Electronic Colloquium on Computational Complexity, Report No. 190 (2023)

# From Trees to Polynomials and Back Again: New Capacity Bounds with Applications to TSP

Leonid Gurvits, Nathan Klein, and Jonathan Leake

November 15, 2023

**Abstract**

We give simply exponential lower bounds on the probabilities of a given strongly Rayleigh distribution, depending only on its expectation. This resolves a weak version of a problem left open by Karlin-Klein-Oveis Gharan in their recent breakthrough work on metric TSP, and this resolution leads to a minor improvement of their approximation factor for metric TSP. Our results also allow for a more streamlined analysis of the algorithm.

To achieve these new bounds, we build upon the work of Gurvits-Leake on the use of the productization technique for bounding the capacity of a real stable polynomial. This technique allows one to reduce certain inequalities for real stable polynomials to products of affine linear forms, which have an underlying matrix structure. In this paper, we push this technique further by characterizing the worst-case polynomials via bipartitioned forests. This rigid combinatorial structure yields a clean induction argument, which implies our stronger bounds.

In general, we believe the results of this paper will lead to further improvement and simplification of the analysis of various combinatorial and probabilistic bounds and algorithms.

ISSN 1433-8092

# 1 Introduction

The theory of real stable and log-concave polynomials has seen numerous applications in combinatorics and theoretical computer science (TCS). This includes bounds and approximation algorithms for various combinatorial quantities [Gur06a; Bar16; AO17; SV17; Ana+18b; AOV18; Ali+21; BLP23], proofs of long-standing log-concavity and sampling conjectures related to matroids [AHK18; Ana+18a; Ana+19; BH20], proofs of the Kadison-Singer conjecture and generalizations [MSS15; AO14; Brä18], an improved approximation factor for the traveling salesperson problem (TSP) [OSS11; KKO21; KKO23], and many more. The power of these polynomial classes comes from two features: (1) their *robustness*, shown in the fact that many natural operations preserve these log-concavity properties, and (2) their *convex analytic properties*, which can be used to prove bounds and other analytic statements. The typical way these polynomials are utilized is by encoding combinatorial objects as real stable and log-concave polynomials, which essentially allows these operations and convexity properties to automatically transfer to the combinatorial objects. This idea, while simple, has led to important breakthroughs in combinatorics, TCS, and beyond.

For example, [Gur06a] utilized real stable polynomials to give a new proof of the Van der Waerden conjecture on the permanent of a doubly stochastic matrix (originally due to [Ego81; Fal81]). This proof led to a vast generalization of Van der Waerden bound, including an improved Schrijver's bound for regular bipartite graphs [Gur06a], an analogous bound for mixed discriminants [Gur06b], and an analogous bound for mixed volumes that led to the development of strongly log-concave polynomials [Gur09a; Gur09b]. One reason the original bound was historically so difficult to prove is a lack of a usable inductive structure coming from the matrices themselves. One of the key insights of the new proof was to use the simple inductive structure of real stable polynomials given by partial derivatives. By encoding the matrices as polynomials, the correct induction becomes straightforward, and the bound follows from a simple calculus argument.

More recently, the approximation factor improvement for the metric traveling salesperson problem (TSP) crucially utilized real stable polynomials [KKO21; KKO23]. The idea is to encode certain discrete probability distributions related to spanning trees as real stable polynomials. The coefficients of these polynomials give probabilities of certain graph-theoretic events (e.g., the number of edges in a given spanning tree incident on a particular vertex), and analytic properties of real stable polynomials allow one to lower bound these probabilities. This in turn implies bounds on the expected cost of the output of a randomized algorithm for metric TSP.

In this paper, we improve upon the polynomial capacity bounds of [GL21], and our applications touch on the two problems discussed above. Specifically, we give:

1. robust coefficient lower bounds for all (not necessarily homogeneous) real stable polynomials,
2. simply exponential lower bounds on probabilities of strongly Rayleigh distributions (solving a weak version of an open problem of [KKO21]), and
3. a further improvement to the approximation factor for metric TSP (predicted by [GL21]).

Interestingly, our approach goes in the opposite direction to that discussed above. Our technical results answer questions regarding real stable polynomials, but to prove these results we use various graph and matrix structures inherent to the polynomials. In [GL21], this was seen in the "productization" technique: bounds on real stable polynomials were achieved by showing that the worst-case bounds come from polynomials associated to certain matrices. In this paper, we push this idea further by showing that these worst-case matrices are bipartite adjacency matrices of forests. This very rigid structure enables a clean induction argument, which implies stronger polynomial capacity bounds. These new bounds lead to the applications discussed above, with the strongest bounds implying the metric TSP improvement.

# 2 Main Results

We first state here our main technical results; see Section 3 for any undefined notation.

Our first main result is a non-homogeneous version of Theorem 2.1 of [GL21] which implies robust coefficient lower bounds for all real stable polynomials as a direct corollary. Crucially, these bounds do not

depend on the total degree of the polynomial. This was one of the main barriers to applying the results of [GL21] to metric TSP.

**Theorem 2.1** (Main non-homogeneous capacity bound)**.** *Let $p \in \mathbb{R}_{\geq 0}[x_1, \ldots, x_n]$ be a real stable polynomial in $n$ variables, and fix any $\boldsymbol{\kappa} \in \mathbb{Z}^n$ with non-negative entries. If $p(\mathbf{1}) = 1$ and $\|\boldsymbol{\kappa} - \nabla p(\mathbf{1})\|_1 < 1$, then*

$$\inf_{x_1,\ldots,x_n>0} \frac{p(\boldsymbol{x})}{x_1^{\kappa_1} \cdots x_n^{\kappa_n}} \geq (1 - \|\boldsymbol{\kappa} - \nabla p(\mathbf{1})\|_1)^n .$$

*This bound is tight for any fixed $\boldsymbol{\kappa}$ with strictly positive entries.*

**Corollary 2.2** (Main non-homogeneous coefficient bound)**.** *Let $p \in \mathbb{R}_{\geq 0}[x_1, \ldots, x_n]$ be a real stable polynomial in $n$ variables, and fix any $\boldsymbol{\kappa} \in \mathbb{Z}^n$ with non-negative entries. If $p(\mathbf{1}) = 1$ and $\|\boldsymbol{\kappa} - \nabla p(\mathbf{1})\|_1 < 1$, then*

$$p_{\boldsymbol{\kappa}} \geq \left( \prod_{i=1}^{n} \frac{\kappa_i^{\kappa_i} e^{-\kappa_i}}{\kappa_i!} \right) (1 - \|\boldsymbol{\kappa} - \nabla p(\mathbf{1})\|_1)^n,$$

*where $p_{\boldsymbol{\kappa}}$ is the coefficient of $\boldsymbol{x}^{\boldsymbol{\kappa}}$ in $p$. The dependence on $(1 - \|\boldsymbol{\kappa} - \nabla p(\mathbf{1})\|_1)$ is tight for any fixed $\boldsymbol{\kappa}$ with strictly positive entries.*

The above results[1] are robust (i.e., resilient to $\ell^1$ perturbations) versions of the results utilized to bound various combinatorial and probabilistic quantities, as discussed above. That said, they are still not quite strong enough to imply an improvement to the metric TSP approximation factor. To obtain this improvement, we resolve a weak version of an open problem from [KKO21], which we discuss below. Stronger versions of Theorem 2.1 and Corollary 2.2 which imply this result can be found in Section 5.

## 2.1 Application: Metric TSP

We first recall an important probabilistic bound from [KKO21] used in the analysis of their metric TSP approximation algorithm (which is a slight modification of the max entropy algorithm from [OSS11] first studied by [Asa+10]). In what follows, we let $A_S := \sum_{i \in S} A_i$ and $\kappa_S := \sum_{i \in S} \kappa_i$. See Section 3 for any undefined notation.

**Theorem 2.3** (Prop. 5.1 of [KKO21])**.** *Let $\mu$ be a strongly Rayleigh distribution on $[m]$, let $A_1, \ldots, A_n$ be random variables counting the number of elements contained in disjoint subsets of $[m]$, and fix $\boldsymbol{\kappa} \in \mathbb{Z}^n$ with non-negative entries. Suppose for all $S \subseteq [n]$ we have*

$$\mathbb{P}_\mu [A_S \geq \kappa_S] \geq \epsilon \qquad \text{and} \qquad \mathbb{P}_\mu [A_S \leq \kappa_S] \geq \epsilon.$$

*Then we have*

$$\mathbb{P}_\mu [A_1 = \kappa_1, \ldots, A_n = \kappa_n] \geq f(\epsilon) \cdot \mathbb{P}_\mu \left[A_{[n]} = \kappa_{[n]}\right],$$

*where $f(\epsilon) \geq \epsilon^{2^n} \prod_{i=2}^{n} \frac{1}{\max\{\kappa_i, \kappa_{[i-1]}\}+1}$.*

In [KKO21], the authors note two things about this bound. First, they note that to apply the bound it is sufficient to have

$$\left|\mathbb{E}_\mu [A_S] - \kappa_S\right| < 1 \qquad \forall S \subseteq [n], \tag{1}$$

since this implies a lower bound on $\mathbb{P}_\mu [A_S = \kappa_S]$ for all $S \subseteq [n]$ for strongly Rayleigh distributions. Second, they note that the bound on $f(\epsilon)$ is doubly exponential in $n$, but they expect the true dependency to only be simply exponential. They leave it as an open problem to determine a tight lower bound on $f(\epsilon)$.

In this paper, we further improve the metric TSP approximation factor by resolving a weak version of this open problem: we give a simply exponential lower bound which depends tightly on $\epsilon$, under the stronger condition of (1). Concretely, we prove the following.

[1] Note that these bounds already appear in the arXiv version of [GL21], but not in the STOC version.

**Theorem 2.4** (Improved probability lower bound)**.** *Let $\mu$ be a strongly Rayleigh distribution on $[m]$, let $A_1, \dots, A_n$ be random variables counting the number of elements contained in disjoint subsets of $[m]$, and fix $\boldsymbol{\kappa} \in \mathbb{Z}^n$ with non-negative entries. Suppose for all $S \subseteq [n]$ we have*

$$\left|\mathbb{E}_\mu [A_S] - \kappa_S\right| \leq 1 - \epsilon.$$

*Then we have*

$$\mathbb{P}_\mu [A_1 = \kappa_1, \dots, A_n = \kappa_n] \geq \epsilon^n \prod_{\kappa_i > 0} \frac{1}{e\sqrt{\kappa_i}}.$$

*The dependence on $\epsilon$ is tight for any fixed $\boldsymbol{\kappa}$ with strictly positive entries.*

In fact, we prove stronger versions of this result which more directly depend on the specific values of $(\mathbb{E}_\mu [A_S] - \kappa_S)$ for all $S \subseteq [n]$; see Theorem 4.1 and Corollary 4.2. These results are analogous to our main coefficient bound Corollary 2.2 because the coefficients and the gradient of probability generating polynomials can be interpreted as the probabilities and the expectation of the associated distribution. That said, our stronger probabilistic results require a more delicate analysis of the expectations (gradient) beyond what is required for Corollary 2.2. In particular, note that the conditions on the expectations in Theorem 2.4 are more general than a bound on the $\ell^1$ norm of $(\mathbb{E}_\mu [A_i] - \kappa_i)_{i=1}^n$. See Section 4 for further details.

Using Theorem 2.4, we improve the metric TSP approximation factor for the algorithm given in [KKO21].

**Theorem 2.5.** *There exists a randomized algorithm for metric TSP with approximation factor $\frac{3}{2} - \epsilon$ for some $\epsilon > 10^{-34}$.*

The guarantee in [KKO22] was not much smaller: there they showed $\epsilon > 10^{-36}$. Thus our improvement in terms of the approximation factor itself is perhaps weaker than what would be expected, given that we were able to improve the probability bound in Theorem 2.4 to a simply exponential factor. The reason for this is that while Theorem 2.3 was useful in [KKO21] to quickly determine which events occurred with constant probability (and indeed provided a single unifying explanation for why one should expect many of their probabilistic bounds to hold), it gave such small guarantees that [KKO21] resorted to ad hoc arguments instead to give their final probabilistic bounds.

We show that Theorem 2.4 alone can be used to give bounds that are comparable to the ad hod methods of [KKO21] (and, in several important cases, much better) whenever the bounds came purely from information on the expectations as in (1). Thus, we believe our main contribution to work on metric TSP is a version of Theorem 2.3 that is "reasonable" to use, allowing one to show a similar approximation factor but with a more streamlined proof.

Unfortunately, not all of the bounds in [KKO21] follow from expectation information, and two of them become bottlenecks for improving the approximation factor after applying Theorem 2.4 to the other statements. Thus, to demonstrate Theorem 2.5 we need to sharpen these bounds using other techniques. For one of these lemmas we show that the existing proof in [KKO21] was far from tight, and in the other we refine their proof. In particular, using Theorem 2.4, we show we can reduce this second lemma to a special case that is possible to analyze more carefully. See Section 4 for further details.

# 3 Technical Overview

In this section we discuss the proof strategy of our main capacity and coefficient bounds, Theorem 2.1 and Corollary 2.2, and their stronger forms.

**Notation.** Given a vector $\boldsymbol{z} \in \mathbb{R}^E$ and a subset $S$ of $E$, let $\boldsymbol{z}^S := \prod_{e \in S} z_e$. Let $\mu : \{0,1\}^E \to \mathbb{R}$ be a probability distribution over subsets of $E$. The generating polynomial $g_\mu \in \mathbb{R}_{\geq 0}[\{z_e\}_{e \in E}]$ of $\mu$ is defined as

$$g_\mu(z) := \sum_{S \subseteq E} \mu(S) \cdot \boldsymbol{z}^S.$$

The distribution $\mu$ is **strongly Rayleigh** if $g_\mu$ is real stable, where a polynomial $p \in \mathbb{R}[z_1, \dots, z_n]$ is **real stable** if $p(\boldsymbol{z}) \neq 0$ whenever $\Im(z_i) > 0$ for all $i \in [n]$ (i.e., when all inputs are in the complex upper half-plane).

See [BBL09] for much more on strongly Rayleigh measures. Further, given a polynomial $p \in \mathbb{R}_{\geq 0}[x_1, \ldots, x_n]$ and $\boldsymbol{\kappa} \in \mathbb{Z}^n_{\geq 0}$, the **capacity** of $p$ is defined as

$$\mathrm{Cap}_{\boldsymbol{\kappa}}(p) := \inf_{\boldsymbol{x}>0} \frac{p(\boldsymbol{x})}{\boldsymbol{x}^{\boldsymbol{\kappa}}}.$$

Finally, we let $p_{\boldsymbol{\kappa}}$ denote the coefficient of $\boldsymbol{x}^{\boldsymbol{\kappa}}$ in $p$.

