together with a combinatorial argument relying on trees and permutations, are the keys in showing that Equation (3) holds for points $x \in \mathcal{B}_n$.

PROPOSITION 3.2. *Let $\mathcal{L}$ be the n -by- n weighted directed Laplacian with (arc) weights $[x_{i,j}]_{n \times n}$, i.e.,*

$$\forall i, j \in [n] : \mathcal{L}_{i,j} = \begin{cases} \sum_{k \neq i} x_{k,i} & \text{if } i = j \\ -x_{i,j} & \text{if } i \neq j \end{cases}$$

Moreover, for any $r \in [n]$, let $\mathcal{L}^{(r)}$ denote the $(n-1)$ -by- $(n-1)$ submatrix of $\mathcal{L}$ obtained by removing the row and the column corresponding to r . Then for any $x \in \mathcal{B}_n$, and any $r, r' \in [n]$:

$$\begin{aligned} \sum_{T \in \mathcal{T}_r(n)} \prod_{(u,v) \in T} x_{u,v} &= \det[\mathcal{L}^{(r)}] \\ &= \det[\mathcal{L}^{(r')}] = \sum_{T \in \mathcal{T}_{r'}(n)} \prod_{(u,v) \in T} x_{u,v}. \end{aligned} \quad (6)$$

In order to prove the above proposition, we rely on two results from algebraic combinatorics. The first result is Tutte's matrix-tree theorem [31], which is essentially an adaptation of the standard Kirchoff's matrix-tree theorem [17] to weighted directed graphs.

THEOREM 3.3 (TUTTE'S MATRIX-TREE THEOREM [31]). *Let $\mathcal{L}$ be the n -by- n weighted directed Laplacian matrix with weights $[x_{i,j}]$. Then for any $r \in [n]$,*

$$\det[\mathcal{L}^{(r)}] = \sum_{T \in \mathcal{T}_r(n)} \prod_{(u,v) \in T} x_{u,v}.$$

The second result implies that all the principal minors of a *zero-line-sum* (ZLS) matrix, i.e., a matrix whose rows and columns all sum to 0, are equal. Its proof can be found in the online version of this paper.

LEMMA 3.4 (ZLS MATRICES HAVE EQUAL COFACTORS). *Let A be an n -by- n matrix satisfying $\sum_{j=1}^n A_{ij} = 0$ for each $i \in [n]$ and satisfying*

$\sum_{i=1}^n A_{ij} = 0$ for each $j \in [n]$. For any $i, j \in [n]$, let $A^{(i,j)}$ denote the $(n-1)$ -by- $(n-1)$ submatrix of A obtained by removing the row and the column corresponding to i and j respectively. Then, for any $i, j, i', j' \in [n]$,

$$(-1)^{(i+j)} \det[A^{(i,j)}] = (-1)^{(i'+j')} \det[A^{(i',j')}].$$

PROOF OF PROPOSITION 3.2. First of all, Theorem 3.3 implies that for any choice of $r \in [n]$, $\det[\mathcal{L}^{(r)}] = \sum_{T \in \mathcal{T}_r(n)} \prod_{(u,v) \in T} x_{u,v}$. Next, note that every row in $\mathcal{L}$ sums to zero by definition. Moreover, every column in $\mathcal{L}$ also sums to zero when $x \in \mathcal{B}_n$, simply because $\mathcal{L}_{i,i} = 1 - x_{i,i}$ for $x \in \mathcal{B}_n$. Hence $\mathcal{L}$ is ZLS for $x \in \mathcal{B}_n$ and $\det[\mathcal{L}^{(r)}] = \det[\mathcal{L}^{(r')}]$ for any $r, r' \in [n]$ (due to Lemma 3.4), as desired. ■

Given Proposition 3.2, we are now ready to prove the main result of this section, that is, our polynomials $\{P_\pi\}_{\pi \in S_n}$ satisfy Equation (3) for all points in the perfect matching polytope.

PROPOSITION 3.5. *Consider the polynomials $\{P_\pi\}_{\pi \in S_n}$ as in Equation (5). Then for any $x \in \mathcal{B}_n$,*

$$\sum_{\pi \in S_n} (\pi - x) P_\pi(x) = 0. \quad (7)$$

REMARK 1. *Every step of the Bernoulli race procedure described in Theorem 2.5 – over the Bernstein polynomials in Equation (5) – can be implemented efficiently by sampling uniform random permutations and uniform random spanning trees in the complete graph K_n . See Algorithm 2.*

3.1 Proof of Proposition 3.5

Recall that Equation (7) is an n -by- n matrix equality. Fix any $(r, c) \in [n] \times [n]$; we will show that this equality holds for its $(r, c)^{\text{th}}$ entry. In other words, using the fact that $\pi_{r,c} = \mathbb{I}\{\pi(r) = c\}$ in the permutation matrix π , we will show that

$$\sum_{\pi | \pi(r)=c} P_\pi(x) = \sum_{\pi} x_{r,c} P_\pi(x). \quad (8)$$

Since $x \in \mathcal{B}_n$, the sum of each row and column of x is equal to 1. In particular, $\sum_{i=1}^n x_{r,i} = 1$. Therefore, by multiplying the LHS of Equation (8) and $\sum_{i=1}^n x_{r,i}$, it suffices to show that

$$\sum_{i=1}^n x_{r,i} \sum_{\pi | \pi(r)=c} P_\pi(x) = x_{r,c} \sum_{\pi} P_\pi(x). \quad (9)$$

Recall from Proposition 3.2 that polynomials $\{P_\pi\}_{\pi \in S_n}$ are invariant to the choice of the root of arborescences used in Equation (5). Suppose r is used as the root for all π when defining $P_\pi(x)$. We will show that Equation (9) is true as a polynomial identity – i.e. it is true not just for $x \in \mathcal{B}_n$, but for all $x \in \mathbb{R}^{n^2}$. To do so, it is enough to show that for any fixed $i \in [n]$,

$$x_{r,i} \sum_{\pi | \pi(r)=c} P_\pi(x) = x_{r,c} \sum_{\pi | \pi(r)=i} P_\pi(x). \quad (10)$$

Summing Equation (10) over all $i \in [n]$, we obtain Equation (9), as desired. Now, recall the definition of $P_\pi(x)$ when r is used as the root:

$$P_\pi(x) = \prod_{i=1}^n x_{i,\pi(i)} \sum_{T \in \mathcal{T}_r(n)} \prod_{(u,v) \in T} x_{u,\pi(v)}.$$

Note that in an arborescence rooted at r , the vertex r has no outgoing edges. Therefore, the only variable of the form $x_{r,*}$ occurring in $P_\pi(x)$ is the variable $x_{r,\pi(r)}$, which divides each term of $P_\pi(x)$ exactly once. Define $Q_\pi(x) \triangleq P_\pi(x)/x_{r,\pi(r)}$. To prove Equation (10), it then suffices to show

$$\sum_{\pi|\pi(r)=c} Q_\pi(x) = \sum_{\pi|\pi(r)=i} Q_\pi(x). \quad (11)$$

In order to prove Equation (11), first note that the LHS of this equation, i.e., the sum of $Q_\pi(x)$ over permutations $\pi \in S_n$ with $\pi(r) = c$, can be written as:

$$\sum_{\pi|\pi(r)=c} \prod_{i \in [n], i \neq r} x_{i,\pi(i)} \sum_{T \in \mathcal{T}_r(n)} \prod_{(u,v) \in T} x_{u,\pi(v)} = \sum_{\pi|\pi(r)=c} \sum_{T \in \mathcal{T}_r(n)} \prod_{(u,v) \in T} x_{u,\pi(u)} x_{u,\pi(v)}, \quad (12)$$

simply because each $u \in [n] \setminus \{r\}$ has exactly one outgoing edge in every arborescence $T \in \mathcal{T}_r(n)$. We will interpret the RHS of Equation (12) as enumerating certain undirected bipartite graphs. Given a permutation π with $\pi(r) = c$ and an r -rooted arborescence T , consider an undirected bipartite graph $G(\pi, T)$ on $2n$ vertices, with n vertices on the left (labelled 1_L through n_L) and n vertices on the right (labelled 1_R through n_R). The edges are constructed as follows (see Figure 2):

- for each $u \in [n] \setminus \{r\}$, add the edge $(u_L, \pi(u)_R)$.
- for each directed edge $u \rightarrow v$ in T , add the edge $(u_L, \pi(v)_R)$.