**Conceptual strategy.** We first give an overarching view of the strategy used to prove our main results, as well as the key similarities and differences compared to that of [GL21]. The general idea for proving our bounds is to find a simple and sparse underlying structure for the worst-case inputs. The space of all real stable polynomials can be complicated, but we show that the worst-case polynomials for our bounds are far simpler: they are "sparse" products of affine linear forms. More concretely, we reduce the space of input polynomials (and the corresponding combinatorial structures) as follows:

| | | | | |
|---|---|---|---|---|
| real stable polynomials | $\Longrightarrow$ | products of affine linears | $\Longrightarrow$ | sparse products of affine linears |
| matroids | $\Longrightarrow$ | matrices | $\Longrightarrow$ | forests |

The first reduction step uses the idea of productization which was the key idea from [GL21]. This allows for one to utilize the matrix structure inherent to products of affine linear forms.

The second reduction step is then new to this paper. We first show that we may restrict to the extreme points of the set of matrices corresponding to products of affine linear forms, and then we show that these extreme matrices are supported on the edges of forests. This implies a significant decrease in density of the matrices: general graphs can have quadratically many edges, whereas forests can only have linearly many. This allows for an intricate but clean induction on the leaf vertices of these forests, which yields the strongest bounds of this paper. Additionally, it is this step that allows for bounds which do not depend on the total degree of the polynomial, and this was a crucial barrier to applying the bounds of [GL21] to metric TSP.

## 3.1 Conceptual Strategy, in More Detail

We now go through the steps of the conceptual strategy described above in more detail. Let us first restate our main capacity and coefficient bounds.

**Theorem 3.1** (= Theorem 2.1 and Corollary 2.2)**.** *Let $p \in \mathbb{R}_{\geq 0}[x_1, \ldots, x_n]$ be a real stable polynomial in $n$ variables, and fix any $\boldsymbol{\kappa} \in \mathbb{Z}^n$ with non-negative entries. If $p(\mathbf{1}) = 1$ and $\|\boldsymbol{\kappa} - \nabla p(\mathbf{1})\|_1 < 1$, then*

$$\mathrm{Cap}_{\boldsymbol{\kappa}}(p) \geq (1 - \|\boldsymbol{\kappa} - \nabla p(\mathbf{1})\|_1)^n$$

*and*

$$p_{\boldsymbol{\kappa}} \geq \left( \prod_{i=1}^{n} \frac{\kappa_i^{\kappa_i} e^{-\kappa_i}}{\kappa_i!} \right) (1 - \|\boldsymbol{\kappa} - \nabla p(\mathbf{1})\|_1)^n,$$

*where $\mathrm{Cap}_{\boldsymbol{\kappa}}(p) := \inf_{\boldsymbol{x}>0} \frac{p(\boldsymbol{x})}{\boldsymbol{x}^{\boldsymbol{\kappa}}}$ is the **capacity** of $p$ and $p_{\boldsymbol{\kappa}}$ is the coefficient of $\boldsymbol{x}^{\boldsymbol{\kappa}}$ in $p$. The dependence on $(1 - \|\boldsymbol{\kappa} - \nabla p(\mathbf{1})\|_1)$ in these bounds is tight for any fixed $\boldsymbol{\kappa}$ with strictly positive entries.*

Analogous bounds required for the metric TSP application then follow from interpreting desired quantities as the coefficients and gradient of certain real stable polynomials. Specifically, the strongly Rayleigh probabilities we wish to lower bound are the coefficients of the corresponding real stable generating polynomial, and the expectations of the associated random variables are given by the gradient of that polynomial. We leave further details to Section 4.

We now discuss the proof of Theorem 3.1. First note that the coefficient bound follows from the capacity bound. This immediately follows from Corollary 3.6 of [Gur09b], which implies

$$p_{\boldsymbol{\kappa}} \geq \left( \prod_{i=1}^{n} \frac{\kappa_i^{\kappa_i} e^{-\kappa_i}}{\kappa_i!} \right) \mathrm{Cap}_{\boldsymbol{\kappa}}(p). \tag{2}$$

Thus what remains to be proven is the capacity bound

$$\mathrm{Cap}_{\boldsymbol{\kappa}}(p) \geq (1 - \|\boldsymbol{\kappa} - \nabla p(\mathbf{1})\|_1)^n,$$

which is precisely the bound of Theorem 2.1, as well as its tightness, which follows from considering a particular example of $p$ (see Lemma 5.14).

The remainder of the proof then has four main steps. We also note here that these proof steps actually imply stronger bounds than Theorem 2.1, see Corollary 5.13 for the formal statement. These stronger bounds are required for the metric TSP application.

**Step 1: Reduce to products of affine linear forms via productization.** We first generalize the productization technique of [GL21] to non-homogeneous real stable polynomials. The upshot of this technique is that it implies it is sufficient to prove Theorem 2.1 for products of affine linear forms with non-negative coefficients (see Corollary 5.4). Such polynomials correspond to $d \times (n+1)$ $\mathbb{R}_{\geq 0}$-valued matrices with row sums $\mathbf{1}$ and column sums $\boldsymbol{\alpha}$ equal the entries of the gradient of the polynomial, via

$$\phi : A \mapsto \prod_{i=1}^{d} \left( a_{i,n+1} + \sum_{j=1}^{n} a_{i,j} x_j \right).$$

This gives far more structure to work with, beyond that of real stable polynomials in general. This part is a straightforward generalization of the analogous result of [GL21].

**Step 2: Reduce to extreme points.** The set of $\mathbb{R}_{\geq 0}$-valued matrices with row sums $\mathbf{1}$ and column sums $\boldsymbol{\alpha}$ forms a convex polytope $P^d_{\boldsymbol{\alpha}}$, and thus the polynomials we now must consider correspond to the points of this polytope via the map $\phi$ defined above. Inspired by a result of Barvinok (see Lemma 5.5), we next show that the function

$$A \mapsto \mathrm{Cap}_{\kappa}(\phi(A))$$

is log-concave on the above described polytope. Since we want to minimize the capacity, this implies we may further restrict to the polynomials associated to the extreme points of the polytope.

**Step 3: Extreme points correspond to bipartitioned forests.** Any $\mathbb{R}_{\geq 0}$-valued matrix $A$ can be interpreted as the weighted bipartite adjacency matrix of a bipartite graph, where the left vertices correspond to the rows of $A$ and the right vertices correspond to the columns of $A$. A matrix $A \in P^d_{\boldsymbol{\alpha}}$ being extreme implies the associated bipartite graph has no cycles. This implies the associated bipartite graph is a forest (see Lemma 5.6). The sparsity properties of such matrices implies a simple structure for the associated polynomials, which is particularly amenable to an intricate but clean induction.

**Step 4: Induction on leaf vertices of the bipartitioned forests.** Leaf vertices in the forest corresponding to a given matrix $A \in P^d_{\boldsymbol{\alpha}}$ indicate rows or columns of the matrix $A$ which have exactly one non-zero entry. If a row of $A$ has exactly one non-zero entry, the induction proceeds in a straightforward fashion, by simply removing the corresponding row of $A$ and recalculating the column sums (see the $d \geq n+1$ case of the proof of Theorem 5.11).

If a column of $A$ (say column $i$) has exactly one non-zero entry, then the induction is more complicated. We prove lemmas showing how much the capacity can change after applying the partial derivative $\partial_{x_i}$ (when $\kappa_i \geq \alpha_i$, see Lemma 5.10) or setting $x_i$ to 0 (when $\kappa_i < \alpha_i$, see Lemma 5.9). Since column $i$ has only one entry, applying $\partial_{x_i}$ corresponds to removing column $i$ and the row of $A$ which contains the non-zero entry, and setting $x_i$ to 0 corresponds to removing column $i$. After renormalizing the row sums and recalculating the column sums, the proof again proceeds by induction. (See the proof of Theorem 5.11 to see the above arguments presented formally.) Example 5.15 and Example 5.16 show that the distinction between the $\kappa_i \geq \alpha_i$ and $\kappa_i < \alpha_i$ cases is not an artifact of the proof.

**Some comments on tightness of the bounds.** The main coefficient bound of Corollary 2.2 is proven via two different bounds, as discussed at the beginning of this section. That is, one first bounds the coefficient in terms of the capacity (2) via Corollary 3.6 of [Gur09b], and then one bounds the capacity via the steps outlined above. Thus while the capacity bound (Theorem 2.1) is tight for $\boldsymbol{\kappa} > \mathbf{0}$, the coefficient bound (Corollary 2.2) may not be. We note that the coefficient bound is likely close to tight in the case that $(1 - \|\boldsymbol{\kappa} - \nabla p(\mathbf{1})\|_1)$ is close to 1, but it seems this tightness deteriorates as $(1 - \|\boldsymbol{\kappa} - \nabla p(\mathbf{1})\|_1)$ gets close to 0. That said, the dependence on $(1 - \|\boldsymbol{\kappa} - \nabla p(\mathbf{1})\|_1)$ in Theorem 2.1 and Corollary 2.2 is tight for $\boldsymbol{\kappa} > \mathbf{0}$ by Lemma 5.14, though there is still potential for improvement in the more refined capacity bounds; see Section 5.5.

That said, tight coefficient lower bounds in the univariate case can be achieved by directly applying Hoeffding's theorem (Theorem A.9), and these lower bounds resemble the coefficient lower bound of Corollary 2.2. Thus one can view Corollary 2.2 as a step towards a multivariate generalization of Hoeffding's theorem. It is an interesting question whether or not the techniques used here can be extended to a full multivariate generalization of Hoeffding's theorem.

### 3.2 Example: The Univariate Case

In this section, we demonstrate the proof of Theorem 2.1 in the univariate case. This will serve as a sort of proof of concept for the more general proof.

Here we consider real stable non-homogeneous polynomials $p \in \mathbb{R}_{\geq 0}[x_1]$ such that $p(1) = 1$ and $\nabla p(1) = \alpha_1$, and we define $\epsilon := 1 - |\alpha_1 - \kappa_1| > 0$. By Step 1 above, we may assume that $p$ is of the form

$$p(x_1) = \prod_{i=1}^{d} (a_{i,1}x_1 + a_{1,2}),$$

where $A$ is an $\mathbb{R}_{\geq 0}$-valued $d \times 2$ matrix with row sums 1 and column sums $(\alpha_1, d - \alpha_1)$. By Steps 2-3, we may further assume $A$ is the weighted bipartite adjacency matrix of a forest. If every row of $A$ contains exactly one non-zero entry then $p(x_1) = x_1^{k_1}$, and the result is trivial in this case. Otherwise, $d - 1$ rows of $M$ have exactly one non-zero entry (see Lemma 5.7). Thus for some $k \leq d - 1$ we have

$$p(x_1) = x_1^k(ax_1 + b),$$

where $a + b = 1$ and $a + k = \alpha_1$. Since $\kappa_1 \in \mathbb{Z}_{\geq 0}$ and

$$\epsilon = 1 - |\alpha_1 - \kappa_1| = 1 - |a + k - \kappa_1|,$$

we have that $k$ is equal to either $\kappa_1$ or $\kappa_1 - 1$. If $k = \kappa_1$, then

$$\mathrm{Cap}_{\kappa_1}(p) = \inf_{x_1 > 0} \frac{x_1^k(ax_1 + b)}{x_1^k} = b \qquad \text{and} \qquad \epsilon = 1 - |a + k - \kappa_1| = 1 - a = b.$$

If $k = \kappa_1 - 1$, then

$$\mathrm{Cap}_{\kappa_1}(p) = \inf_{x_1 > 0} \frac{x_1^k(ax_1 + b)}{x_1^{k+1}} = a \qquad \text{and} \qquad \epsilon = 1 - |a + k - \kappa_1| = 1 - (1 - a) = a.$$

Therefore in both cases we have

$$\mathrm{Cap}_{\kappa_1}(p) = \epsilon = (1 - |\alpha_1 - \kappa_1|),$$

which proves Theorem 2.1 in the univariate case and demonstrates the tight dependence on $(1 - \|\boldsymbol{\kappa} - \nabla p(\mathbf{1})\|_1)$ in this case.

## 4 Proof of Application: Metric TSP

The following is the strongest probability lower bound, which we will use for the metric TSP application. In what follows, we let $A_S := \sum_{i \in S} A_i$ and $\kappa_S := \sum_{i \in S} \kappa_i$. Note that if

$$\left| \mathbb{E}_\mu [A_S] - \kappa_S \right| \leq 1 - \epsilon,$$

then Theorem 2.4 immediately follows from Theorem 4.1, Lemma 5.14, and a standard computation.

**Theorem 4.1** (Strongest form of the probability bound)**.** *Let $\mu$ be a strongly Rayleigh distribution on $[m]$, let $A_1, \dots, A_n$ be random variables counting the number of elements contained in some associated disjoint subsets of $[m]$, and fix $\boldsymbol{\kappa} \in \mathbb{Z}^n_{\geq 0}$. Suppose for all $S \subseteq [n]$ we have $|\mathbb{E}_\mu [A_S] - \kappa_S| < 1$. Define $\boldsymbol{\epsilon}, \boldsymbol{\delta} \in \mathbb{R}^n_{>0}$ via*

$$\delta_k := 1 + \min_{S \in \binom{[n]}{k}} \left(\mathbb{E}_\mu [A_S] - \kappa_S\right) \qquad \text{and} \qquad \epsilon_k := 1 - \max_{S \in \binom{[n]}{k}} \left(\mathbb{E}_\mu [A_S] - \kappa_S\right).$$

*Then we have*

$$\mathbb{P}_\mu [A_1 = \kappa_1, \dots, A_n = \kappa_n] \geq \prod_{i=1}^n \frac{\kappa_i^{\kappa_i} e^{-\kappa_i}}{\kappa_i!} \cdot \min_{0 \leq \ell \leq n} \prod_{k=1}^{\ell} \epsilon_k \prod_{k=1}^{n-\ell} \delta_k.$$

*Proof.* Let $q$ be the probability generating polynomial of $\mu$, and let $p$ be the polynomial obtained by setting the variables of $q$ associated to $A_i$ to $x_i$ for all $i \in [n]$, and setting all other variables to equal 1. Since $\mu$ is strongly Rayleigh, $p$ is real stable. Further, $p(\mathbf{1}) = 1$, $\nabla p(\mathbf{1}) = (\mathbb{E}_\mu [A_i])_{i=1}^n$, and $\mathbb{P}_\mu [A_1 = \kappa_1, \dots, A_n = \kappa_n]$ is the $\boldsymbol{x}^{\boldsymbol{\kappa}}$ coefficient of $p$. Thus Gurvits' capacity inequality (2) and Corollary 5.13 imply the desired result. □

We also give a slightly weaker bound which is a bit easier to use in practice. Note that Theorem 2.4 also follows from Corollary 4.2.