Now, it is straightforward to verify that the summation in Equation (12) can be written as

$$\sum_{\pi|\pi(r)=c} Q_\pi(x) = \sum_{\pi|\pi(r)=c} \sum_{T \in \mathcal{T}_r(n)} \prod_{(u_L, v_R) \in G(\pi, T)} x_{u,v}.$$

Next, we define a collection of bipartite graphs $\mathcal{G}_r$ on vertices $\{1_L, \dots, n_L\} \cup \{1_R, \dots, n_R\}$ for a fixed root r – which we call *r-bi-trees* (see the definition below). We then claim there is a bijection between (π, T) pairs in the above summation (with $\pi(r) = c$) and bipartite graphs $G' \in \mathcal{G}_r$ where $G' = G(\pi, T)$. If the claim holds, we have

$$\sum_{\pi|\pi(r)=c} Q_\pi(x) = \sum_{G' \in \mathcal{G}_r} \prod_{(u_L, v_R) \in G'} x_{u,v},$$

and since the RHS does not depend on the identity of c , it immediately implies Equation (11).

Definition 3.6 (*r*-bi-tree). For any root $r \in [n]$, an undirected bipartite graph G on $2n$ vertices $\{1_L, \dots, n_L\}$ and $\{1_R, \dots, n_R\}$ is an *r-bi-tree* if it satisfies the following conditions:

- The vertex r_L is an isolated vertex.
- The remainder of the vertices (aside from r_L) belong to a single connected component.
- Each vertex u_L (where $u \neq r$) on the left side has degree exactly equal to 2.

We finish the proof by sketching why the above bijection claim holds in the following lemma. A more detailed proof of this lemma can be found in the online version of this paper.

LEMMA 3.7 (BIJECTION). *For any $r, c \in [n]$, there exists a one-to-one correspondence between pairs $(\pi, T) \in S_n \times \mathcal{T}_r(n)$ where $\pi(r) = c$ and *r*-bi-trees G' in $\mathcal{G}_r$ where $G' = G(\pi, T)$.*

PROOF SKETCH. We prove the bijection in two parts:

Part (i): first, we claim $G(\pi, T)$ is an *r*-bi-tree. Notice that $G' = G(\pi, T)$ can be constructed from (π, T) by a *reverse* breadth-first search (BFS) walk on T starting from root r , and then adding both edges $(u_L, \pi(v)_R)$ and $(u_L, \pi(u)_R)$ to G' each time the walk moves from a vertex v to a vertex u (this is possible only if $u \rightarrow v$ is a directed edge in T). This step can alternatively be seen as adding a path of length 2 from $\pi(v)_R$ to $\pi(u)_R$, passing through u_L , in G' . See Figure 2 (left to right) for a pictorial demonstration. Now the claim can be proved as follows. As root r has no outgoing edges in T , r_L does not appear in any edges of G' and remains isolated. Moreover, each $u \neq r$ is visited exactly once in the reverse BFS walk, which adds exactly two edges incident to u_L in G' . Therefore, each u_L for $u \neq r$ has degree 2. Finally, G' has no cycles, as we basically replace each edge in the undirected version of T with a path of length 2 to construct G' . As the forest G' has $2n - 2$ edges, the remaining $2n - 1$ vertices aside from r_L should belong to a single connected component, which finishes the proof of our first claim.

Part (ii): second, we show the mapping $G(\pi, T)$ has an inverse. In other words, we propose an inverse mapping that given an *r*-bi-tree G' uniquely returns a permutation π (satisfying $\pi(r) = c$) and *r*-rooted arborescence $T \in \mathcal{T}_r(n)$, so that $G(\pi, T) = G'$. To construct such a pair (π, T) , consider a BFS walk on the given undirected bipartite graph G' starting from $c_R = \pi(r)_R$ (index the BFS tree layers by $0, 1, 2, \dots$). We first show how to construct a permutation π satisfying $\pi(r) = c$ from the walk. As $G' \setminus \{r_L\}$ is a single connected component, the BFS walk will visit all the vertices in G' except for r_L . Moreover, in each odd layer of the BFS walk, it visits a left vertex u_L with degree exactly 2, as G' is an *r*-bi-tree. Once the walk enters u_L , there is only one remaining indecent edge (u_L, v_R) that can be added next to the BFS tree. Add this edge to the “matching” π by setting $\pi(v) = u$. At the end of the walk, the constructed π (together with setting $\pi(r) = c$) gives a permutation as desired, simply because the BFS tree visits every right hand side vertex exactly once. Next, revisit the BFS walk and construct an arborescence T by adding a directed edge $u \rightarrow v$ to T for every edge $(u_L, \pi(v)_R)$ going from an even layer to an odd layer of the BFS tree (or equivalently, for every path of length 2 in the BFS walk from an even layer vertex $\pi(v)_R$ to another even layer vertex $\pi(u)_R$ add a directed edge $u \rightarrow v$ to T). See Figure 2 (right to left) for a pictorial demonstration. As the BFS tree visits every vertex on the right side of G' exactly once and π is a permutation, the directed graph T will be an arborescence rooted at $\pi^{-1}(c) = r$, as desired. Moreover, from the construction it is clear that a reverse BFS walk as described in the Part (i) of the proof using (π, T) will return G' . Hence, $G' = G(\pi, T)$. ■

4 NECESSARY CONDITIONS FOR FACTORIES FOR POLYTOPES

We now begin our exploration of the general combinatorial Bernoulli factory problem: for which polytopes $\mathcal{P} \subseteq [0, 1]^n$ does there exist a Bernoulli factory for $\mathcal{P}$? In this section we provide a necessary condition: any such $\mathcal{P}$ must be the intersection of $[0, 1]^n$ with an affine subspace. Recall that an affine subspace $\mathcal{H}$ of $\mathbb{R}^d$ is a set of points x satisfying $Wx = b$ for some full-rank k -by- d matrix W and $b \in \mathbb{R}^k$ (in this case, we say the *codimension* of $\mathcal{H}$ is k).

Figure 2: An example of the bijection between pairs $(\pi, T) \in S_n \times \mathcal{T}_r(n)$ with $\pi(r) = c$ (left hand side) and r -bi-trees $G' \in \mathcal{G}_r$ (right hand side): $n = 5$, root $r = 1$, $\pi = (3, 1, 2, 5, 4)$, $c = \pi(r) = 3$, and $T = \{2 \rightarrow 1, 3 \rightarrow 1, 5 \rightarrow 3, 4 \rightarrow 3\}$; solid black edges belong to T ; dashed purple edges are matching edges corresponding to π excluding the green dashed edge $(r, \pi(r))$, i.e., $\{(u, \pi(u))\}_{u \in \{2,3,4,5\}}$;

THEOREM 4.1. *Let $\mathcal{P} \subseteq [0, 1]^n$ be a polytope such that $\mathcal{P} \cap (0, 1)^n \neq \emptyset$. If $\mathcal{P}$ is not of the form $\mathcal{P} = [0, 1]^n \cap \mathcal{H}$ for some affine subspace $\mathcal{H}$ then no Bernoulli factory for $\mathcal{P}$ exists.*

Since (non-strong) Bernoulli factories for $\mathcal{P}$ are only required to work for $x \in \mathcal{P} \cap (0, 1)^n$, the constraint that $\mathcal{P} \cap (0, 1)^n \neq \emptyset$ is necessary. For strong Bernoulli factories that work for all $x \in \mathcal{P}$, we have the following stronger theorem.

THEOREM 4.2. *Let $\mathcal{P}$ be a polytope. If $\mathcal{P}$ is not of the form $\mathcal{P} = [0, 1]^n \cap \mathcal{H}$ for some affine subspace $\mathcal{H}$, then no strong Bernoulli factory for $\mathcal{P}$ exists.*

The full proofs of Theorems 4.1 and 4.2 can be found in the online version of this paper. In the remainder of this section, we provide a sketch of the main ideas in this proof.