**Corollary 4.2.** *Let $\mu$ be a strongly Rayleigh distribution on $[m]$, let $A_1, \dots, A_n$ be random variables counting the number of elements contained in some associated disjoint subsets of $[m]$, and fix $\boldsymbol{\kappa} \in \mathbb{Z}^n_{\geq 0}$. Suppose for all $S \subseteq [n]$ we have $|\mathbb{E}_\mu [A_S] - \kappa_S| < 1$. Define $\boldsymbol{\epsilon} \in \mathbb{R}^n_{>0}$ via*

$$\epsilon_k := 1 - \max_{\substack{S \subseteq [n] \\ |S| \leq k}} \left|\mathbb{E}_\mu [A_S] - \kappa_S\right|$$

*Then we have*

$$\mathbb{P}_\mu [A_1 = \kappa_1, \dots, A_n = \kappa_n] \geq \prod_{i=1}^n \frac{\kappa_i^{\kappa_i} e^{-\kappa_i}}{\kappa_i!} \cdot \prod_{k=1}^{n} \epsilon_k.$$

*Proof.* Follows from Theorem 4.1; see the proof of Corollary 5.12 for more details. □

The remainder of this section is devoted to demonstrating how one can use the above results to improve the approximation factor for metric TSP.

## 4.1 A Simple Application

We recall some definitions from [KKO21]. There, we have a graph $G = (V, E)$ and a strongly Rayleigh (SR) distribution $\mu : 2^E \to \mathbb{R}_{\geq 0}$ supported on spanning trees of $G$. We let $x_e = \mathbb{P}_{T \sim \mu} [e \in T]$ and for a set of edges $F$ let $x(F) = \sum_{e \in F} x_e$. Furthermore we let $\delta(S) = \{e = \{u, v\} \mid |e \cap S| = 1\}$. The guarantee on $x$ is that $x \in P_{\text{Sub}}$,[2] where

$$P_{\text{Sub}} := \begin{cases} x(\delta(S)) \geq 2 & \forall\, S \subsetneq V \\ x(\delta(v)) = 2 & \forall v \in V \\ x_{\{u,v\}} \geq 0 & \forall\, u, v \in V \end{cases} \tag{3}$$

In the algorithm they analyze, one first samples a spanning tree $T$ from $\mu$ and then add the minimum cost matching on the odd vertices of $T$. Their work involves analyzing the expected cost of this matching over the randomness of the sampled tree $T$. Unsurprisingly, the *parity* of vertices is therefore very important, as it determines which vertices are involved in the matching.

In this section, to give some sense of the utility of Theorem 4.1, we show an application in a simplified setting. Namely, we show that for any two vertices $u, v$, except in the special case that $x_{\{u,v\}} \approx \frac{1}{2}$, we have $\mathbb{P}_{T \sim \mu} [|\delta(u) \cap T| = |\delta(v) \cap T| = 2] \geq \Omega(1)$, i.e. for any two vertices that do not share an edge of value $\frac{1}{2}$ there is a constant probability that they have even parity simultaneously. This is helpful because (under

[2] Technically, a spanning tree plus an edge is sampled, as otherwise one cannot exactly have $x \in P_{\text{Sub}}$, but we ignore that here.

some conditions on the point $x \in P_{\text{Sub}}$) the event that $u, v$ are even simultaneously indicates one can strictly decrease the cost of the matching proportional to the cost of the edge $e$.

We now prove $\mathbb{P}_{T\sim\mu}\left[|\delta(u)\cap T| = |\delta(v)\cap T| = 2\right] \geq \Omega(1)$ whenever $x_{\{u,v\}} \not\approx \frac{1}{2}$. To do this, we split into two cases: when $x_{\{u,v\}} \geq \frac{1}{2} + \epsilon$ and when $x_{\{u,v\}} \leq \frac{1}{2} - \epsilon$.

**Lemma 4.3.** *Let $u, v$ be two vertices such that $x_{\{u,v\}} \geq \frac{1}{2} + \epsilon$ for some $\epsilon > 0$. Then,*

$$\mathbb{P}_{T\sim\mu}\left[|\delta(u)\cap T| = |\delta(v)\cap T| = 2\right] \geq \frac{\epsilon}{2e^3}.$$

*Proof.* Let $e = (u, v)$. For $T \sim \mu$, let $A_1 = \mathbb{I}[e \in T]$, $A_2 = |(\delta(u) \setminus \{e\}) \cap T|$, $A_3 = |(\delta(v) \setminus \{e\}) \cap T|$. We are now interested in the event $A_i = \kappa_i, \forall i$ for the vector $\kappa = (1, 1, 1)$ as this implies $|\delta(u) \cap T| = |\delta(v) \cap T| = 2$. We have $\mathbb{E}[A_1] = x_e$, $\mathbb{E}[A_2] = \mathbb{E}[A_3] = 2 - x_e$. Therefore, to apply Theorem 4.1 we can set:

$$\delta_1 = x_e, \delta_2 = 1, \delta_3 = 2 - x_e, \qquad \epsilon_1 = x_e, \epsilon_2 = 2x_e - 1, \epsilon_3 = x_e$$

In this case the worst case is using all of the $\epsilon$ terms in the bound, giving $e^{-3}x_e^2(2x_e - 1) \geq \frac{\epsilon}{2e^3}$ as desired. □

**Lemma 4.4.** *Let $u, v$ be two vertices such that $x_{\{u,v\}} \leq \frac{1}{2} - \epsilon$ for some $\epsilon > 0$. Then,*

$$\mathbb{P}_{T\sim\mu}\left[|\delta(u)\cap T| = |\delta(v)\cap T| = 2\right] \geq \frac{2\epsilon}{e^4}.$$

*Proof.* As above, $e = (u, v)$, and for $T \sim \mu$, let $A_1 = \mathbb{I}[e \in T]$, $A_2 = |(\delta(u) \setminus \{e\}) \cap T|$, $A_3 = |(\delta(v) \setminus \{e\}) \cap T|$. We are now interested in the event $A_i = \kappa_i, \forall i$ for the vector $\kappa = (0, 2, 2)$ as this implies $|\delta(u) \cap T| = |\delta(v) \cap T| = 2$. We have $\mathbb{E}[A_1] = x_e$, $\mathbb{E}[A_2] = \mathbb{E}[A_3] = 2 - x_e$. Therefore, to apply Theorem 4.1 we can set:

$$\delta_1 = 1 - x_e, \delta_2 = 1 - 2x_e, \delta_3 = 1 - x_e, \qquad \epsilon_1 = 1 - x_e, \epsilon_2 = 1, \epsilon_3 = 1 + x_e$$

In this case the worst case is using all of the $\delta$ terms in the bound, giving $4e^{-4}(1 - x_e)^2(1 - 2x_e) \geq \frac{2\epsilon}{e^4}$. □

We leave further background about TSP and the proofs of our improved probabilistic statements to Appendix A, as understanding their importance requires some knowledge of the (highly technical) proof in [KKO21]. However, in the next section we summarize the new probabilistic bounds we get and their consequences.

## 4.2 Summary of Probabilistic Bounds and New Approximation Factor

The current bound on the performance of the max entropy algorithm is $\frac{3}{2} - 4.11 \cdot 10^{-36}$. This is primarily governed by a constant $p$ that is determined by the minimum probability over a number of events. In [KKO21], $p$ was equal to $2 \cdot 10^{-10}$, and these events are described by the following statements in [KKO21]:

1. Corollary 5.9, which gives a bound of $1.5 \cdot 10^{-9}$. We do not modify this bound, although we note that Lemma 5.7 can be slightly improved which would lead to a small improvement here.
2. Lemma 5.21, which gives a bound of $0.005\epsilon_{1/2}^2 = 2 \cdot 10^{-10}$. We improve this to $0.039\epsilon_{1/2}^2 = 1.56 \cdot 10^{-9}$.
3. Lemma 5.22, which gives a bound of $0.006\epsilon_{1/2}^2 = 2.4 \cdot 10^{-10}$. We improve this to $0.038\epsilon_{1/2}^2 = 1.52 \cdot 10^{-9}$.
4. Lemma 5.23, which gives a bound of $0.005 \cdot \epsilon_{1/2} = 2 \cdot 10^{-10}$. We observe here that arguments already in [KKO21] can be used to give $0.0498\epsilon_{1/2}^2 \geq 1.9 \cdot 10^{-9}$.
5. Lemma 5.24, which gives a bound of $0.02\epsilon_{1/2}^2 = 8 \cdot 10^{-10}$. We improve this to $0.0498\epsilon_{1/2}^2 \geq 1.99 \cdot 10^{-9}$.
6. Lemma 5.27, which gives a bound of 0.01. This lemma actually uses that the threshold $p$ is *small*, and therefore this bound decreases slightly upon raising $p$. However, as it is quite far from being the bottleneck in these bounds, we omit the proof that the probability remains above $1.5 \cdot 10^{-9}$.

We prove these bounds formally using our new strongly Rayleigh probability bounds in Appendix A.2 (which directly uses our new results) and Appendix A.3 (which mixes in more ad-hoc methods).

Therefore, we may increase $p$ to the minimum of all these probabilities, $1.5 \cdot 10^{-9}$. In Appendix A, we observe that using statements from [KKO21; KKO22], the following holds:

**Lemma 4.5.** *Let $p$ be a lower bound on the probabilities guaranteed by (1) - (6) for $\epsilon_{1/2} \leq 0.0002$, and suppose $p \leq 10^{-4}$. Then given $x \in P_{\text{Sub}}$, the max entropy algorithm returns a solution of expected cost at most $(\frac{3}{2} - 9.7p^2 \cdot 10^{-17}) \cdot c(x)$.*

As we improve the bounds on $p$ to $1.5 \cdot 10^{-9}$, an immediate corollary is the following:

**Corollary 4.6.** *The max entropy algorithm is a $\frac{3}{2} - 2.18 \cdot 10^{-34}$ approximation algorithm for metric TSP.*

Using [KKO23] this guarantee can be made deterministic, as we do not require any modifications to the algorithm.

In Appendix A, we also observe a lower bound on Lemma 5.21 for strongly Rayleigh distributions of $\Omega(\epsilon_{1/2}^2)$. The fact that $\epsilon_{1/2} \leq 0.0002$ is used in many places in [KKO21] and thus decreasing it may require more effort. Thus without modifying other parts of the argument, it may not be possible to improve the bound below $1.5 \cdot 10^{-31}$.

In the rest of the body of the paper we prove our main capacity bound.

# 5 Proofs of the Main Capacity Bounds

In this section we prove the strongest forms of the main capacity bounds, which give Theorem 2.1 and Corollary 2.2 as corollaries. For this section, we utilize the following notation.

**Definition 5.1.** For $n, d \in \mathbb{N}$ and $\boldsymbol{\alpha} \in \mathbb{R}^n_{\geq 0}$, we define the following:

1. $\text{NHMat}_n^d(\boldsymbol{\alpha})$ is the set of all $\mathbb{R}_{\geq 0}$-valued $d \times (n+1)$ matrices with row sums all equal to 1 and column sums equal to $\alpha_1, \ldots, \alpha_n, d - \|\boldsymbol{\alpha}\|_1$.
2. $\text{NHProd}_n^d(\boldsymbol{\alpha})$ is the set of all polynomials of the form
$$p(x) = \prod_{i=1}^{d} \left( a_{i,n+1} + \sum_{j=1}^{n} a_{i,j} x_j \right),$$
where $A \in \text{NHMat}_n^d(\boldsymbol{\alpha})$. In this case, we call $p$ the *polynomial associated to* $A$. Note that $p(\mathbf{1}) = 1$ and $\nabla p(\mathbf{1}) = \boldsymbol{\alpha}$ for all such polynomials.
3. $\text{NHStab}_n^d(\boldsymbol{\alpha})$ is the set of all real stable polynomials in $\mathbb{R}_{\geq 0}[x_1, \ldots, x_n]$ of degree at most $d$ for which $p(\mathbf{1}) = 1$ and $\nabla p(\mathbf{1}) = \boldsymbol{\alpha}$. (Recall that a polynomial is stable if it is never zero when all inputs are in the open complex upper half-plane.)

We also define the following for $n \in \mathbb{N}$, $\boldsymbol{\alpha} \in \mathbb{R}^n_{\geq 0}$, and $\boldsymbol{\kappa} \in \mathbb{Z}^n_{\geq 0}$:

$$L_n^{\text{NHProd}}(\boldsymbol{\alpha}; \boldsymbol{\kappa}) := \min_{d \in \mathbb{N}} \min_{p \in \text{NHProd}_n^d(\boldsymbol{\alpha})} \text{Cap}_{\boldsymbol{\kappa}}(p).$$

We now follow the steps of the proof given in Section 3.

## 5.1 Productization for Non-homogeneous Stable Polynomials

We first show how we can reduce the problem of bounding the capacity to products of affine linear forms. We recall the main productization result from [GL21], which gives the non-homogeneous productization result as an immediate corollary.

**Theorem 5.2** (Thm. 6.2, [GL21]). *Fix $n, d \in \mathbb{N}$, $\boldsymbol{u}, \boldsymbol{\alpha} \in \mathbb{R}^n_{\geq 0}$, and $p \in \mathbb{R}_{\geq 0}[x_1, \ldots, x_n]$ of homogeneous degree $d$, such that $p(\mathbf{1}) = 1$ and $\nabla p(\mathbf{1}) = \boldsymbol{\alpha}$. There exists an $\widetilde{\mathbb{R}}_{\geq 0}$-valued $d \times n$ matrix $A$ such that the rows sums of $A$ are all equal to $\mathbf{1}$, the column sums of $A$ are given by $\boldsymbol{\alpha}$, and $p(\boldsymbol{u}) = \prod_{i=1}^d (A\boldsymbol{u})_i$.*

**Corollary 5.3.** *Fix $n, d \in \mathbb{N}$, $\boldsymbol{u}, \boldsymbol{\alpha} \in \mathbb{R}^n_{\geq 0}$, and $p \in \mathrm{NHStab}^d_n(\boldsymbol{\alpha})$. There exists $f \in \mathrm{NHProd}^d_n(\boldsymbol{\alpha})$ such that $p(\boldsymbol{u}) = f(\boldsymbol{u})$.*

*Proof.* Let $q(\boldsymbol{x}) = x_{n+1}^d \cdot p\left(\frac{x_1}{x_{n+1}}, \ldots, \frac{x_n}{x_{n+1}}\right)$ be the homogenization of $p$, and define $\boldsymbol{\beta} := \nabla q(\mathbf{1})$. So $q \in \mathbb{R}_{\geq 0}[x_1, \ldots, x_{n+1}]$ of homogeneous degree $d$ such that $q(\mathbf{1}) = 1$ and $\nabla q(\mathbf{1}) = \boldsymbol{\beta} = (\alpha_1, \ldots, \alpha_n, d - \|\boldsymbol{\alpha}\|_1)$. Define $u_{n+1} := 1$, apply Theorem 5.2 to $q$ and $\boldsymbol{u}$, and dehomogenize to obtain the desired result. □

We now use this result to reduce the problem of bounding the capacity to products of affine linear forms.

**Corollary 5.4.** *For $p \in \mathrm{NHStab}^d_n(\boldsymbol{\alpha})$, we have*

$$\mathrm{Cap}_\kappa(p) \geq L_n^{\mathrm{NHProd}}(\boldsymbol{\alpha}; \boldsymbol{\kappa}).$$

*Proof.* For any $\boldsymbol{x} \in \mathbb{R}^n_{\geq 0}$, let $f \in \mathrm{NHProd}^d_n(\boldsymbol{\alpha})$ be such that $p(\boldsymbol{x}) = f(\boldsymbol{x})$ according to Corollary 5.3. With this, we have

$$\mathrm{Cap}_{\boldsymbol{\kappa}}(p) = \inf_{\boldsymbol{x}>0} \frac{p(\boldsymbol{x})}{\boldsymbol{x}^{\boldsymbol{\kappa}}} \geq \inf_{\boldsymbol{x}>0} \min_{d \in \mathbb{N}} \min_{f \in \mathrm{NHProd}^d_n(\boldsymbol{\alpha})} \frac{f(\boldsymbol{x})}{\boldsymbol{x}^{\boldsymbol{\kappa}}} = L_n^{\mathrm{NHProd}}(\boldsymbol{\alpha}; \boldsymbol{\kappa}).$$

□

## 5.2 The Extreme Points of $\mathrm{NHMat}^d_n(\alpha)$

The next result implies we can reduce to the extreme points of $\mathrm{NHMat}^d_n(\boldsymbol{\alpha})$ to lower bound $L_n^{\mathrm{NHProd}}(\boldsymbol{\alpha}; \boldsymbol{\kappa})$.

**Lemma 5.5** (See Thm. 3.1 of [Bar08]). *Given $\boldsymbol{\kappa} \in \mathbb{Z}^n_{\geq 0}$ and $\boldsymbol{\alpha} \in \mathbb{R}^n_{\geq 0}$, let $\phi : \mathrm{NHMat}^d_n(\boldsymbol{\alpha}) \to \mathbb{R}_{\geq 0}$ be the function which maps $M$ to $\mathrm{Cap}_{\boldsymbol{\kappa}}(p)$ where $p$ is the polynomial associated to $M$. Then $\phi$ is log-concave on $\mathrm{NHMat}^d_n(\boldsymbol{\alpha})$.*

*Proof.* Fix any $A, B \in \mathrm{NHMat}^d_n(\boldsymbol{\alpha})$, and let $p, q, f$ be the polynomials associated to $A, B, \frac{A+B}{2}$ respectively. By the AM-GM inequality, we compute

$$\begin{aligned}
\mathrm{Cap}_{\boldsymbol{\kappa}}(f) &= \inf_{\boldsymbol{x}>0} \frac{\prod_{i=1}^d \left(\frac{a_{i,n+1}+b_{i,n+1}}{2} + \sum_{j=1}^n \frac{a_{i,j}+b_{i,j}}{2} \cdot x_j\right)}{\boldsymbol{x}^{\boldsymbol{\kappa}}} \\
&= \inf_{\boldsymbol{x}>0} \frac{\prod_{i=1}^d \left[\frac{\left(a_{i,n+1}+\sum_{j=1}^n a_{i,j}x_j\right)+\left(b_{i,n+1}+\sum_{j=1}^n b_{i,j}x_j\right)}{2}\right]}{\boldsymbol{x}^{\boldsymbol{\kappa}}} \\
&\geq \inf_{\boldsymbol{x}>0} \left[\frac{\prod_{i=1}^d \left(a_{i,n+1} + \sum_{j=1}^n a_{i,j}x_j\right)}{\boldsymbol{x}^{\boldsymbol{\kappa}}} \cdot \frac{\prod_{i=1}^d \left(b_{i,n+1} + \sum_{j=1}^n b_{i,j}x_j\right)}{\boldsymbol{x}^{\boldsymbol{\kappa}}}\right]^{\frac{1}{2}} \\
&\geq \left[\mathrm{Cap}_{\boldsymbol{\kappa}}(p) \cdot \mathrm{Cap}_{\boldsymbol{\kappa}}(q)\right]^{\frac{1}{2}}
\end{aligned}$$

□

We next describe the extreme points of $\mathrm{NHMat}^d_n(\boldsymbol{\alpha})$ via bipartitioned forests.

**Lemma 5.6.** *Any extreme point of $\mathrm{NHMat}^d_n(\boldsymbol{\alpha})$ has support given by a bipartite forest on $d$ left vertices and $n+1$ right vertices.*

*Proof.* Let $M$ be an extreme point of $\mathrm{NHMat}^d_n(\boldsymbol{\alpha})$, and suppose its bipartite support graph $G$ does not give a forest. Then $G$ must contain an even simple cycle. Group the edges of this cycle into two groups such that the odd edges make up one group, and the even edges make up the other (with any starting point). Add $\epsilon > 0$ to all matrix entries corresponding to even edges and subtract $\epsilon$ to all matrix entries corresponding to odd edges, to construct $M_+ \in \mathrm{NHMat}^d_n(\boldsymbol{\alpha})$. Do the same thing, but reverse the signs, to construct $M_- \in \mathrm{NHMat}^d_n(\boldsymbol{\alpha})$. Thus $M = \frac{M_+ + M_-}{2}$, contradicting the fact that $M$ is an extreme point. □

In what follows, we will also need the following basic graph theoretic result.

**Lemma 5.7.** *Let $G$ be a bipartite forest on $m$ left vertices and $n$ right vertices such that $G$ has no vertices of degree 0. Then $G$ has at least $m - n + 1$ left leaves.*

*Proof.* We prove this by induction, where the base case is any path graph. For this case, we have $m \in \{n-1, n, n+1\}$. If $m = n-1$ then $G$ has 0 left leaves, if $m = n$ then $G$ has 1 left leaf, and if $m = n+1$ then $G$ has two left leaves. Thus the desired result holds in this case.

For the inductive step, $G$ is not a path graph. Let $v$ be any leaf of $G$, and construct a new graph $G'$ as follows. Let $v_0 := v$ and remove $v_0$ from $G_0 := G_0$ to create the graph $G_1$. Let $v_1$ be the one neighbor of $v_0$ in $G$. If $v_1$ is a leaf or vertex of degree 0 in $G_1$, then remove $v_1$ from $G_1$ to create the graph $G_2$. If $v_1$ was a vertex of degree 0, then stop and define $G' := G_1$. Otherwise let $v_2$ be the one neighbor of $v_1$, and continue this process inductively until $v_k$ is not a leaf or a vertex of degree 0 in $G_k$. Once the process stops, define $G' := G_k$. Note that $G'$ is a bipartite forest which has no vertices of degree 0, and $G'$ is non-empty since $G$ is not a path graph. We now have two cases: $v$ is a left leaf of $G$, or $v$ is a right leaf of $G$.

First suppose $v$ is a left leaf of $G$. Then for some $i$, $G'$ has $m-i$ left vertices and at most $n-i+1$ right vertices. Thus by induction, the number of left leaves of $G$ is at least

$$1 + (m-i) - (n-i+1) + 1 = m - n + 1$$

since $v$ is a left leaf. Thus the result holds in this case.

Next suppose $v$ is a right leaf of $G$. Then for some $i$, $G'$ has $m-i$ left vertices and at most $n-i$ right vertices. Thus by induction, the number of left leaves of $G$ is at least

$$(m-i) - (n-i) + 1 = m - n + 1.$$

Thus the result holds in this case as well. □

## 5.3 Capacity Bounds via Induction

Here we complete the proof of Theorem 2.1. We first prove a simple lemma, which bears some resemblance to the probabilistic union bound.

**Lemma 5.8.** *Given $\boldsymbol{c} \in \mathbb{R}^d_{\geq 0}$ such that $c_i < 1$ for all $i \in [d]$, we have $1 - \sum_{i=1}^d c_i \leq \prod_{i=1}^d (1 - c_i)$.*

*Proof.* We prove the desired result by induction, where the case of $d = 1$ is trivial. Now note that

$$\frac{1 - \sum_{i=1}^d c_i}{1 - c_d} = \frac{(1-c_d)\left(1 - \sum_{i=1}^{d-1} c_i\right) + c_d\left(1 - \sum_{i=1}^{d-1} c_i\right) - c_d}{1 - c_d} = 1 - \sum_{i=1}^{d-1} c_i - \frac{c_d}{1-c_d}\sum_{i=1}^{d-1} c_i \leq 1 - \sum_{i=1}^{d-1} c_i.$$

Thus by induction we have

$$1 - \sum_{i=1}^{d} c_i \leq (1-c_d)\left(1 - \sum_{i=1}^{d-1} c_i\right) \leq (1-c_d)\prod_{i=1}^{d-1}(1-c_i),$$

and this completes the proof. □

The next lemma handles the case from Step 4 in Section 3 of setting some variable equal to 0. To see how these next two lemmas actually are actually used, see Theorem 5.11 below.

**Lemma 5.9.** *For $n \geq 1$, let $p \in \mathrm{NHProd}^d_n(\boldsymbol{\alpha})$ be the polynomial associated to a $d \times (n+1)$ matrix $M$, and suppose $\boldsymbol{\kappa} \in \mathbb{Z}^n_{\geq 0}$ such that $\kappa_n = 0$ and $\alpha_n - \kappa_n \leq 1 - \epsilon$ for some $\epsilon > 0$. Then there exists $q \in \mathrm{NHProd}^d_{n-1}(\boldsymbol{\beta})$ such that*

$$\mathrm{Cap}_{\boldsymbol{\kappa}}(p) \geq \epsilon \cdot \mathrm{Cap}_{\boldsymbol{\gamma}}(q),$$

*where $\boldsymbol{\gamma} = (\kappa_1, \ldots, \kappa_{n-1}) \in \mathbb{Z}^{n-1}_{\geq 0}$ and $\boldsymbol{\beta} \in \mathbb{R}^{n-1}_{\geq 0}$ is such that for all $S \subseteq [n-1]$ we have*

$$\sum_{j \in S}(\alpha_j - \kappa_j) \leq \sum_{j \in S}(\beta_j - \gamma_j) \leq (\alpha_n - \kappa_n) + \sum_{j \in S}(\alpha_j - \kappa_j).$$

*Proof.* Define $q_0(\boldsymbol{x}) = q_0(x_1, \ldots, x_{n-1})$ via

$$q_0(\boldsymbol{x}) = p(x_1, \ldots, x_{n-1}, 0),$$

so that

$$\mathrm{Cap}_{\boldsymbol{\kappa}}(p) = \mathrm{Cap}_{\boldsymbol{\gamma}}(q_0).$$

Note that $q_0$ is the polynomial associated to a $d \times n$ matrix $A$, where the row sums of $A$ are given by $1 - m_{in}$ for all $i$. Note that $m_{in} \leq \alpha_n \leq 1 - \epsilon < 1$ implies $1 - m_{in} > 0$ for all $i$. We now construct a new matrix $B \in \mathrm{NHMat}^d_{n-1}(\boldsymbol{\beta})$ by dividing row $i$ of $A$ by $1 - m_{in}$ for all $i$. Defining $q(\boldsymbol{x})$ to be the polynomial associated to the matrix $B$, we have

$$\mathrm{Cap}_{\boldsymbol{\kappa}}(p) = \mathrm{Cap}_{\boldsymbol{\gamma}}(q_0) = \prod_{i=1}^{d}(1 - m_{in}) \cdot \mathrm{Cap}_{\boldsymbol{\gamma}}(q) \geq \epsilon \cdot \mathrm{Cap}_{\boldsymbol{\gamma}}(q)$$

by Lemma 5.8. Finally, for all $S \subseteq [n-1]$ we compute

$$\begin{aligned}
\sum_{j \in S}(\beta_j - \gamma_j) &= \sum_{j \in S}\left(\alpha_j - \kappa_j + \sum_{i=1}^{d}\left(\frac{m_{ij}}{1 - m_{in}} - m_{ij}\right)\right) \\
&= \sum_{i=1}^{d}\frac{m_{in}}{1 - m_{in}}\sum_{j \in S} m_{ij} + \sum_{j \in S}(\alpha_j - \kappa_j) \\
&\leq \sum_{i=1}^{d}\frac{m_{in}}{1 - m_{in}}(1 - m_{in}) + \sum_{j \in S}(\alpha_j - \kappa_j) \\
&= (\alpha_n - \kappa_n) + \sum_{j \in S}(\alpha_j - \kappa_j).
\end{aligned}$$

The other inequality then follows from the fact that $\sum_{i=1}^{d}\frac{m_{in}}{1-m_{in}}\sum_{j \in S} m_{ij} \geq 0$. □

The next lemma handles the case from Step 4 in Section 3 of taking the partial derivative with respect to some variable.