Before we proceed, it will prove illustrative to understand some of the obstacles to producing one-parameter Bernoulli factories for certain functions $f : [0, 1] \rightarrow [0, 1]$ (i.e., the classic Bernoulli factory setting studied in [16]). Consider, for example, the function $f(x) = |x - 0.5|$. On first glance, since $f(x) \in [0, 1]$ for all $x \in [0, 1]$, it might appear possible to construct a one-bit Bernoulli factory $\mathcal{F}$ for f . However, this is impossible. One reason why is that since $f(0.6) = 0.1 > 0$, there must be some finite sequence of coin flips where $\mathcal{F}$ outputs 1 (i.e., a leaf ℓ labelled 1 in the tree for $\mathcal{F}$ where $\Pr[\mathcal{F}(0.6) \rightarrow \ell] > 0$). But this finite sequence of coin flips must also occur with positive probability when $x = 0.5$, so $f(0.5)$ must also be strictly positive. In general, if any non-constant

Figure 3: The factory should output v at x_2 but not at x_1 .

$f(x) : [0, 1] \rightarrow [0, 1]$ achieves the value 0 or 1 on $(0, 1)$, this argument shows it is not possible to construct a Bernoulli factory for f .

Similar obstructions appear when designing Bernoulli factories for polytopes. Consider, for example, the polytope $\mathcal{P} \subseteq [0, 1]^2$ with vertices $(0, 0)$, $(0, 1)$, and $(1, 1)$ (see Figure 3) and assume to the contrary we have a Bernoulli factory $\mathcal{F}$ for $\mathcal{P}$. Let $v = (0, 1)$. Note that for a point x_2 in the interior of $\mathcal{P}$, $\mathcal{F}(x_2)$ must output v with positive probability (since x_2 cannot be written as a convex combination of the two other vertices). Similarly, for a point x_1 in the middle of the edge connecting $(0, 0)$ and $(1, 1)$, $\mathcal{F}(x_1)$ must output v with zero probability. But these two statements are incompatible; if $\mathcal{F}(x_1)$ outputs v with positive probability, there is some leaf ℓ labelled v in the protocol tree for $\mathcal{F}$ such that $\Pr[\mathcal{F}(x_2) \rightarrow \ell] > 0$. But $\Pr[\mathcal{F}(x) \rightarrow \ell]$ is just a Bernstein monomial in x , so if $\Pr[\mathcal{F}(x_2) \rightarrow \ell] > 0$ for an $x_2 \in (0, 1)^2$, it follows that $\Pr[\mathcal{F}(x_1) \rightarrow \ell] > 0$ (since x_1 also lies in $(0, 1)^2$). This means no Bernoulli factory for $\mathcal{P}$ can exist.

The general proof of Theorem 4.1 proceeds along these lines. We formalize this by looking at the *open faces* of $\mathcal{P}$. A face of $\mathcal{P}$ is a set of points in $\mathcal{P}$ which maximize a linear functional. The polytope $\mathcal{P}$ in Figure 3 contains 7 faces: one 2-dimensional face (all of $\mathcal{P}$), three 1-dimensional faces (the edges of $\mathcal{P}$) and three 0-dimensional faces (the vertices of $\mathcal{P}$). The faces of $\mathcal{P}$ form a lattice; an open face is the set of points that belong to some face in $\mathcal{P}$ but no sub-faces (e.g. the interior of $\mathcal{P}$).

We begin by showing that if there are two different open faces of $\mathcal{P}$ contained in $(0, 1)^n$, then there is no Bernoulli factory for $\mathcal{P}$. For example, the $\mathcal{P}$ in Figure 3 has two open faces that are subsets of $(0, 1)^n$: the 2-dimensional open face $\text{int}(\mathcal{P})$ and the 1-dimensional open edge between $(0, 0)$ and $(1, 1)$. The proof of this is similar to the reasoning above; if we have two points x_1, x_2 in the interior of $(0, 1)^n$ that belong to different open faces of $\mathcal{P}$, we can show that there is some vertex which must occur with positive probability

in $\mathcal{F}(x_1)$ but with zero probability in $\mathcal{F}(x_2)$. This, however, is impossible for the same reason as above (since a non-zero Bernstein monomial is positive everywhere on $(0, 1)^n$).

We then show that if the unique open face of $\mathcal{P}$ contained in $(0, 1)^n$ is the interior of $\mathcal{P}$, $\mathcal{P}$ is the intersection of $[0, 1]^n$ and an affine space. To see this, we prove the contrapositive – assume $\mathcal{P}$ is not the intersection of $[0, 1]^n$ with an affine subspace. Then look at the affine span $\mathcal{H}$ of $\mathcal{P}$, and let Q be the polytope formed by the intersection of $[0, 1]^n$ and $\mathcal{H}$. We now know $\mathcal{P}$ is strictly contained in Q – using this, we can show that there is a boundary face of $\mathcal{P}$ in the interior of Q . But then there are two open faces of $\mathcal{P}$ in the interior of $(0, 1)^n$: this boundary face and the interior of $\mathcal{P}$. Combining these two results, we arrive at Theorem 4.1. The proof of Theorem 4.2 proceeds similarly – it suffices to look at the smallest face of $[0, 1]^n$ containing $\mathcal{P}$.

5 BERNOULLI FACTORIES FOR GENERIC POLYTOPES

We will start by building Bernoulli factories for polytopes of the form $\mathcal{P} = [0, 1]^n \cap \mathcal{H}$ for generic affine spaces $\mathcal{H}$. Later, we will extend this construction to non-generic spaces. A k -dimensional affine subspace $\mathcal{H}$ can be written in the form

$$\mathcal{H} = \{x \in \mathbb{R}^n \mid Wx = b\}$$

where W is a $k \times n$ -matrix of rank k and $b \in \mathbb{R}^k$. We will let w^i denote the i -th column of W . Given a subset $S \subseteq [n]$ we define the matrix W_S to be the submatrix formed by the columns of W indexed by elements in S .

Generic subspaces. An affine subspace $\mathcal{H}$ is said to be *generic* if for each subset S of size k such that W_S is non-singular and for each subset $B \subseteq [n] \setminus S$ the solution $W_S^{-1}(b - \sum_{i \in B} w^i)$ has no coordinates in $\{0, 1\}$. Equivalently, $\mathcal{H}$ is generic if every vertex in $\mathcal{P}$ has exactly k coordinates in the open interval $(0, 1)$.

It is easy to check that for any fixed matrix W , the set of b such that $Wx = b$ is non-generic forms a set of measure zero. So, by slightly perturbing b it is always possible to obtain a generic subspace from a non-generic one. Many subspaces of interest in combinatorial optimization (e.g. k -subset, matchings, flows) are non-generic. We will later study these spaces as limits of generic affine spaces.

Vertices and partitions. It is useful to represent each vertex of the polytope $\mathcal{P}$ with a partition of the set $[n]$ into three parts, indicating which coordinates of the vertex are equal to 0, 1, or lie in the open interval $(0, 1)$. We define the set of relevant partitions as follows:

$$\text{Part}_{[n],k} \triangleq \{(A, S, B) \mid A \cup S \cup B = [n], |S| = k \text{ and } |A| + |S| + |B| = n\}$$

We define the set of *valid* partitions for the polytope $\mathcal{P}$ as:

$$\mathcal{U} \triangleq \{(A, S, B) \in \text{Part}_{[n],k} \mid \det W_S \neq 0 \text{ and } W_S^{-1}(b - \sum_{i \in B} w^i) \in (0, 1)^k\}$$

For polytopes formed from generic subspaces, there is a bijective mapping between vertices $v \in V$ and valid partitions $\pi \in \mathcal{U}$. Given a valid partition $\pi = (A, S, B)$ consider the vertex $v \in \mathcal{P}$ such that $v_i = 0$ for $i \in A$, $v_i = 1$ for $i \in B$ and $v_S = W_S^{-1}(b - \sum_{i \in B} w^i)$ (where v_S is a shorthand for the coordinates of x corresponding to indices in S). Similarly, given a vertex $v \in V$ we can represent it by the

partition $\pi = (A, S, B)$ where A corresponds to the indices where $v_i = 0$, B corresponds to the indices where $v_i = 1$ and S corresponds to the remaining indices.

Given a partition $\pi \in \mathcal{U}$ we will let $v^\pi \in V$ denote the vertex associated with this partition; likewise, given a vertex $v \in V$, we will let $\pi^v \in \mathcal{U}$ be the partition corresponding to this vertex. We write A_π, S_π and B_π to refer to the subsets in the partition π .