**Lemma 5.10.** *For $n \geq 1$, let $p \in \mathrm{NHProd}^d_n(\boldsymbol{\alpha})$ be the polynomial associated to a $d \times (n+1)$ matrix $M$ such that column $n$ has exactly one non-zero entry, and suppose $\boldsymbol{\kappa} \in \mathbb{Z}^n_{\geq 0}$ is such that $\kappa_n = 1$ and $\kappa_n - \alpha_n \leq 1 - \epsilon$ for some $\epsilon > 0$. Then there exists $q \in \mathrm{NHProd}^{d-1}_{n-1}(\boldsymbol{\beta})$ such that*

$$\mathrm{Cap}_{\boldsymbol{\kappa}}(p) \geq \epsilon \cdot \mathrm{Cap}_{\boldsymbol{\gamma}}(q),$$

*where $\boldsymbol{\gamma} = (\kappa_1, \ldots, \kappa_{n-1}) \in \mathbb{Z}^{n-1}_{\geq 0}$ and $\boldsymbol{\beta} \in \mathbb{R}^{n-1}_{\geq 0}$ is such that for all $S \subseteq [n-1]$ we have*

$$\sum_{j \in S}(\kappa_j - \alpha_j) \leq \sum_{j \in S}(\gamma_j - \beta_j) \leq (\kappa_n - \alpha_n) + \sum_{j \in S}(\kappa_j - \alpha_j).$$

*Proof.* Let $i_0$ be the row index of the one non-zero entry in column $n$. Note that $\epsilon \leq \alpha_n = m_{i_0,n} \leq 1$. Define $q(\boldsymbol{x}) = q(x_1, \ldots, x_{n-1})$ via

$$p(\boldsymbol{x}) = \left(m_{i_0,n+1} + \sum_{j=1}^{n} m_{i_0,j}x_j\right) \cdot q(\boldsymbol{x}).$$

That is, $q(\boldsymbol{x})$ is the polynomial associated to the matrix obtained by removing row $i_0$ of $M$. We then have

$$\mathrm{Cap}_{\boldsymbol{\kappa}}(p) \geq \mathrm{Cap}_{\boldsymbol{e}_n}\left(m_{i_0,n+1} + \sum_{j=1}^{n} m_{i_0,j}x_j\right) \cdot \mathrm{Cap}_{\boldsymbol{\gamma}}(q) = m_{i_0,n} \cdot \mathrm{Cap}_{\boldsymbol{\gamma}}(q) \geq \epsilon \cdot \mathrm{Cap}_{\boldsymbol{\gamma}}(q).$$

Note that $q$ is the polynomial associated to a $(d-1)\times n$ matrix $A$ with row sums all equal to 1. Let $\boldsymbol{\beta}$ be the vector of the first $n-1$ column sums of this matrix, so that $q \in \mathrm{NHProd}_{n-1}^{d-1}(\boldsymbol{\beta})$. For all $S \subseteq [n-1]$ we have

$$\begin{aligned}\sum_{j\in S}(\gamma_j - \beta_j) &= \sum_{j\in S}(\kappa_j - \alpha_j + m_{i_0,j}) \\ &= \sum_{j\in S} m_{i_0,j} + \sum_{j\in S}(\kappa_j - \alpha_j) \\ &\le (1 - m_{i_0,n}) + \sum_{j\in S}(\kappa_j - \alpha_j) \\ &= (\kappa_n - \alpha_n) + \sum_{j\in S}(\kappa_j - \alpha_j).\end{aligned}$$

The other inequality then follows from the fact that $\sum_{j\in S} m_{i_0,j} \ge 0$. □

We now combine the lemmas above to prove our strongest capacity lower bound for polynomials in $\mathrm{NHProd}_n^d(\boldsymbol{\alpha})$.

**Theorem 5.11.** *Fix any $\boldsymbol{\kappa} \in \mathbb{Z}_{\ge 0}^n$, $\boldsymbol{\alpha} \in \mathbb{R}_{\ge 0}^n$, and $\boldsymbol{\epsilon}, \boldsymbol{\delta} \in \mathbb{R}_{>0}^n$ such that*

$$\max_{S\in\binom{[n]}{k}} \sum_{j\in S}(\alpha_j - \kappa_j) \le 1 - \epsilon_k \qquad \text{and} \qquad \max_{S\in\binom{[n]}{k}} \sum_{j\in S}(\kappa_j - \alpha_j) \le 1 - \delta_k$$

*for all $k \in [n]$. Then*

$$\mathrm{Cap}_{\boldsymbol{\kappa}}(p) \ge \min_{0\le \ell \le n} \prod_{k=1}^{\ell} \epsilon_k \prod_{k=1}^{n-\ell} \delta_k$$

*for every $p \in \mathrm{NHProd}_n^d(\boldsymbol{\alpha})$.*

*Proof.* We prove the desired result by induction on $(n, d)$ with lexicographical order. The base case is the case where $n = 0$, which corresponds to $p \equiv 1$ for any $d$. In this case, $\boldsymbol{\kappa}$ is the empty vector and $\mathrm{Cap}_{\boldsymbol{\kappa}}(p) = 1$, which implies the desired result. (See Section 3.2 for the $n = 1$ case written out in detail.)

For $n \ge 1$, we first handle the case that $\kappa_j = 0$ for some $j \in [n]$, and by permuting the variables we may assume that $\kappa_n = 0$. By Lemma 5.9, there exists $q \in \mathrm{NHProd}_{n-1}^d(\boldsymbol{\beta})$ such that

$$\mathrm{Cap}_{\boldsymbol{\kappa}}(p) \ge \epsilon_1 \cdot \mathrm{Cap}_{\boldsymbol{\gamma}}(q),$$

where $\boldsymbol{\gamma} = (\kappa_1, \ldots, \kappa_{n-1}) \in \mathbb{Z}_{\ge 0}^{n-1}$ and $\boldsymbol{\beta} \in \mathbb{R}_{\ge 0}^{n-1}$ is such that for all $S \subseteq [n-1]$ we have

$$\sum_{j\in S}(\alpha_j - \kappa_j) \le \sum_{j\in S}(\beta_j - \gamma_j) \le (\alpha_n - \kappa_n) + \sum_{j\in S}(\alpha_j - \kappa_j).$$

Thus for all $k \in [n-1]$, we have

$$\max_{S\in\binom{[n-1]}{k}} \sum_{j\in S}(\beta_j - \gamma_j) \le (\alpha_n - \kappa_n) + \max_{S\in\binom{[n-1]}{k}} \sum_{j\in S}(\alpha_j - \kappa_j) \le 1 - \epsilon_{k+1}$$

and

$$\max_{S\in\binom{[n-1]}{k}} \sum_{j\in S}(\gamma_j - \beta_j) \le \max_{S\in\binom{[n-1]}{k}} \sum_{j\in S}(\kappa_j - \alpha_j) \le 1 - \delta_k.$$

Thus by induction we have that

$$\mathrm{Cap}_{\boldsymbol{\kappa}}(p) \ge \epsilon_1 \cdot \mathrm{Cap}_{\boldsymbol{\gamma}}(q) \ge \min_{0\le \ell \le n-1} \epsilon_1 \prod_{k=1}^{\ell} \epsilon_{k+1} \prod_{k=1}^{n-1-\ell} \delta_k = \min_{1\le \ell \le n} \prod_{k=1}^{\ell} \epsilon_k \prod_{k=1}^{n-\ell} \delta_k,$$

and this implies the desired result.

Otherwise, $\kappa_j \geq 1$ for all $j \in [n]$. Now fix any extreme point $M \in \mathrm{NHMat}_n^d(\boldsymbol{\alpha})$, and let $p$ be the polynomial associated to $M$. By Lemma 5.6, $M$ is the (weighted) bipartite adjacency matrix of a forest $G$. We call the vertices of $G$ corresponding to the rows of $M$ the *left vertices* of $G$, and we call the vertices of $G$ corresponding to the columns of $M$ the *right vertices* of $G$. Since $|\alpha_j - \kappa_j| < 1$ for all $j \in [n]$, we further have that none of the first $n$ columns of $M$ are equal to the zero vector, and thus $G$ has at most one (right) vertex of degree 0 (possibly the vertex corresponding to column $n+1$ of $M$).

We first consider the case where $d \geq n+1$. Then $G$ has at least as many left vertices as right vertices. After possibly removing the right vertex in $G$ of degree 0, Lemma 5.7 implies $M$ has at least $d-n$ rows with exactly one non-zero entry. Letting $\boldsymbol{x}^{\boldsymbol{v}}$ be the polynomial associated to these rows of $M$, we have

$$\mathrm{Cap}_{\boldsymbol{\kappa}}(p) = \inf_{\boldsymbol{x}>0} \frac{\boldsymbol{x}^{\boldsymbol{v}} f(\boldsymbol{x})}{\boldsymbol{x}^{\boldsymbol{\kappa}}} = \mathrm{Cap}_{\boldsymbol{\kappa}-\boldsymbol{v}}(f).$$

With this, we now redefine $\boldsymbol{\kappa}$ to be $\boldsymbol{\kappa} - \boldsymbol{v}$, $\boldsymbol{\alpha}$ to be $\boldsymbol{\alpha} - \boldsymbol{v}$, $p$ to be $f$, and $M$ to be the matrix with the (at least) $d-n$ rows removed. Since $(\alpha_j - v_j) - (\kappa_j - v_j) = \alpha_j - \kappa_j$ for all $j \in [n]$, the result in this case then follows by induction.

We now consider the case where $d \leq n$. Then $G$ has at least one more right vertex than left vertices. After possibly removing the right vertex in $G$ of degree 0 (corresponding to column $n+1$ of $M$), Lemma 5.7 implies at least one of the first $n$ columns of $M$ has exactly one non-zero entry, and by permuting the variables we may assume that it is column $n$. Thus $\alpha_n \leq 1$, and $|\alpha_n - \kappa_n| < 1$ then implies $\kappa_n = 1$. By Lemma 5.10, there exists $q \in \mathrm{NHProd}_{n-1}^{d-1}(\boldsymbol{\beta})$ such that

$$\mathrm{Cap}_{\boldsymbol{\kappa}}(p) \geq \delta_1 \cdot \mathrm{Cap}_{\boldsymbol{\gamma}}(q),$$

where $\boldsymbol{\gamma} = (\kappa_1, \ldots, \kappa_{n-1}) \in \mathbb{Z}_{\geq 0}^{n-1}$ and $\boldsymbol{\beta} \in \mathbb{R}_{\geq 0}^{n-1}$ is such that for all $S \subseteq [n-1]$ we have

$$\sum_{j \in S} (\kappa_j - \alpha_j) \leq \sum_{j \in S} (\gamma_j - \beta_j) \leq (\kappa_n - \alpha_n) + \sum_{j \in S} (\kappa_j - \alpha_j).$$

Thus for all $k \in [n-1]$, we have

$$\max_{S \in \binom{[n-1]}{k}} \sum_{j \in S} (\beta_j - \gamma_j) \leq \max_{S \in \binom{[n-1]}{k}} \sum_{j \in S} (\alpha_j - \kappa_j) \leq 1 - \epsilon_k$$

and

$$\max_{S \in \binom{[n-1]}{k}} \sum_{j \in S} (\gamma_j - \beta_j) \leq (\kappa_n - \alpha_n) + \max_{S \in \binom{[n-1]}{k}} \sum_{j \in S} (\kappa_j - \alpha_j) \leq 1 - \delta_{k+1}.$$

Thus by induction we have that

$$\mathrm{Cap}_{\boldsymbol{\kappa}}(p) \geq \delta_1 \cdot \mathrm{Cap}_{\boldsymbol{\gamma}}(q) \geq \min_{0 \leq \ell \leq n-1} \delta_1 \prod_{k=1}^{\ell} \epsilon_k \prod_{k=1}^{n-1-\ell} \delta_{k+1} = \min_{0 \leq \ell \leq n-1} \prod_{k=1}^{\ell} \epsilon_k \prod_{k=1}^{n-\ell} \delta_k,$$

and this implies the desired result.

We have proven the result for polynomials associated to the extreme points of $\mathrm{NHMat}_n^n(\boldsymbol{\alpha})$. Applying Lemma 5.5 then completes the proof. □

The next result gives the bound which we will use in Section 4 to prove Theorem 2.4, the simply exponential improvement to the probability bound used for the metric TSP application.

**Corollary 5.12.** *Fix any $\boldsymbol{\kappa} \in \mathbb{Z}_{\geq 0}^n$ and $\boldsymbol{\alpha} \in \mathbb{R}_{\geq 0}^n$ such that $\boldsymbol{\epsilon} \in \mathbb{R}_{>0}^n$ can be defined via*

$$\epsilon_k := 1 - \max_{\substack{S \subseteq [n] \\ |S| \leq k}} \left| \sum_{j \in S} (\kappa_j - \alpha_j) \right|$$

*for all $k \in [n]$. Then*

$$\mathrm{Cap}_{\boldsymbol{\kappa}}(p) \geq \prod_{k=1}^{n} \epsilon_k$$

*for every $p \in \mathrm{NHProd}_n^d(\boldsymbol{\alpha})$.*

*Proof.* Note that

$$\max_{\substack{S\subseteq[n]\\|S|\leq k}} \sum_{j\in S}(\alpha_j - \kappa_j) \leq 1-\epsilon_k \qquad \text{and} \qquad \max_{\substack{S\subseteq[n]\\|S|\leq k}} \sum_{j\in S}(\kappa_j - \alpha_j) \leq 1-\epsilon_k$$

for all $k \in [n]$, and thus

$$\mathrm{Cap}_{\boldsymbol{\kappa}}(p) \geq \min_{0\leq \ell \leq n} \prod_{k=1}^{\ell} \epsilon_k \prod_{k=1}^{n-\ell} \epsilon_k$$

by Theorem 5.11. Since $\epsilon_1 \geq \epsilon_2 \geq \cdots \geq \epsilon_n$, this implies the result. □

By Corollary 5.4, the above results hold for all $p \in \mathrm{NHStab}_n^d(\boldsymbol{\alpha})$, and we state this formally now.

**Corollary 5.13.** *Theorem 5.11 and Corollary 5.12 hold for all $p \in \mathrm{NHStab}_n^d(\boldsymbol{\alpha})$.*

## 5.4 Proving Theorem 2.1 and Corollary 2.2

We now complete the proof of Theorem 2.1, and thus also of Corollary 2.2. Fix $p \in \mathrm{NHStab}_n^d(\boldsymbol{\alpha})$. Thus

$$1 - \max_{\substack{S\subseteq[n]\\|S|\leq k}} \left|\sum_{j\in S}(\kappa_j - \alpha_j)\right| \geq 1 - \|\boldsymbol{\kappa} - \boldsymbol{\alpha}\|_1$$

and Corollary 5.13 (via Corollary 5.12) imply

$$\mathrm{Cap}_{\boldsymbol{\kappa}}(p) \geq (1 - \|\boldsymbol{\kappa} - \boldsymbol{\alpha}\|_1)^n,$$

which completes the proof. The tightness claim follows from Lemma 5.14.

## 5.5 Examples and Tightness

The $\delta$ parameters in Theorem 5.11 are tight, and this is shown in Example 5.15. However, the $\epsilon$ parameters are not, as shown in Example 5.16. Example 5.15 also demonstrates tightness of the dependence on the error parameter for some of our results, and we state this formally now.

**Lemma 5.14.** *The dependence on $(1-\|\boldsymbol{\kappa}-\nabla p(\mathbf{1})\|_1)$ in Theorem 2.1 and Corollary 2.2 and the dependence on $\epsilon$ in Theorem 2.4 are all tight for any fixed $\boldsymbol{\kappa} > \mathbf{0}$.*

*Proof.* Define $\alpha_1 := \kappa_1 - (1-\epsilon) > 0$ and $\alpha_i := \kappa_i \geq 1$ for all $i \geq 2$. Let $p$ be the polynomial described by Example 5.15, given explicitly by

$$p(\boldsymbol{x}) = \left(\prod_{i=1}^{n} x_i^{\kappa_i - 1}\right) \cdot \left(\prod_{i=1}^{n-1} (\epsilon x_i + (1-\epsilon)x_{i+1})\right) \cdot (\epsilon x_n + (1-\epsilon)).$$

Then $\boldsymbol{\kappa}$ is a vertex of of the Newton polytope of $p$. Thus

$$p_{\boldsymbol{\kappa}} = \mathrm{Cap}_{\boldsymbol{\kappa}}(p) = \prod_{k=1}^{n} \left(1 - \sum_{j=1}^{k}(\kappa_j - \alpha_j)\right) = \epsilon^n = (1 - \|\boldsymbol{\kappa} - \nabla p(\mathbf{1})\|_1)^n.$$

To see that $p$ can be the probability generating polynomial for some random variables associated to a strongly Rayleigh distribution (as in Theorem 2.4), note that the polarization of $p$ is real stable and gives the probability generating polynomial for such a strongly Rayleigh distribution. □

**Example 5.15.** Fix any $\boldsymbol{\kappa} \in \mathbb{Z}^n_{\geq 0}$ and $\boldsymbol{\alpha} \in \mathbb{R}^n_{\geq 0}$ such that $\kappa_j - \alpha_j \geq 0$ for all $j \in [n]$ and $\|\boldsymbol{\kappa} - \boldsymbol{\alpha}\|_1 < 1$. Thus $\kappa_j \geq 1$ for all $j \in [n]$. For $d = \|\boldsymbol{\kappa}\|_1$, consider the matrix $M = \left[\begin{smallmatrix} A \\ B \end{smallmatrix}\right]$, where $A$ is the $(\|\boldsymbol{\kappa}\| - n) \times (n+1)$ matrix given by

$$A = \begin{bmatrix} \mathbf{1}_{\kappa_1 - 1} \boldsymbol{e}_1^\top \\ \mathbf{1}_{\kappa_2 - 1} \boldsymbol{e}_2^\top \\ \vdots \\ \mathbf{1}_{\kappa_n - 1} \boldsymbol{e}_n^\top \end{bmatrix},$$

and $B$ is the $n \times (n+1)$ matrix for which

$$b_{kk} = 1 - \sum_{j=1}^{k} (\kappa_j - \alpha_j), \qquad b_{k,k+1} = \sum_{j=1}^{k} (\kappa_j - \alpha_j),$$

for all $k \in [n]$ and $b_{ij} = 0$ otherwise. Note that every row sum of $A$ is equal to 1, and the row sums of $B$ are given by

$$\sum_{j=1}^{n} b_{kj} = 1 - \sum_{j=1}^{k} (\kappa_j - \alpha_j) + \sum_{j=1}^{k} (\kappa_j - \alpha_j) = 1$$

for all $k \in [n]$. The column sums of $M$ are then given by

$$\sum_{i=1}^{d} m_{ik} = (\kappa_k - 1) + 1 - \sum_{j=1}^{k} (\kappa_j - \alpha_j) + \sum_{j=1}^{k-1} (\kappa_j - \alpha_j) = \kappa_k - (\kappa_k - \alpha_k) = \alpha_k$$

for all $k \in [n]$. Thus $M \in \mathrm{NHMat}_n^d(\boldsymbol{\alpha})$. Let $p$ be the polynomial associated to $M$, and let $q$ be the polynomial associated to $B$. We then have $\mathrm{Cap}_{\boldsymbol{\kappa}}(p) = \mathrm{Cap}_{\mathbf{1}}(q)$. Note that $\mathbf{1}$ is a vertex of the Newton polytope of $q$, and thus

$$\mathrm{Cap}_{\boldsymbol{\kappa}}(p) = \mathrm{Cap}_{\mathbf{1}}(q) = \prod_{k=1}^{n} b_{kk} = \prod_{k=1}^{n} \left(1 - \sum_{j=1}^{k} (\kappa_j - \alpha_j)\right).$$

By possibly permuting the variables to put $\kappa_j - \alpha_j$ in non-increasing order, this is precisely the lower bound guaranteed by Theorem 5.11.

**Example 5.16.** Consider the case that $\boldsymbol{\kappa} = \mathbf{0}$ and $\|\boldsymbol{\alpha}\|_1 < 1$. Given any $d$, let $M$ be any extreme point of $\mathrm{NHMat}_n^d(\boldsymbol{\alpha})$. Then the column sum of column $n+1$ of $M$ is equal to $d - \sum_{j=1}^{n} \alpha_j > d - 1$. Since every row sum equals 1, all entries of column $n+1$ of $M$ are strictly positive. Since $M$ is an extreme point, this further implies that each column of $M$ has at most 1 positive entry except column $n+1$. Letting $p$ be the polynomial associated to $M$, there exists a partition $S_1 \sqcup \cdots \sqcup S_k = [n]$ such that

$$p(\boldsymbol{x}) = \prod_{i=1}^{k} \left( \left( \sum_{j \in S_i} \alpha_j x_j \right) + \left( 1 - \sum_{j \in S_i} \alpha_j \right) \right).$$

Thus by Lemma 5.8,

$$\mathrm{Cap}_{\mathbf{0}}(p) = p_{\mathbf{0}} = \prod_{i=1}^{k} \left( 1 - \sum_{j \in S_i} \alpha_j \right) \geq 1 - \sum_{j=1}^{n} \alpha_j.$$

By Lemma 5.5, this gives a lower bound on the capacity of every $p \in \mathrm{NHProd}_n^d(\boldsymbol{\alpha})$. However, this lower bound is strictly better than the one guaranteed by Theorem 5.11.

As a note, this can be partially remedied by removing all $\kappa_j = 0$ columns at the same time (i.e., adjusting Lemma 5.9 to remove many columns at once). However, it is currently unclear how to inductively do this correctly.

**Acknowledgements.** The third author would like to thank Vijay Bhattiprolu for many helpful and interesting discussions. The second author was supported by the National Science Foundation under Grant No. DMS-1926686. The third author acknowledges the support of the Natural Sciences and Engineering Research Council of Canada (NSERC), [funding reference number RGPIN-2023-03726]. Cette recherche a été partiellement financée par le Conseil de recherches en sciences naturelles et en génie du Canada (CRSNG), [numéro de référence RGPIN-2023-03726].

# References

[AHK18] Kareem Adiprasito, June Huh, and Eric Katz. "Hodge theory for combinatorial geometries". In: *Annals of Mathematics* 188.2 (2018), pp. 381–452 (cit. on p. 1).

[Ali+21] Yeganeh Alimohammadi, Nima Anari, Kirankumar Shiragur, and Thuy-Duong Vuong. "Fractionally log-concave and sector-stable polynomials: counting planar matchings and more". In: *Proceedings of the 53rd Annual ACM SIGACT Symposium on Theory of Computing*. 2021, pp. 433–446 (cit. on p. 1).

[Ana+18a] Nima Anari, Kuikui Liu, Shayan Oveis Gharan, and Cynthia Vinzant. "Log-concave polynomials III: Mason's ultra-log-concavity conjecture for independent sets of matroids". In: *arXiv preprint arXiv:1811.01600* (2018) (cit. on p. 1).

[Ana+18b] Nima Anari, Tung Mai, Shayan Oveis Gharan, and Vijay V Vazirani. "Nash social welfare for indivisible items under separable, piecewise-linear concave utilities". In: *Proceedings of the Twenty-Ninth Annual ACM-SIAM Symposium on Discrete Algorithms*. SIAM. 2018, pp. 2274–2290 (cit. on p. 1).

[Ana+19] Nima Anari, Kuikui Liu, Shayan Oveis Gharan, and Cynthia Vinzant. "Log-concave polynomials II: High-dimensional walks and an FPRAS for counting bases of a matroid". In: *Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing*. 2019, pp. 1–12 (cit. on p. 1).

[AO14] Nima Anari and Shayan Oveis Gharan. "The Kadison-Singer problem for strongly Rayleigh measures and applications to Asymmetric TSP". In: *arXiv preprint arXiv:1412.1143* (2014) (cit. on p. 1).

[AO17] Nima Anari and Shayan Oveis Gharan. "A generalization of permanent inequalities and applications in counting and optimization". In: *Proceedings of the 49th Annual ACM SIGACT Symposium on Theory of Computing*. 2017, pp. 384–396 (cit. on p. 1).

[AOV18] Nima Anari, Shayan Oveis Gharan, and Cynthia Vinzant. "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. 2018, pp. 35–46 (cit. on p. 1).

[Asa+10] Arash Asadpour, Michel X Goemans, Aleksander Madry, Shayan Oveis Gharan, and Amin Saberi. "An O (log n/log log n)-approximation algorithm for the asymmetric traveling salesman problem". In: *Proceedings of the twenty-first annual ACM-SIAM symposium on Discrete algorithms*. 2010, pp. 379–389 (cit. on p. 2).

[Bar08] Alexander Barvinok. "Enumerating contingency tables via random permanents". In: *Combinatorics, Probability and Computing* 17.1 (2008), pp. 1–19 (cit. on p. 10).

[Bar16] Alexander Barvinok. "Computing the permanent of (some) complex matrices". In: *Foundations of Computational Mathematics* 16 (2016), pp. 329–342 (cit. on p. 1).

[BBL09] Julius Borcea, Petter Brändén, and Thomas Liggett. "Negative dependence and the geometry of polynomials". In: *Journal of the American Mathematical Society* 22.2 (2009), pp. 521–567 (cit. on p. 4).

[BH20] Petter Brändén and June Huh. "Lorentzian polynomials". In: *Annals of Mathematics* 192.3 (2020), pp. 821–891 (cit. on p. 1).

[BLP23] Petter Brändén, Jonathan Leake, and Igor Pak. "Lower bounds for contingency tables via Lorentzian polynomials". In: *Israel Journal of Mathematics* 253.1 (2023), pp. 43–90 (cit. on p. 1).

[Brä18] Petter Brändén. "Hyperbolic polynomials and the Kadison-Singer problem". In: *arXiv preprint arXiv:1809.03255* (2018) (cit. on p. 1).

[Ego81] Gregory P Egorychev. "The solution of van der Waerden's problem for permanents". In: *Advances in Mathematics* 42.3 (1981), pp. 299–305 (cit. on p. 1).

[Fal81] Dmitry I Falikman. "Proof of the van der Waerden conjecture regarding the permanent of a doubly stochastic matrix". In: *Mathematical notes of the Academy of Sciences of the USSR* 29 (1981), pp. 475–479 (cit. on p. 1).

[GL21] Leonid Gurvits and Jonathan Leake. "Capacity lower bounds via productization". In: *Proceedings of the 53rd Annual ACM SIGACT Symposium on Theory of Computing*. 2021, pp. 847–858 (cit. on pp. 1, 2, 4, 5, 9, 10).

[Gur06a] Leonid Gurvits. "Hyperbolic polynomials approach to Van der Waerden/Schrijver-Valiant like conjectures: sharper bounds, simpler proofs and algorithmic applications". In: *Proceedings of the thirty-eighth annual ACM symposium on Theory of computing*. 2006, pp. 417–426 (cit. on p. 1).

[Gur06b] Leonid Gurvits. "The van der Waerden conjecture for mixed discriminants". In: *Advances in Mathematics* 200.2 (2006), pp. 435–454 (cit. on p. 1).

[Gur09a] Leonid Gurvits. "A polynomial-time algorithm to approximate the mixed volume within a simply exponential factor". In: *Discrete & Computational Geometry* 41 (2009), pp. 533–555 (cit. on p. 1).

[Gur09b] Leonid Gurvits. "On multivariate Newton-like inequalities". In: *Advances in Combinatorial Mathematics: Proceedings of the Waterloo Workshop in Computer Algebra 2008*. Springer. 2009, pp. 61–78 (cit. on pp. 1, 4, 6).

[Hoe56] Wassily Hoeffding. "On the distribution of the number of successes in independent trials". In: *The Annals of Mathematical Statistics* (1956), pp. 713–721 (cit. on p. 22).

[KKO21] Anna R. Karlin, Nathan Klein, and Shayan Oveis Gharan. "A (Slightly) Improved Approximation Algorithm for Metric TSP". In: *STOC*. ACM, 2021 (cit. on pp. 1–3, 7–9, 19–25).

[KKO22] Anna Karlin, Nathan Klein, and Shayan Oveis Gharan. "A (Slightly) Improved Bound on the Integrality Gap of the Subtour LP for TSP". In: *FOCS*. IEEE Computer Society, 2022, pp. 844–855 (cit. on pp. 3, 9, 25).

[KKO23] Anna R. Karlin, Nathan Klein, and Shayan Oveis Gharan. "A Deterministic Better-than-3/2 Approximation Algorithm for Metric TSP". In: *Integer Programming and Combinatorial Optimization*. Ed. by Alberto Del Pia and Volker Kaibel. Cham: Springer International Publishing, 2023, pp. 261–274. ISBN: 978-3-031-32726-1 (cit. on pp. 1, 9).

[MSS15] Adam W Marcus, Daniel A Spielman, and Nikhil Srivastava. "Interlacing families ii: Mixed characteristic polynomials and the kadison—singer problem". In: *Annals of Mathematics* (2015), pp. 327–350 (cit. on p. 1).

[OSS11] Shayan Oveis Gharan, Amin Saberi, and Mohit Singh. "A randomized rounding approach to the traveling salesman problem". In: *2011 IEEE 52nd Annual Symposium on Foundations of Computer Science*. IEEE. 2011, pp. 550–559 (cit. on pp. 1, 2).

[SV17] Damian Straszak and Nisheeth K Vishnoi. "Real stable polynomials and matroids: Optimization and counting". In: *Proceedings of the 49th Annual ACM SIGACT Symposium on Theory of Computing*. 2017, pp. 370–383 (cit. on p. 1).

# A Improving the TSP Bound

## A.1 Further Background on TSP

In [KKO21], instead of only looking at vertices as we did in Section 4, they consider sets of edges going between two *sets* of vertices. The notation $\mathbf{e} = (\mathbf{u}, \mathbf{v})$ is used, for $\mathbf{e} \subseteq E$, $\mathbf{u}, \mathbf{v} \subseteq V$. Formally, if $\mathbf{e} = (\mathbf{u}, \mathbf{v})$ then $\mathbf{e} = \{e = \{u, v\} \in E : |e \cap \mathbf{u}| = |e \cap \mathbf{v}| = 1\}$. They call $\mathbf{e}$ a "top edge bundle."[3] $\mathbf{e} = (\mathbf{u}, \mathbf{v})$ is only defined for sets $\mathbf{u}, \mathbf{v}$ for which $2 \leq x(\delta(\mathbf{u})) \leq 2 + \epsilon_\eta$ and $2 \leq x(\delta(\mathbf{v})) \leq 2 + \epsilon_\eta$, where $\epsilon_\eta \leq 10^{-10}$ is a very small constant.