Factory Construction. We will construct a Bernoulli factory for $\mathcal{P}$ by defining a Bernstein polynomial for each partition $\pi = (A, S, B) \in \text{Part}_{[n],k}$:

$$P_\pi(x) \triangleq |\det W_S| \cdot \prod_{i \in A} (1 - x_i) \cdot \prod_{i \in B} x_i \cdot \prod_{i \in S} x_i (1 - x_i) \quad (13)$$

Now, for each vertex $v \in V$ we define $P_v(x)$ to be the polynomial associated with the corresponding partition $\pi^v \in \mathcal{U}$, i.e. $P_v(x) = P_{\pi^v}(x)$. At this point it is useful to note that for constructing the factory we only need P_π for $\pi \in \mathcal{U}$, but we define the polynomials more generally since they will be useful in the proof.

THEOREM 5.1. *For a generic affine subspace $\mathcal{H}$, the Bernoulli race over Bernstein polynomials given by $P_v(x) = P_{\pi^v}(x)$ is a strong Bernoulli factory for $\mathcal{P} = [0, 1]^n \cap \mathcal{H}$.*

5.1 Proof of Theorem 5.1

To show the race over our Bernstein polynomials is a strong Bernoulli factory for $\mathcal{P}$, we need to check Conditions (4) and (3) in Theorem 2.8.

5.1.1 Checking Condition (4). We start with the easier condition, that $\sum P_v(x)$ does not vanish on $\mathcal{P}$. Fix an $x \in \mathcal{P}$ and let

$$A_x = \{i \mid x_i = 0\} \quad S_x = \{i \mid 0 < x_i < 1\} \quad B_x = \{i \mid x_i = 1\}.$$

Write x as a convex combination of vertices and pick any vertex v with positive weight in this combination. Note that if $x_i = 0$, then $v_i = 0$; likewise, if $x_i = 1$, then $v_i = 1$ (since $\mathcal{P} \subseteq [0, 1]^n$). It follows that if vertex v corresponds to partition $\pi \in \mathcal{U}$, then

$$A_x \subseteq A_\pi \quad S_\pi \subseteq S_x \quad B_x \subseteq B_\pi.$$

Now, observe that

$$\begin{aligned} P_\pi(x) &= |\det W_{S_\pi}| \cdot \prod_{i \in A_\pi} (1 - x_i) \cdot \prod_{i \in B_\pi} x_i \cdot \prod_{i \in S_\pi} x_i (1 - x_i) \\ &= |\det W_{S_\pi}| \prod_{i \in S_\pi} x_i (1 - x_i) > 0. \end{aligned}$$

It follows that $\sum_{\pi \in \mathcal{U}} P_\pi(x) > 0$.

5.1.2 Rewriting Condition (3). The interesting part of the proof is to show that Condition (3) holds. Recall that Condition (3) states that

$$\sum_{v \in V} P_v(x)(v - x) = 0$$

must hold for all $x \in \mathcal{P}$. Since x and v are n -dimensional vectors, this is a vector equation. We will check this condition for each coordinate. Fix a coordinate $j \in [n]$ and split the sum $\sum_{\pi \in \mathcal{U}} P_\pi(x)(v^\pi - x)$ over all partitions $\pi \in \mathcal{U}$ depending on whether j belongs to A_π ,

B_π or S_π :

$$\sum_{\pi \in \mathcal{U} | j \in A_\pi} -x_j P_\pi(x) + \sum_{\pi \in \mathcal{U} | j \in B_\pi} (1 - x_j) P_\pi(x) + \sum_{\pi \in \mathcal{U} | j \in S_\pi} (v_j^\pi - x_j) P_\pi(x) = 0 \quad (14)$$

We will now rewrite each of the terms below as sums over partitions in $\text{Part}_{[n] \setminus j, k}$ (i.e. partitions of the set $[n] \setminus j$ into three parts (A, B, S) where S has k elements).

First term of Equation (14). Given a partition $\pi = (A, S, B) \in \mathcal{U}$ with $j \in A$ consider a partition $\pi' = (A \setminus j, S, B)$. This establishes a bijective mapping between $\{\pi \in \mathcal{U} \mid j \in A_\pi\}$ and the set:

$$\mathcal{U}_A^j \triangleq \{(A', S', B') \in \text{Part}_{[n] \setminus j, k} \mid \det W_{S'} \neq 0 \text{ and } W_{S'}^{-1}(b - \sum_{t \in B'} w^t) \in (0, 1)^k\}$$

which allows us to rewrite the first term in Equation (14) as follows:

$$\sum_{\pi \in \mathcal{U} | j \in A_\pi} -x_j P_\pi(x) = \sum_{\pi' \in \mathcal{U}_A^j} -x_j (1 - x_j) \cdot P_{\pi'}(x) \quad (15)$$

Here we define $P_{\pi'}(x)$ analogously to the definition of $P_\pi(x)$ in (13). Observe that $P_{\pi'}(x)$ does not have any terms depending on x_j ; it is a polynomial in the $(n - 1)$ other variables.

Second term of Equation (14). Similarly for the second term, we can establish a bijective mapping between $\{\pi \in \mathcal{U} \mid j \in B_\pi\}$ and the set:

$$\mathcal{U}_B^j \triangleq \{(A', S', B') \in \text{Part}_{[n] \setminus j, k} \mid \det W_{S'} \neq 0 \text{ and } W_{S'}^{-1}(b - w^j - \sum_{t \in B'} w^t) \in (0, 1)^k\}.$$

This allows us to rewrite:

$$\sum_{\pi \in \mathcal{U} | j \in B_\pi} (1 - x_j) P_\pi(x) = \sum_{\pi' \in \mathcal{U}_B^j} x_j (1 - x_j) P_{\pi'}(x). \quad (16)$$

Last term of Equation (14). Let's first examine the term $v_j^\pi - x_j$ in the last expression of Equation (14). For this, it is useful to establish a bit of additional notation. Given a set S of size k , recall that the matrix W_S is the square matrix formed by taking the columns with indices in S (in increasing order of the indices). Given coordinates $j \in S$ and $i \notin S$ we will define $W_{S[j \rightarrow i]}$ to be the matrix formed by replacing column w^j by w^i . For example, if $S = \{2, 3, 5, 7\}$ then:

$$W_S = [w^2 \ w^3 \ w^5 \ w^7] \quad \text{and} \quad W_{S[5 \rightarrow 11]} = [w^2 \ w^3 \ w^{11} \ w^7]$$

Note that the order where the i th column is inserted matters. With that, we are ready to state the next lemma:

LEMMA 5.2. *If $x \in \mathcal{P}$ and v^π is a vertex corresponding to partition $\pi = (A, S, B)$ then for any coordinate $j \in S$ we have that*

$$v_j^\pi - x_j = \sum_{i \in A} \frac{\det W_{S[j \rightarrow i]}}{\det W_S} x_i - \sum_{i \in B} \frac{\det W_{S[j \rightarrow i]}}{\det W_S} (1 - x_i).$$

PROOF. We can write the S -components of v as:

$$v_S = W_S^{-1} \left(b - \sum_{i \in B} w^i \right).$$

Since $x \in \mathcal{P}$ we know that

$$b = \sum_i w^i x_i = W_S x_S + \sum_{i \in A \cup B} w^i x_i.$$

Replacing this in the expression above we get:

$$v_S = x_S + W_S^{-1} \left(\sum_{i \in A} w^i \cdot x_i - \sum_{i \in B} w^i (1 - x_i) \right)$$

Since $j \in S$ we can look at the j th component of the expression above. Observe that the j -th component of $W_S^{-1} w^i$ can be obtained via Cramer's rule and is given by

$$[W_S^{-1} w^i]_j = \frac{\det W_{S[j \rightarrow i]}}{\det W_S}.$$

■

The previous lemma allows us to rewrite the last term in Equation (14) as follows:

$$\begin{aligned} \sum_{\pi \in \mathcal{U} | j \in S_\pi} (v_j^\pi - x_j) P_\pi(x) &= \sum_{i \neq j} \sum_{\substack{\pi \in \mathcal{U} | j \in S_\pi, \\ i \in A_\pi}} \frac{\det W_{S_\pi[j \rightarrow i]}}{\det W_{S_\pi}} \cdot x_i P_\pi(x) \\ &\quad - \sum_{i \neq j} \sum_{\substack{\pi \in \mathcal{U} | j \in S_\pi, \\ i \in B_\pi}} \frac{\det W_{S_\pi[j \rightarrow i]}}{\det W_{S_\pi}} \\ &\quad \cdot (1 - x_i) P_\pi(x) \end{aligned} \quad (17)$$

As before we will rewrite each of these terms as sums of partitions over $[n] \setminus j$. Starting with the first term, observe that for a fixed $i \neq j$ we can establish a bijective mapping between $\{\pi \in \mathcal{U} \mid j \in S_\pi \text{ and } i \in A_\pi\}$ and the set:

$$\mathcal{U}_A^i \triangleq \{(A', S', B') \in \text{Part}_{[n] \setminus j, k} \mid i \in S', \det W_{S'}[i \rightarrow j] \neq 0, \text{ and } W_{S'}^{-1}[i \rightarrow j](b - \sum_{i \in B'} w^i) \in (0, 1)^k\}$$

by mapping $\pi = (A, S, B)$ to $\pi' = (A \setminus i, S \cup i \setminus j, B)$. We now note that:

$$x_i P_\pi(x) = \frac{|\det W_S|}{|\det W_{S[j \rightarrow i]}|} x_j (1 - x_j) P_{\pi'}(x).$$

If we define $\sigma_{ij}(S)$ for a set S of size k with $i \in S$ and $j \notin S$ as

$$\sigma_{ij}(S) \triangleq \text{sign} \left(\frac{\det W_{S[i \rightarrow j]}}{\det W_S} \right) \in \{-1, 0, +1\}$$

then we can rewrite the first term in Equation (17) in the form

$$\begin{aligned} \sum_{\pi \in \mathcal{U} | j \in S_\pi, i \in A_\pi} \frac{\det W_{S_\pi[j \rightarrow i]}}{\det W_{S_\pi}} \cdot x_i P_\pi(x) &= \\ x_j (1 - x_j) \sum_{\pi' \in \mathcal{U}_A^i} \sigma_{ij}(S_{\pi'}) \cdot P_{\pi'}(x) \end{aligned} \quad (18)$$

The second term of Equation (17) can be handled similarly.

Combining the terms. We have now rewritten all the terms of Equation (14) as sums of polynomials defined over partitions of $[n] \setminus j$. Combining Equations (15), (16), (17), (18), we can rewrite (14) as

$$\sum_{\pi' \in \mathcal{U}_A^j} -P_{\pi'}(x) + \sum_{\pi' \in \mathcal{U}_B^j} P_{\pi'}(x) + \sum_{i \neq j} \left[\sum_{\pi' \in \mathcal{U}_A^i} \sigma_{ij}(S_{\pi'}) \cdot P_{\pi'}(x) - \sum_{\pi' \in \mathcal{U}_B^i} \sigma_{ij}(S_{\pi'}) \cdot P_{\pi'}(x) \right] = 0$$

after cancelling all $x_j(1 - x_j)$ terms. Our main goal is to prove this identity. It is useful to group together all partitions π' for which $S_{\pi'}$ is the same. We will then show the following lemma:

LEMMA 5.3. *For any fixed $S' \subseteq [n] \setminus j$ with $|S'| = k$ the following is an identity:*

$$\begin{aligned} & \sum_{\substack{\pi' \in \mathcal{U}_A^j \\ S_{\pi'} = S'}} -P_{\pi'}(x) + \sum_{\substack{\pi' \in \mathcal{U}_B^j \\ S_{\pi'} = S'}} P_{\pi'}(x) + \sum_{i \neq j} \sum_{\substack{\pi' \in \mathcal{U}_A^i \\ S_{\pi'} = S'}} \sigma_{ij}(S') P_{\pi'}(x) \\ & - \sum_{i \neq j} \sum_{\substack{\pi' \in \mathcal{U}_B^i \\ S_{\pi'} = S'}} \sigma_{ij}(S') P_{\pi'}(x) = 0. \end{aligned} \quad (19)$$

We will prove Lemma 5.3 by showing that each term $P_{\pi'}(x)$ appears twice in the expression, once with a positive sign and one with a negative sign. One nice aspect of focusing on a fixed S' is that the magnitude of all leading coefficients in the Bernstein monomials are the same, so we need only worry about signs of these coefficients.

Interestingly, the proof that will follow will be geometric and will be based on decompositions of zonotopes.

5.1.3 Partitions and Zonotopes. A *zonotope* is a polytope formed by the Minkowski sum of line segments. In other words, given vectors $w^1, \dots, w^k$ we will define their associated zonotope as:

$$\text{Zon}(w^1, \dots, w^k) \triangleq \{w^1 x_1 + \dots + w^k x_k \mid x_i \in [0, 1]\}$$

and the open zonotope as $\text{Zon}^0(w^1, \dots, w^k)$ as the interior of $\text{Zon}(w^1, \dots, w^k)$. Whenever $\det[w^1, \dots, w^k] \neq 0$, this is given by:

$$\text{Zon}^0(w^1, \dots, w^k) \triangleq \{w^1 x_1 + \dots + w^k x_k \mid x_i \in (0, 1)\}$$

Otherwise $\text{Zon}^0(w^1, \dots, w^k)$ is empty. Note that since each column vector $w^i \in \mathbb{R}^k$, these zonotopes are subsets of $\mathbb{R}^k$.

We can now rewrite the sets $\mathcal{U}_A^j, \mathcal{U}_B^j, \mathcal{U}_A^i$ and $\mathcal{U}_B^i$ in terms of membership in certain zonotopes. Since we are focusing on S' , let us focus on only the partitions that have S' . Let:

$$\text{Part}_{[n] \setminus j, k}(S') \triangleq \{\pi' \in \text{Part}_{[n] \setminus j, k} \mid S_{\pi'} = S'\}$$

$$\mathcal{U}_A^j(S') \triangleq \mathcal{U}_A^j \cap \text{Part}_{[n] \setminus j, k}(S')$$

and similarly for the other sets. We will refer to the columns in S' as $w^1, \dots, w^k$ (i.e., $W_{S'} = [w^1, \dots, w^k]$). Additionally, we will assume $\det W_{S'} \neq 0$ (otherwise Lemma 5.3 is trivial). Finally, associate with each partition π' the following vector:

$$q(\pi') \triangleq b - \sum_{t \in B_{\pi'}} w^t$$

We can now write our sets in terms of membership of $q(\pi')$ in a corresponding zonotope. In particular, we have that:

$$\begin{aligned} \mathcal{U}_A^j(S') &= \{\pi' \in \text{Part}_{[n] \setminus j, k}(S') \mid q(\pi') \in Z_A^j\}, \\ Z_A^j &\triangleq \text{Zon}^0(w^1, \dots, w^k) \\ \mathcal{U}_B^j(S') &= \{\pi' \in \text{Part}_{[n] \setminus j, k}(S') \mid q(\pi') \in Z_B^j\}, \\ Z_B^j &\triangleq w^j + \text{Zon}^0(w^1, \dots, w^k) \end{aligned} \quad (20)$$

For $\mathcal{U}_A^i(S')$ and $\mathcal{U}_B^i(S')$ we have the condition that $i \in S'$. Hence $\mathcal{U}_A^i(S') = \mathcal{U}_B^i(S') = \emptyset$ if $i \notin S'$ or $W_{S'}[i \rightarrow j] = 0$ and otherwise:

$$\begin{aligned} \mathcal{U}_A^i(S') &= \{\pi' \in \text{Part}_{[n] \setminus j, k}(S') \mid q(\pi') \in Z_A^i\}, \\ Z_A^i &\triangleq \text{Zon}^0(w^1, \dots, w^{i-1}, w^j, w^{i+1}, \dots, w^k) \\ \mathcal{U}_B^i(S') &= \{\pi' \in \text{Part}_{[n] \setminus j, k}(S') \mid q(\pi') \in Z_B^i\}, \\ Z_B^i &\triangleq w^i + \text{Zon}^0(w^1, \dots, w^{i-1}, w^j, w^{i+1}, \dots, w^k) \end{aligned} \quad (21)$$

When we loop over all partitions in $\pi' \in \text{Part}_{[n] \setminus j, k}(S')$, we will observe that its corresponding $q(\pi')$ either belongs to none of these zonotopes or to exactly two. In the latter case, we will show that it gets assigned opposite signs.

5.1.4 Proof of Lemma 5.3 for general k . We now prove the general case of Lemma 5.3. The main step will be to prove the following geometric lemma.