In [KKO21], the event that the sets $\mathbf{u}$ and $\mathbf{v}$ are subtrees of $T$ is ubiquitous and therefore we always work in the conditional measure $\mu' = \mu_{|\mathbf{u},\mathbf{v} \text{ trees}}$. By Lemma 2.23 in [KKO21], $\mathbf{u}, \mathbf{v}$ are subtrees with probability at least $1-2\epsilon_\eta$, $\mu'$ is a strongly Rayleigh distribution, and if $x'_e = \mathbb{P}_{T \sim \mu'}[e \in T]$ then $x(F) \leq x'(F) \leq x(F) + 2\epsilon_\eta$ for any $F \subseteq \delta(\mathbf{u}) \cup \delta(\mathbf{v})$. Thus, since $\epsilon_\eta$ is such a small constant, there is very little difference between $\mu$ and $\mu'$.

For a top edge bundle $\mathbf{e} = (\mathbf{u}, \mathbf{v})$, they partition the edges of $\mathbf{u}$ into sets $A, B, C$. This is called the "$(A, B, C)$ degree partition" of $\mathbf{u}$ (this is analogously done for $\mathbf{v}$). The details of this operation are not relevant in this work. What is relevant is that in the measure $\mu'$ we have the bounds

$$\begin{aligned} x'(A), x'(B) &\in [1 - \epsilon_{1,1}, 1 + 3\epsilon_\eta], \\ x'(C) &\leq 2\epsilon_{1/1} + 3\epsilon_\eta. \end{aligned} \tag{4}$$

where $\epsilon_{1/1} = \epsilon_{1/2}/12$ and $\epsilon_{1/2} = 0.0002$ are small constants that will be relevant in this section.[4] $\mathbf{e}$ is called a **half edge bundle** if $\frac{1}{2} - \epsilon_{1/2} \leq x(\mathbf{e}) \leq \frac{1}{2} + \epsilon_{1/2}$. As already observed in Section 4, these edge bundles are particularly difficult to deal with. Thus they are the subject of many probabilistic statements in [KKO21].

The final relevant definition is the event that a top edge bundle $\mathbf{e} = (\mathbf{u}, \mathbf{v})$ is "2-1-1 happy with respect to $\mathbf{u}$." Where for sets of edges $F, T \subseteq E$ we let $F_T = |F \cap T|$, this indicates that

$$A_T = 1, B_T = 1, C_T = 0, \delta(\mathbf{v})_T = 2, \mathbf{u}, \mathbf{v} \text{ trees}$$

where $A, B$ and $C$ are the $(A, B, C)$ degree partition of $\mathbf{u}$. The event "2-1-1 happy with respect to $\mathbf{v}$" is analogous, and also relevant, however since the assumptions on the partitions of $\mathbf{u}$ and $\mathbf{v}$ are the same we do not distinguish between them in this work.

The lemmas we improve all involve showing that the event that an edge is 2-1-1 happy (with respect to $\mathbf{u}$ or $\mathbf{v}$) has a non-negligible probability.

We will require one additional fact about our spanning tree distribution, which is a simple consequence of negative association and homogeneity:

**Lemma A.1** (Explained in e.g. Section 2 of [KKO21])**.** *Let $\mu : 2^E \to \mathbb{R}_{\geq 0}$ be a strongly Rayleigh distribution over spanning trees with $x_e = \mathbb{P}_{T \sim \mu}[e \in T]$ as above. Let $A, B \subseteq E$ be disjoint sets. Then, $\mathbb{E}_{T \sim \mu}[A_T \mid B_T = 0] \leq x(A) + x(B)$. Furthermore, the conditional measure $\mu_{|B_T = 0}$ is strongly Rayleigh. Similarly, for an edge $e$, we can condition $e \in T$ and the measure $\mu_{|e \in T}$ is SR, and for any set $A \subseteq E$ with $e \notin A$ we have $x(A) - x_e \leq \mathbb{E}_{T \sim \mu}[A_T \mid e \in T] \leq x(A)$.*

## A.2 Using Capacity to Improve Probabilistic Bounds

Theorem 4.1 and Corollary 4.2 can be applied straightforwardly to strengthen and simplify the proof of a number of lemmas in [KKO21]. In this section we demonstrate this for Lemmas 5.21 and 5.22, as they are bottlenecks for the approximation factor. However, they can also be applied to simplify and improve the factor for Lemma 5.7 and achieve simpler proofs for Lemmas 5.16 and 5.17 at the price of a slightly worse constant. (Of course, with more effort, one can likely use these statements to strengthen the bounds for these as well similar to what is done for Lemma 5.24.)

[3] They also consider "bottom edge bundles," which have a slightly different definition, however they do not arise in the lemmas we consider.

[4] Note in [KKO21], these bounds are given with respect to $x$, not $x'$. We use $x'$ for convenience since as noted everything is done in the conditional measure $\mu'$.

We then show that Lemma 5.23 can be easily strengthened using existing arguments. Finally, we use Corollary 4.2 to reduce Lemma 5.24 to a special case in which ad hoc methods can further improve the bound.

### A.2.1 Improving Lemma 5.21

Here we improve upon the bound of $0.005\epsilon_{1/2}^2$ via a straightforward application of our new capacity bound, Corollary 4.2.

**Lemma A.2.** *Let $\mathbf{e} = (\mathbf{u}, \mathbf{v})$ be a top edge bundle such that $x(\mathbf{e}) \leq 1/2 - \epsilon_{1/2}$. If $\epsilon_{1/2} \leq 0.001$, $\epsilon_{1/1} \leq \epsilon_{1/2}/12$, and $\epsilon_\eta \leq \epsilon_{1/2}^2$ then $\mathbf{e}$ is 2-1-1 happy with probability at least $0.039\epsilon_{1/2}^2$.*

*Proof.* Consider the $(A, B, C)$ degree partitioning of $\mathbf{u}$ (the case for $\mathbf{v}$ is exactly the same). We first condition on $\mu'$, $C_T = 0$ and $((A \cap C) \cup (B \cap C))_T = 0$. Let $V = \delta(\mathbf{u})$. We now want to bound the probability that $A_T = B_T = 1$ and $V_T = 2$ in the resulting measure $\nu$. As we are working in the same setting as the proof of 5.21 in [KKO21], we re-use their bounds on the expectations:

$$\begin{array}{rcccl} -0.5 & \leq & \mathbb{E}_\nu[A_T] - 1 & \leq & 0.5 \\ -0.5 & \leq & \mathbb{E}_\nu[B_T] - 1 & \leq & 0.5 \\ -0.5 & \leq & \mathbb{E}_\nu[V_T] - 2 & \leq & 0.01 \\ -0.5 & \leq & \mathbb{E}_\nu[A_T + B_T] - 2 & \leq & 0.01 \\ -1 + 1.8\epsilon_{1/2} & \leq & \mathbb{E}_\nu[A_T + V_T] - 3 & \leq & 0.01 \\ -1 + 1.8\epsilon_{1/2} & \leq & \mathbb{E}_\nu[B_T + V_T] - 3 & \leq & 0.01 \\ -1 + 1.75\epsilon_{1/2} & \leq & \mathbb{E}_\nu[A_T + B_T + V_T] - 4 & \leq & -0.49 \end{array}$$

Now, note that since $V$ defines a cut in the graph and we have a distribution over spanning trees, $V_T \geq 1$ with probability 1. Therefore we can consider the random variable $V_T - 1$ instead (this corresponds to dividing the generating polynomial by the variable representing $V_T$). Thus in Corollary 4.2 we may choose (considering $\kappa = (1, 1, 1)$ on the variables $A_T, B_T, V_T - 1$):

$$\epsilon_1 = 0.5, \qquad \epsilon_2 = 1.8\epsilon_{1/2}, \qquad \epsilon_3 = 1.75\epsilon_{1/2}$$

which implies (using that $\nu$ occurs with probability at least $1/2$):

$$0.5 \cdot \mathbb{P}_\nu \left[A_T = 1, B_T = 1, V_T = 2\right] \geq 0.5 \cdot e^{-3} \cdot 0.5 \cdot 1.8\epsilon_{1/2} \cdot 1.75\epsilon_{1/2} \geq 0.0392 \cdot \epsilon_{1/2}^2.$$

as desired. □

We also note that this bound is tight up to the constant factor for strongly Rayleigh distributions (where one can set $\gamma \approx \epsilon_{1/2}$):

**Lemma A.3.** *There is a strongly Rayleigh distribution $\mu$ for which $x(A) = x(B) = 1, x(V) = 2$ and $x((A \cap V) \cup (B \cap V)) = \frac{1}{2} - \gamma$ such that*

$$\mathbb{P}_{T \sim \mu}\left[A_T = 1, B_T = 1, V_T = 2\right] = (1 + 2\gamma)\gamma^2$$

*Proof.* We will let $A$ be represented by $w$ and $x$, $B$ by $y$, and $V$ by $w$ and $z$. Consider the real stable generating polynomial

$$p_\mu(x, y, z, w) = z\left(\left(\frac{1}{2} + \gamma\right) + \left(\frac{1}{2} - \gamma\right)w\right)\left(\frac{1}{2}x + \frac{1}{2}z\right)(\gamma x + \gamma z + (1 - 2\gamma)y)\left((1 - 2\gamma) + 2\gamma y\right)$$

One can check that the expectations are correct. Now we calculate the coefficient of $wyz$ and $xyz^2$, corresponding to the desired event. The distribution $\mu$ consists of 5 independent choices corresponding to each of the 5 terms. The first is of course trivial, we always take $z$. In the second, suppose we take $w$, so we have collected terms $wz$. Then, note we must take some variable from the third and fourth terms. So, $wyz$ has a coefficient of 0.

Therefore we must take $(\frac{1}{2} + \gamma)$ in the second term, and we must now collect $xyz$ from the remainder. It follows that we must collect $y$ in the last term, as there are three terms and three variables, so we now have $(\frac{1}{2} + \gamma)2\gamma yz$ and must collect $xz$. There are two ways to do this, each with a coefficient of $\frac{\gamma}{2}$. Therefore the coefficient of $xyz^2$ is $(1 + \gamma)\gamma^2$. □

### A.2.2 Improving Lemma 5.22

We now improve upon Lemma 5.22 from [KKO21], which had a bound of $0.006\epsilon_{1/2}^2$. This is again a straightforward application of Corollary 4.2.

**Lemma A.4.** *Let $\mathbf{e} = (\mathbf{u}, \mathbf{v})$ be a top edge bundle such that $x_{\mathbf{e}} \geq 1/2 + \epsilon_{1/2}$. If $\epsilon_{1/2} \leq 0.001$, then, $\mathbf{e}$ is 2-1-1 happy with respect to $u$ with probability at least $0.038 \cdot \epsilon_{1/2}^2$.*

*Proof.* The setup of this proof is similar to Lemma A.2. We again look at the $(A, B, C)$ degree partition of $\mathbf{u}$. First, we condition on $\mu'$ and $C_T = 0$. Then, we condition on $\mathbf{u} \cup \mathbf{v}$ to be a tree. All these events occur with probability at least $1/2$. Let $\nu$ be the resulting measure. Under $\nu$, notice that there is exactly one edge between $\mathbf{u}$ and $\mathbf{v}$, or in other words that $(A \cap \delta(\mathbf{v})) \cup (B \cap \delta(\mathbf{v}))$ is 1.

Therefore, to obtain $A_T = B_T = 1$ and $\delta(\mathbf{v})_T = 2$, where $V = \delta(\mathbf{v}) \setminus \delta(\mathbf{u})$ it suffices to prove that $A_T = B_T = 1$ and $V_T = 1$ (where note that $A, B, V$ are now disjoint).

As this is the same setup as Lemma 5.22 in [KKO21], we simply re-use their bounds on the expectations:

$$\begin{array}{rcccl}
-0.5 & \leq & \mathbb{E}_\nu[A_T] - 1 & \leq & 0.5 \\
-0.5 & \leq & \mathbb{E}_\nu[B_T] - 1 & \leq & 0.5 \\
-0.005 & \leq & \mathbb{E}_\nu[V_T] - 1 & \leq & 0.5 \\
-0.005 & \leq & \mathbb{E}_\nu[A_T + B_T] - 2 & \leq & 0.5 \\
-0.01 & \leq & \mathbb{E}_\nu[A_T + V_T] - 2 & \leq & 1 - 1.75\varepsilon_{1/2} \\
-0.01 & \leq & \mathbb{E}_\nu[B_T + V_T] - 2 & \leq & 1 - 1.75\varepsilon_{1/2} \\
-0.01 & \leq & \mathbb{E}_\nu[A_T + B_T + V_T] - 3 & \leq & 1 - 1.75\varepsilon_{1/2}
\end{array}$$

Therefore in Corollary 4.2 we can set

$$\epsilon_1 = 0.5, \qquad \epsilon_2 = 1.75\varepsilon_{1/2}, \qquad \epsilon_3 = 1.75\varepsilon_{1/2}$$

which implies

$$0.5 \cdot \mathbb{P}_\nu\left[A_T = 1, B_T = 1, V_T = 1\right] \geq 0.5 \cdot e^{-3} \cdot 0.5 \cdot 1.75\varepsilon_{1/2} \cdot 1.75\varepsilon_{1/2} \geq 0.038 \cdot \varepsilon_{1/2}^2$$

as desired. □

### A.2.3 Improving Lemma 5.23

First we recall some notation from [KKO21]. For $\mathbf{e} = (\mathbf{u}, \mathbf{v})$ and a set $A \subseteq E$, (where recall $\mathbf{e} \subseteq E$) we let $A_{-\mathbf{e}} = A \setminus \mathbf{e}$ and $x_{\mathbf{e}(A)} = x(A \cap \mathbf{e})$.

Unfortunately, this lemma does not easily fit into the framework from this paper, as it does not follow from expectation information. Instead, to improve Lemma 5.23, we first note that Lemma A.1 in [KKO21] can be parameterized in the following way. We simply replace the constant 5 by $k$ in the statement and exactly follow the previous proof.