LEMMA 5.4 (TILING). *The zonotopes with a positive sign in Lemma 5.3 have disjoint interior. Moreover, the union of their closures is the zonotope $\text{Zon}(w^1, \dots, w^k, w^j)$. The same is true for all the regions with a negative sign.*

Intuitively, the (proof of the) Tiling Lemma says the following: there are $2(k+1)$ distinct terms in Lemma 5.3; $(k+1)$ of these are positive and $(k+1)$ of these are negative. Each of these terms correspond to a zonotope (in fact, a parallelotope) in $q(\pi)$ space. The $(k+1)$ positive parallelotopes partition the zonotope $\text{Zon}(w^1, \dots, w^k, w^j)$, as do the $(k+1)$ negative parallelotopes.

Given the Tiling Lemma it is straightforward to show Lemma 5.3:

PROOF OF LEMMA 5.3. Given π' in Equation (19) then it appears in a given term iff $q(\pi')$ is in the corresponding zonotope. Since all the zonotopes corresponding to positive terms are non-overlapping, it can appear at most once with a positive term.

If it does appear with a positive term, then $q(\pi') \in \text{Zon}(w^1, \dots, w^k, w^j)$. By the fact that the instance is generic, $q(\pi')$ can't be in the boundary of any of the smaller zonotopes. Since the union of zonotopes with negative sign is also $\text{Zon}(w^1, \dots, w^k, w^j)$ then it must also appear in exactly one such zonotope. Hence it also appears exactly once with negative sign in Equation (19). ■

We now devote the rest of the section to a proof of Lemma 5.4. The regions which get assigned a positive sign are i) Z_B^j , ii) Z_A^i if $\sigma_i = +1$, and iii) Z_B^i if $\sigma_i = -1$ (here we suppress j and S' in $\sigma_{ij}(S')$). It is convenient to assign to j a sign $\sigma_j = -1$ so to get a more uniform treatment of j and $\{1, \dots, k\}$; in particular, we are now simply considering all Z_B^i where $\sigma_i = -1$, and all Z_A^i where $\sigma_i = 1$. With this notation in mind, we first show that Z_B^i and Z_A^t must be disjoint if $\sigma_i \neq \sigma_t$.

LEMMA 5.5. Consider $i, t \in I \triangleq \{1, \dots, k, j\}$ with $i \neq t$. If $\sigma_i = -\sigma_t$ then regions Z_B^i and Z_A^t are disjoint.

PROOF. Consider a point x in Z_B^i . This point can be written in the form:

$$x = w^i + \sum_{r \in I \setminus \{i\}} \lambda_r w^r \quad \text{for } \lambda_r \in (0, 1)$$

Now consider a point x' in Z_A^t , which can similarly be written in the form:

$$x' = \sum_{r \in I \setminus \{t\}} \lambda'_r w^r \quad \text{for } \lambda'_r \in (0, 1)$$

Assume to the contrary that $x = x'$. We then have that

$$(1 - \lambda'_i)w^i + \lambda'_t w^t = \sum_{r \in I \setminus \{i, t\}} (\lambda'_r - \lambda_r)w^r. \quad (22)$$

Note that the term on the right hand side of (22) belongs to the hyperplane $\mathcal{H}$ spanned by the $(k-1)$ vectors w^r where $r \notin \{i, t\}$. We will now write:

$$w^i = h_i + c_i h_\perp \quad \text{and} \quad w^t = h_t + c_t h_\perp \quad (23)$$

for $h_i, h_j \in \mathcal{H}$ and $h_\perp$ a unit vector orthogonal to $\mathcal{H}$.

Observe now that c_i and c_t must have the same sign. First consider the case where neither is j . Then:

$$-1 = \frac{\sigma_i}{\sigma_t} = \text{sign} \left(\frac{\det[W' w^j w^t]}{\det[W' w^i w^j]} \right) = -\text{sign} \left(\frac{\det[W' w^j w^t]}{\det[W' w^j w^i]} \right)$$

where W' is a matrix containing all columns w^r except w^i, w^j , and w^t . Finally note that:

$$\begin{aligned} \det[W' w^j w^t] &= c_t \det[W' w^j h_\perp] \\ \det[W' w^j w^i] &= c_i \det[W' w^j h_\perp] \end{aligned}$$

which shows that $\text{sign}(c_i/c_j) = +1$. The case where one of the indices $\{i, t\}$ equals j is analogous. If t is j then we have that:

$$-1 = \frac{\sigma_i}{\sigma_j} = \sigma_i = -\text{sign} \left(\frac{\det[W' w^i]}{\det[W' w^j]} \right)$$

where now the columns of W' are formed by all the other w^r except w^i, w^j . We again reach the same conclusion that $\text{sign}(c_i/c_j) = +1$.

Now, assume without loss of generality that c_i and c_j are both positive. But then, since $(1 - \lambda'_i) > 0$ and $\lambda_t > 0$, the left hand side of (22) will have a positive $h_\perp$ component and cannot lie entirely in $\mathcal{H}$. Thus it is not possible that $x = x'$, as desired. ■

LEMMA 5.6. Consider $i, t \in I := \{1, \dots, k, j\}$ with $i \neq t$. If $\sigma_i = \sigma_t$ then regions Z_A^i and Z_A^t are disjoint. The regions Z_B^i and Z_B^t are also disjoint

PROOF. The proof follows a similar pattern as the proof of Lemma 5.6. We first choose points $x = \sum_{r \in I \setminus \{i\}} \lambda_r w^r \in Z_A^i$ and $x' = \sum_{r \in I \setminus \{t\}} \lambda'_r w^r \in Z_A^t$. Now, if $x = x'$ then $\lambda_i w^i - \lambda'_t w^t$ should be contained in the hyperplane $\mathcal{H}$ spanned by the vectors w^r for $r \neq i, t$. This again allows us to write w^i and w^t as in Equation (23), but this time with $\text{sign}(c_i/c_t) = -1$ since $\sigma_i/\sigma_t = +1$. With this sign pattern it is impossible to have $\lambda_i w^i - \lambda'_t w^t$ in $\mathcal{H}$ since it will have a non-zero $h_\perp$ component.

The argument for Z_B^i and Z_B^t is the same. ■

With these two lemmas, we are ready to prove the Tiling Lemma:

PROOF OF LEMMA 5.4. Lemmas 5.5 and 5.6 show that the regions assigned positive sign are disjoint. It is straightforward to see that they are all contained in $\text{Zon}(\{w^r\}_{r \in I})$ for $I = \{1, \dots, k, j\}$ since it is simple to express a point in each region as $\sum_{r \in I} \lambda_r w^r$ with $\lambda_r \in [0, 1]$. To show that their closures are exactly the zonotope it is enough to argue that their volumes sum up to the volume of $\text{Zon}(\{w^r\}_{r \in I})$. Note that (since the zonotopes Z_A^i and Z_B^i are actually parallelotopes generated by k vectors):

$$\text{Vol}(Z_A^i) = \text{Vol}(Z_B^i) = |\det W_{I \setminus \{i\}}|$$

where $W_{I \setminus \{i\}}$ is the matrix formed by columns w^r for $r \in I \setminus \{i\}$. Finally the formula for computing the volume of a zonotope (see e.g. [10]) is:

$$\text{Vol}(\text{Zon}(\{w^r\}_{r \in I})) = \sum_{T \subseteq I, |T|=k} |\det W_T|$$

which is equal to the sum of volumes of the smaller parallelotopes. ■

6 BERNOULLI FACTORIES FOR NON-GENERIC POLYTOPES

In this section, we demonstrate how to obtain a factory for a non-generic polytope as a limit of factories for generic polytopes. We then look at the specific case of the k -subset polytope ($\{x \in [0, 1]^n \mid \sum_i x_i = k\}$), where we recover the statistical method known as Sampford sampling.

THEOREM 6.1. Consider a polytope of the form $\mathcal{P} = [0, 1]^n \cap \mathcal{H}$, where $\mathcal{H}$ is a possibly non-generic affine subspace. Then there exist Bernstein polynomials $P_v(x)$ for each vertex v of $\mathcal{P}$ such that the corresponding Bernoulli race over Bernstein polynomials is a Bernoulli factory for $\mathcal{P}$.

This gives us a valid Bernoulli factory that always terminates as long as all coins belong to the open set $(0, 1)^n$, i.e., no coin is deterministically 0 or 1.

PROOF OF THEOREM 6.1. Let $\mathcal{H} = \{x \in \mathbb{R}^n \mid Wx = b\}$ for a $k \times n$ matrix W and a vector $b \in \mathbb{R}^k$. For each $t = 1, 2, 3, \dots$ sample a vector $b_t \in \mathbb{R}^k$ uniformly from the ball of radius $1/t$ around b and define $\mathcal{H}_t$ as the hyperplane: $\mathcal{H}_t = \{x \in \mathbb{R}^n \mid Wx = b_t\}$. Since the set $\{b \in \mathbb{R}^k \mid Wx = b \text{ is generic}\}$ has measure zero, then $\mathcal{H}_t$ is generic almost surely for all t . Let $\mathcal{P}_t = [0, 1]^n \cap \mathcal{H}_t$ and $\mathcal{U}_t$ be the set of valid partitions for polytope $\mathcal{P}_t$. Since $\mathcal{U}_t$ is a subset of the finite set $\text{Part}_{[n], k}$, there are finitely many possibilities, so one of them must occur infinitely often. Passing to this subsequence if necessary, we can assume $\mathcal{U}_t$ is the same for all t . With this, observe that the Bernoulli factory is exactly the same for all polytopes $\mathcal{P}_t$ since the polynomials in Theorem 2.8 depend on W and $\mathcal{U}$ but not directly on b . It follows that:

$$\sum_{\pi \in \mathcal{U}_t} P_\pi(x) \cdot (x - v^{t, \pi}) = 0, \quad \forall x \in \mathcal{P}_t$$

where $v^{t, \pi}$ is the vertex in $\mathcal{P}_t$ associated with partition π . Note that as $t \rightarrow \infty$ vertex $v^{t, \pi} \rightarrow v^\pi$ for the vertex v^π in $\mathcal{P}$ associated with partition π . However, note that the correspondence is no longer 1-to-1, i.e., two different partitions π_1, π_2 may map to the same vertex in $\mathcal{P}$.

Now take any $x \in \mathcal{P}$ and write it as a limit $x_t \rightarrow x$ with $x_t \in \mathcal{P}_t$. We know that: $\sum_{\pi \in \mathcal{U}_t} P_\pi(x) \cdot (x_t - v^{t,\pi}) = 0$. Taking the limit $t \rightarrow \infty$ we obtain $\sum_{\pi \in \mathcal{U}_t} P_\pi(x) \cdot (x - v^\pi) = 0$ which establishes condition (3). Note that the polynomial associated with each vertex is:

$$P_v(x) = \sum_{\pi \in \mathcal{U}_t | v=v^\pi} P_\pi(x). \quad (24)$$

Finally, observe that condition (4) is trivial for $x \in (0, 1)^n$ since all Bernstein monomials are strictly positive at such points. ■

Warning. The proof of the previous theorem (and in particular Equation (24)) give a recipe for constructing Bernoulli factories for non-generic hyperplanes. In practice, to construct a factory one can add a tiny perturbation to b , compute set $\mathcal{U}_t$ and use the formula in Equation (24).

One may be tempted to ignore the perturbation and try to apply Equation (24) directly using $\mathcal{U}$ instead of $\mathcal{U}_t$. For this, one also needs to change $(0, 1)^n$ to $[0, 1]^n$ in the definition of $\mathcal{U}$ so that all vertices are represented. This approach, however, fails. One example is the 3-by-3 perfect matching polytope. We implement³ the factory for matching both with and without the perturbation. With the perturbation we obtain a multiple⁴ of the polynomial in Section 3. If instead we don't add a perturbation, we obtain a family of polynomials that do not satisfy Equation (3).

6.1 Sampling a k -subset (Sampford Sampling)

We now show that for k -subset sampling, the recipe in Theorem 6.1 recovers the procedure known as Sampford sampling [26]. Consider the polytope

$$\mathcal{P}_{\alpha,n} = \left\{ x \in [0, 1]^n \mid \sum_{i=1}^n x_i = \alpha \right\} \quad (25)$$

for $\alpha \in (0, n)$. The vertices of $\mathcal{P}_{\alpha,n}$ are the vectors $v \in [0, 1]^n$ having $k = \lfloor \alpha \rfloor$ coordinates equal to 1, one coordinate equal to $\alpha - k$ and the remaining coordinates equal to 0.

If α is not an integer, then $\mathcal{P}_{\alpha,n}$ is generic and we can apply the construction in (13) directly. The polynomial associated with vertex $v = (\alpha - k, 1, \dots, 1, 0, \dots, 0)$ is:

$$P_v(x) = x_1(1 - x_1)x_2 \dots x_{k+1}(1 - x_{k+2}) \dots (1 - x_n)$$

Following the recipe for generic factories we obtain:

Algorithm 3 BERNOULLI FACTORY FOR $\mathcal{P}_{\alpha,n}$ FOR NON-INTEGER α (version 1)

Pick a random vertex v
 For each index such that $v_i = 1$, sample the x_i -coin and restart if it is 0.
 For each index such that $v_i = 0$, sample the x_i -coin and restart if it is 1.
 For the remaining index i sample two coins x_i -coins and restart unless their outcome is 0 and 1.
 Output vertex v

³See the code in <https://gist.github.com/renatoppl/f9151d44e8ef798737e9ce75efbf0d1d>. The implementation is done in the computational algebra system SageMath.

⁴The polynomial obtained by the generic recipe has degree $n^2 + 2n - 1$ while the one in Section 3 has degree $2n - 1$. They differ by a factor of $\prod_{ij} (1 - x_{ij})$.

We can slightly optimize this procedure by sampling the coins first and then picking a random vertex that matches the coins:

Algorithm 4 BERNOULLI FACTORY FOR $\mathcal{P}_{\alpha,n}$ FOR NON-INTEGER α (version 2)

Sample the x_i -coin for each $i \in [n]$. Let S be the indices that returned 1.
 If $|S| \neq k + 1$, restart.
 Choose $i \in S$ uniformly at random and flip the x_i -coin again. If the coin returns 1, restart.
 Output the vertex such that $v_i = \alpha - k$, $v_j = 1$ for $j \in S \setminus \{i\}$ and $v_j = 0$ for $j \notin S$.

Now, if α is an integer, the polytope $\mathcal{P}_{\alpha,n}$ becomes non-generic. According to the recipe given in Theorem 6.1 we should perturb k to $\alpha = k \pm \epsilon$ and take the limit as ϵ goes to zero. Depending on the sign of the perturbation this can lead to two different factories. One option is to look at the factory for $k - \epsilon$ and round vertices $(1 - \epsilon, 1, \dots, 1, 0, \dots, 0)$ to $(1, 1, \dots, 1, 0, \dots, 0)$. Following Equation (24) if we have a vertex v where $A = \{i | v_i = 1\}$ and $B = \{i | v_i = 0\}$ then the associated Bernstein polynomial is:

$$P_v(x) = \prod_{i \in A} x_i \cdot \prod_{i \in B} (1 - x_i) \cdot \left(k - \sum_{i \in A} x_i \right) \quad (26)$$

Then factory obtained recovers the Sampford sampling [26] procedure :

Algorithm 5 SAMPFORD SAMPLING (MINUS ϵ VERSION)

Sample the x_i -coin for each $i \in [n]$. Let S be the indices that returned 1.
 If $|S| \neq k$, restart.
 Choose $i \in S$ uniformly at random and flip the x_i -coin again. If the coin returns 1, restart.
 Output the vertex such that $v_j = 1$ for $j \in S$ and $v_j = 0$ for $j \notin S$.

A second option is to look at the factory for $k + \epsilon$ and round vertices $(\epsilon, 1, \dots, 1, 0, \dots, 0)$ to $(0, 1, 1, \dots, 1, 0, \dots, 0)$, which leads to the following polynomial:

$$P_v(x) = \prod_{i \in A} x_i \cdot \prod_{i \in B} (1 - x_i) \cdot \left(\sum_{i \in B} x_i \right) \quad (27)$$

which is an alternative implementation of Sampford sampling. The factory then becomes:

Algorithm 6 SAMPFORD SAMPLING (PLUS ϵ VERSION)

Sample the x_i -coin for each $i \in [n]$. Let S be the indices that returned 1.
 If $|S| \neq k + 1$, restart.
 Choose $i \in S$ uniformly at random and flip the x_i -coin again. If the coin returns 1, restart.
 Output the vertex such that $v_j = 1$ for $j \in S \setminus \{i\}$ and $v_j = 0$ for $j \notin S \cup \{i\}$.

Note that polynomials (26) and (27) are different polynomials and hence lead to different algorithms, but evaluate the same within the polytope $\mathcal{P}_k$.

Can we terminate on the vertices? One interesting observation is that for both polynomials (26) and (27), it is the case that for any vertex $v' \in V$ we have $\sum_{v \in V} P_v(v') = 0$. This means that although this factory terminates a.s. in the interior of $\mathcal{P}_{k,n}$ it does not terminate on the vertices (one can also directly check that the algorithms described above never terminate in the case that x is a vertex of $\mathcal{P}_{k,n}$). Hence these are Bernoulli factories but not strong Bernoulli factories.

Can this be fixed? For the special cases of $k = 1$ and $k = n - 1$ it is possible to obtain a strong Bernoulli factory. In the case $k = 1$ take $P_i(x) = x_i$ (where P_i is the polynomial corresponding to the vertex e_i with 1 in the i -th coordinate). It is easy to verify that $\sum_i x_i(x - e_i) = 0$ since $\sum_i x_i = 1$ and $\sum_i x_i e_i = x$. Similarly for $k = n - 1$ we can take $P_i(x) = 1 - x_i$ (where now P_i corresponds to the vertex $1 - e_i$ with 1 in each coordinate except i).

ACKNOWLEDGMENTS

The authors would like to thank Shaddin Dughmi, Jan Vondrak, Jason Hartline, and Bobby Kleinberg for several helpful insights and comments at early phases of this work. The first author was visiting Google Research NYC during parts of this project.

REFERENCES

- [1] Nima Anari, Shayan Oveis Gharan, and Cynthia Vinzant. 2018. Log-concave polynomials, entropy, and a deterministic approximation algorithm for counting bases of matroids. In *2018 IEEE 59th Annual Symposium on Foundations of Computer Science (FOCS)*. IEEE, 35–46.
- [2] Søren Asmussen, Peter W Glynn, and Hermann Thorisson. 1992. Stationarity detection in the initial transient problem. *ACM Transactions on Modeling and Computer Simulation (TOMACS)* 2, 2 (1992), 130–157.
- [3] Jose Blanchet and Fan Zhang. 2017. Exact simulation for multivariate Itô diffusions. *arXiv preprint arXiv:1706.05124* (2017).
- [4] Yang Cai, Argyris Oikonomou, Grigoris Velezgas, and Mingfei Zhao. 2019. An Efficient epsilon-BIC to BIC Transformation and Its Application to Black-Box Reduction in Revenue Maximization. *arXiv preprint arXiv:1911.10172* (2019).
- [5] Howard Dale, David Jennings, and Terry Rudolph. 2015. Provable quantum advantage in randomness processing. *Nature communications* 6, 1 (2015), 1–4.
- [6] Shaddin Dughmi, Jason D Hartline, Robert Kleinberg, and Rad Niazadeh. 2017. Bernoulli factories and black-box reductions in mechanism design. In *Proceedings of the 49th Annual ACM SIGACT Symposium on Theory of Computing*. 158–169.
- [7] James MFlegal, Radu Herbei, et al. 2012. Exact sampling for intractable probability distributions via a Bernoulli factory. *Electronic Journal of Statistics* 6 (2012), 10–37.
- [8] Flávio B Gonçalves, Krzysztof Łatuszyński, Gareth O Roberts, et al. 2017. Barker's algorithm for Bayesian inference with intractable likelihoods. *Brazilian Journal of Probability and Statistics* 31, 4 (2017), 732–745.
- [9] Flávio B Gonçalves, Krzysztof Łatuszyński, and Gareth O Roberts. 2017. Exact Monte Carlo likelihood-based inference for jump-diffusion processes. *arXiv preprint arXiv:1707.00332* (2017).
- [10] Eugene Gover and Nishan Krikorian. 2010. Determinants and the volumes of parallelotopes and zonotopes. *Linear Algebra Appl.* 433, 1 (2010), 28–40.
- [11] Vineet Goyal and Karl Sigman. 2012. On simulating a class of Bernstein polynomials. *ACM Transactions on Modeling and Computer Simulation (TOMACS)* 22, 2 (2012), 1–5.
- [12] Radu Herbei and L Mark Berliner. 2014. Estimating ocean circulation: an MCMC approach with approximated likelihoods via the Bernoulli factory. *J. Amer. Statist. Assoc.* 109, 507 (2014), 944–954.
- [13] Mark Huber. 2016. Nearly Optimal Bernoulli Factories for Linear Functions. *Combinatorics, Probability and Computing* 25, 4 (2016), 577–591.
- [14] Mark Huber. 2017. Optimal linear Bernoulli factories for small mean problems. *Methodology and Computing in Applied Probability* 19, 2 (2017), 631–645.
- [15] Mark Jerrum and Alistair Sinclair. 1996. The Markov chain Monte Carlo method: an approach to approximate counting and integration. *Approximation Algorithms for NP-hard problems*. PWS Publishing (1996).
- [16] MS Keane and George L O'Brien. 1994. A Bernoulli factory. *ACM Transactions on Modeling and Computer Simulation (TOMACS)* 4, 2 (1994), 213–219.
- [17] Gustav Kirchhoff. 1847. Ueber die Auflösung der Gleichungen, auf welche man bei der Untersuchung der linearen Vertheilung galvanischer Ströme geführt wird. *Annalen der Physik* 148, 12 (1847), 497–508.
- [18] Krzysztof Łatuszyński, Ioannis Kosmidis, Omiros Papaspiliopoulos, and Gareth O Roberts. 2011. Simulating events of unknown probabilities via reverse time martingales. *Random Structures & Algorithms* 38, 4 (2011), 441–452.
- [19] Luis Mendo. 2019. An asymptotically optimal Bernoulli factory for certain functions that can be expressed as power series. *Stochastic Processes and their Applications* 129, 11 (2019), 4366–4384.
- [20] Giulio Morina, Krzysztof Łatuszyński, Piotr Nayar, and Alex Wendland. 2019. From the Bernoulli factory to a dice enterprise via perfect sampling of Markov chains. *arXiv preprint arXiv:1912.09229* (2019).
- [21] Elchanan Mossel, Yuval Peres, et al. 2005. New coins from old: computing with unknown bias. *Combinatorica* 25, 6 (2005), 707–724.
- [22] Şerban Nacu and Yuval Peres. 2005. Fast simulation of new coins from old. *The Annals of Applied Probability* 15, 1A (2005), 93–115.
- [23] Raj B Patel, Terry Rudolph, and Geoff J Pryde. 2019. An experimental quantum Bernoulli factory. *Science advances* 5, 1 (2019), eaau6668.
- [24] James Propp and David Wilson. 1998. Coupling from the past: a user's guide. *Microsurveys in discrete probability* 41 (1998), 181–192.
- [25] James Gary Propp and David Bruce Wilson. 1996. Exact sampling with coupled Markov chains and applications to statistical mechanics. *Random Structures & Algorithms* 9, 1-2 (1996), 223–252.
- [26] MR Sampford. 1967. On sampling without replacement with unequal probabilities of selection. *Biometrika* 54, 3-4 (1967), 499–513.
- [27] Sebastian M Schmon, Arnaud Doucet, and George Deligiannidis. 2019. Bernoulli Race Particle Filters. In *The 22nd International Conference on Artificial Intelligence and Statistics*. 2350–2358.
- [28] Alexander Schrijver. 2003. *Combinatorial optimization: polyhedra and efficiency*. Vol. 24. Springer Science & Business Media.
- [29] Mohit Singh and Nisheeth K Vishnoi. 2014. Entropy, optimization and counting. In *Proceedings of the forty-sixth annual ACM symposium on Theory of computing*. 50–59.
- [30] Damian Straszak and Nisheeth K Vishnoi. 2017. Real stable polynomials and matroids: Optimization and counting. In *Proceedings of the 49th Annual ACM SIGACT Symposium on Theory of Computing*. 370–383.
- [31] WT Tutte. 1948. The dissection of equilateral triangles into equilateral triangles. In *Mathematical Proceedings of the Cambridge Philosophical Society*, Vol. 44. Cambridge University Press, 463–482.
- [32] John Von Neumann. 1951. 13. various techniques used in connection with random digits. *Appl. Math Ser* 12, 36-38 (1951), 5.
- [33] Xiao Yuan, Ke Liu, Yuan Xu, Weiting Wang, Yuwei Ma, Fang Zhang, Zhaopeng Yan, R Vijay, Luyan Sun, and Xiongfeng Ma. 2016. Experimental quantum randomness processing using superconducting qubits. *Physical review letters* 117, 1 (2016), 010502.