**Lemma A.5.** *For a good half top edge bundle $\mathbf{e} = (\mathbf{u}, \mathbf{v})$, let $A, B, C$ be the degree partitioning of $\delta(u)$, and let $V = \delta(\mathbf{v})_{-\mathbf{e}}$. If $\epsilon_{1/2} \leq 0.001$, $x_{\mathbf{e}(B)} \leq \epsilon_{1/2}$, and $\mathbb{P}\left[(A_{-\mathbf{e}})_T + V_T \leq 1\right] \geq k\epsilon_{1/2}$ then $\mathbf{e}$ is 2-1-1 good,*

$$\mathbb{P}\left[\mathbf{e} \text{ 2-1-1 happy w.r.t. } u\right] \geq 0.001(\min\{100, k\} - 0.2)\epsilon_{1/2}^2$$

Since we do not modify the proof in any meaningful way, we defer it to Appendix A. Now we can show that in fact without any modifications, Lemma 5.23 is true with a much larger constant as long as one sets $\epsilon_{1/2} = 0.0002$. In particular:

**Lemma A.6.** *Let $\mathbf{e} = (\mathbf{v}, \mathbf{u})$ and $\mathbf{f} = (\mathbf{v}, \mathbf{w})$ be good half top edge bundles and let $A, B, C$ be the degree partitioning of $\delta(v)$ such that $x_{\mathbf{e}(B)}, x_{\mathbf{f}(B)} \leq \epsilon_{1/2}$. Then, if $\epsilon_{1/2} \leq 0.0002$, one of $\mathbf{e}, \mathbf{f}$ is 2-1-1 happy with probability at least $0.0498\epsilon_{1/2}^2$.*

*Proof.* Let $U = \delta(\mathbf{u}) \setminus \mathbf{e}$ and $A_{-\mathbf{e}} = A \setminus \mathbf{e}$. Then, Lemma 5.23 in [KKO21] shows that $\mathbb{P}\left[U_T + (A_{-\mathbf{e}})_T \leq 1\right] \geq 0.01$. Using that $\epsilon_{1/2} \leq 0.0002$, this is at least $50\epsilon_{1/2}$. Now by Lemma A.6 we obtain a bound of $0.001(50 - 0.2)\epsilon_{1/2}^2$ as desired. □

### A.2.4 Improving Lemma 5.24

This lemma also does not follow directly from expectation information. However, we show that our new capacity bound is still helpful in improving the bound. We include the new part of the lemma here and defer the rest of the proof to the appendix as it is similar to [KKO21].

**Lemma A.7.** *Let $\mathbf{e} = (\mathbf{u}, \mathbf{v})$ be a good half edge bundle and let $A, B, C$ be the degree partitioning of $\delta(u)$. If $\epsilon_{1/2} \leq 0.001$ and $x_{\mathbf{e}(A)}, x_{\mathbf{e}(B)} \geq \epsilon_{1/2}$, then*

$$\mathbb{P}\left[\mathbf{e} \text{ 2-1-1 happy w.r.t } u\right] \geq 0.0498\epsilon_{1/2}^2.$$

*Proof.* Condition $C_T$ to be zero, $u, v$ and $u \cup v$ be trees. Let $\nu$ be the resulting measure. Let $X = A_{-\mathbf{e}} \cup B_{-\mathbf{e}}, Y = \delta(v)_{-\mathbf{e}}$. Then

$$\mathbb{E}_\nu\left[X_T\right], \mathbb{E}_\nu\left[Y_T\right] \in [1 - 3\epsilon_{1/1}, 1.5 + \epsilon_{1/2} + 2\epsilon_{1/1} + 3\epsilon_\eta] \subset [0.995, 1.51]$$
$$\mathbb{E}_\nu\left[X_T + Y_T\right] \in [2.5 - 3\epsilon_{1/1} - 3\epsilon_{1/2}, 3 + 2\epsilon_{1/2} + 2\epsilon_{1/1} + 4\epsilon_\eta] \in [2.495, 3.01]$$

Therefore using Corollary 4.2, if $\mathbb{E}_\nu\left[X_T + Y_T\right] \leq 2.999$, we could set $\epsilon_1 = 0.49, \epsilon_2 = 0.001$ and obtain a bound of $\mathbb{P}_\nu\left[X_T = Y_T = 1\right] \geq e^{-2}(0.49)(0.001) \geq 0.06\epsilon_{1/2}$. Therefore, we either have $\mathbb{P}_\nu\left[X_T = Y_T = 1\right] \geq 0.06\epsilon_{1/2}$, or it is the case that:

$$\mathbb{E}_\nu\left[X_T\right], \mathbb{E}_\nu\left[Y_T\right] \in [1.489, 1.51], \qquad \mathbb{E}_\nu\left[X_T + Y_T\right] \in [2.999, 3.01]$$

Given this, we obtain the lemma using Lemma A.12. □

## A.3 Further Probabilistic Statements

While in the previous section, we focused on the usage of Corollary 4.2, here we require some of the ad hoc methods of [KKO21] to make progress.

A key fact about real stable polynomials we will use in the remaining statements is that their univariate restrictions are real rooted. As a consequence, the following standard fact can be shown:

**Lemma A.8.** *Let $\mu : 2^E \to \mathbb{R}_{\geq 0}$ be a strongly Rayleigh distribution and let $F \subseteq E$. Then there exists independent Bernoulli random variables $B_1, \ldots, B_m$ such that $\mathbb{P}_{T\sim\mu}\left[|F \cap T| = k\right] = \mathbb{P}\left[\sum_{i=1}^m B_i = k\right]$ for all $k \in \mathbb{Z}_{\geq 0}$.*

Bernoulli random variables are quite easy to work with due to the following theorem of Hoeffding.

**Theorem A.9** (Corollary 2.1 from [Hoe56])**.** *Let $g : \{0, 1, \ldots, n\} \to \mathbb{R}$ and $0 \leq q \leq n$ for some integer $n \geq 0$. Let $B_1, \ldots, B_n$ be $n$ independent Bernoulli random variables with success probabilities $p_1, \ldots, p_n$, where $\sum_{i=1}^n p_n = q$ that minimizes (or maximizes)*

$$\mathbb{E}[g(B_1 + \cdots + B_n)]$$

*over all such distributions. Then, $p_1, \ldots, p_n \in \{0, x, 1\}$ for some $0 < x < 1$. In particular, if only $m$ of $p_i$'s are nonzero and $\ell$ of $p_i$'s are 1, then the remaining $m - \ell$ are $\frac{q-\ell}{m-\ell}$.*

Finally, we will need the following two statements from [KKO21].

**Lemma A.10** (Lemma 5.4 from [KKO21])**.** *Given a strongly Rayleigh distribution $\mu : 2^{[n]} \to \mathbb{R}_{\geq 0}$, let $A, B$ be two (nonnegative) random variables corresponding to the number of elements sampled from two* disjoint *sets such that $\mathbb{P}\left[A + B = n\right] > 0$ where $n = n_A + n_B$. Then,*

$$\mathbb{P}\left[A \geq n_A | A + B = n\right] = \mathbb{P}\left[B \leq n_B | A + B = n\right] \geq \mathbb{P}\left[A \geq n_A\right]\mathbb{P}\left[B \leq n_B\right], \tag{5}$$
$$\mathbb{P}\left[A \leq n_A | A + B = n\right] = \mathbb{P}\left[B \geq n_B | A + B = n\right] \geq \mathbb{P}\left[A \leq n_A\right]\mathbb{P}\left[B \geq n_B\right]. \tag{6}$$

**Corollary A.11** (Corollary 5.5 from [KKO21]). *Let $\mu : 2^{[n]} \to \mathbb{R}_{\geq 0}$ be a SR distribution. Let $A, B$ be two random variables corresponding to the number of elements sampled from two disjoint sets of elements such that $A \geq k_A$ with probability 1 and $B \geq k_B$ with probability 1. If $\mathbb{P}[A \geq n_A], \mathbb{P}[B \geq n_B] \geq \epsilon_1$ and $\mathbb{P}[A \leq n_A], \mathbb{P}[B \leq n_B] \geq \epsilon_2$, then, letting $n'_A = n_A - k_A, n'_B = n_B - k_B$,*

$$\mathbb{P}[A = n_A | A + B = n_A + n_B] \geq \epsilon \min\{\frac{1}{n'_A + 1}, \frac{1}{n'_B + 1}\},$$
$$\mathbb{P}[A = n_A | A + B = n_A + n_B] \geq \min\left\{p_m, \epsilon(1 - (\epsilon/p_m)^{1/\max\{n'_A, n'_B\}})\right\}$$

*where $\epsilon = \epsilon_1\epsilon_2$ and $p_m \leq \max_{k_A \leq k \leq n_A + n_B - k_B} \mathbb{P}[A = k | A + B = n_A + n_B]$ is a lower bound on the mode of $A$.*

The following proof is essentially identical to the the proof of Lemma A.1 in [KKO21], we simply parameterize the statement by $k$.

**Lemma A.5.** *For a good half top edge bundle $\mathbf{e} = (\mathbf{u}, \mathbf{v})$, let $A, B, C$ be the degree partitioning of $\delta(u)$, and let $V = \delta(\mathbf{v})_{-\mathbf{e}}$. If $\epsilon_{1/2} \leq 0.001$, $x_{\mathbf{e}(B)} \leq \epsilon_{1/2}$, and $\mathbb{P}[(A_{-\mathbf{e}})_T + V_T \leq 1] \geq k\epsilon_{1/2}$ then $\mathbf{e}$ is 2-1-1 good,*

$$\mathbb{P}[\mathbf{e} \text{ 2-1-1 happy w.r.t. } u] \geq 0.001(\min\{100, k\} - 0.2)\epsilon_{1/2}^2$$

*Proof.* We first condition on $\mu'$, $C_T = 0$, $u \cup v$ to be a tree and let $\nu$ be the resulting SR measure on edges in $A, B, V$. Since $\mathbf{e}$ is 2-2 good, by Lemma 5.15 in [KKO21] and negative association,

$$\mathbb{P}_\nu[(\delta(u)_{-\mathbf{e}})_T + V_T \leq 2] \geq \mathbb{P}[(\delta(u)_{-\mathbf{e}})_T + V_T \leq 2] - \mathbb{P}[C_T = 0] \geq 0.4\epsilon_{1/2} - 2\epsilon_{1/1} - \epsilon_\eta \geq 0.22\epsilon_{1/2},$$

where we used $\epsilon_{1/1} \leq \epsilon_{1/2}/12$. Letting $p_i = \mathbb{P}[(\delta(u)_{-\mathbf{e}})_T + V_T = i]$, we therefore have $p_{\leq 2} \geq 0.22\epsilon_{1/2}$. In addition, by Theorem A.9, $p_3 \geq 1/4$. If $p_2 < 0.2\epsilon_{1/2}$, then from $p_2/p_3 \leq 0.8\epsilon_{1/2}$, we could use log-concavity to derive a contradiction to $p_{\leq 2} \geq 0.22\epsilon_{1/2}$. Therefore, we must have

$$\mathbb{P}_\nu[A_T + B_T + V_T = 3] = \mathbb{P}_\nu[(\delta(u)_{-\mathbf{e}})_T + V_T = 2] \geq 0.2\epsilon_{1/2}.$$

Next, notice since $\mathbb{P}[u, v, u \cup v \text{ trees}, C_T = 0] \geq 0.49$, by the lemma's assumption, $\mathbb{P}_\nu[\mathbf{e}(B)] \leq 2.01\epsilon_{1/2}$. Therefore,

$$\mathbb{E}_\nu[B_T + V_T] \leq x(V) + x(B) + 1.01\epsilon_{1/2} + 2\epsilon_{1/1} + \epsilon_\eta \leq 2.51.$$

So, by Markov, $\mathbb{P}_\nu[B_T + V_T \leq 2] \geq 0.15$. Finally, by negative association,

$$\mathbb{P}_\nu[A_T + V_T \leq 2] \geq \mathbb{P}_\nu[(A_{-\mathbf{e}})_T + V_T \leq 1] \geq \mathbb{P}[(A_{-\mathbf{e}})_T + V_T \leq 1] - \mathbb{P}[C_T = 0] \geq (k - 0.2)\epsilon_{1/2}$$

where we used the lemma's assumption.

Using arguments from the proof of Lemma 5.21 in [KKO21], we have

$$\mathbb{P}_\nu[A_T + B_T = 2 \mid A_T + B_T + V_T = 3] \geq 0.12$$

Now by Lemma A.10,

$$\begin{aligned}\mathbb{P}_\nu[B_T \geq 1 \mid A_T + B_T + V_T = 3] &= \mathbb{P}_\nu[A_T + B_T \leq 2 \mid A_T + B_T + V_T = 3] \\ &\geq \mathbb{P}_\nu[A_T + B_T \leq 2]\,\mathbb{P}_\nu[V_T \geq 1] \geq 0.63(k - 0.2)\epsilon_{1/2}\end{aligned}$$

using that similar to Lemma 5.21 in [KKO21] we can show $\mathbb{P}_\nu[V_T \geq 1] \geq 0.63$. We also use similar arguments to show $\mathbb{P}_\nu[B_T \leq 1 \mid A_T + B_T + V_T = 3] \geq 0.147$.

Therefore, by Corollary A.11, $\mathbb{P}_\nu[B_T = 1 \mid A_T + B_T = 2, V_T = 1] \geq 0.147 \cdot 0.63(k - 0.2)\epsilon_{1/2} \cdot p_m$ where $p_m \leq \max_{k \in \{0,1,2\}} \mathbb{P}[B = k \mid A_T + B_T = 2, V_T = 1]$. For $k \leq 100$ and $\epsilon_{1/2} \leq 0.0002$, our bound for $\mathbb{P}_\nu[B_T = 1 \mid A_T + B_T = 2, V_T = 1]$ is at most 0.001 for $p_m = 1$. Therefore, we may assume it is at most 0.001. By log concavity we can set $p_m = 0.998$. So,

$$\mathbb{P}[\mathbf{e} \text{ 2-1-1 happy}] \geq (0.092(\max\{k, 100\} - 0.2)\epsilon_{1/2})0.12(0.2\epsilon_{1/2})0.498 \geq 0.001(\max\{k, 100\} - 0.2)\epsilon_{1/2}^2$$

as desired. □

Here we use the additional assumptions provided by Lemma A.7 to improve upon the original bound.

**Lemma A.12.** *Let $\mathbf{e} = (\mathbf{u}, \mathbf{v})$ be a good half edge bundle and let $A, B, C$ be the degree partitioning of $\delta(u)$. Suppose $\epsilon_{1/2} \leq 0.001$ and $x_{\mathbf{e}(A)}, x_{\mathbf{e}(B)} \geq \epsilon_{1/2}$. Let $\nu$ be the measure conditioned on $C_T = 0$, $u, v$ and $u \cup v$ be trees. Suppose that either $\mathbb{P}_\nu[X_T = Y_T = 1] \geq 0.06\epsilon_{1/2}$, or*

$$\mathbb{E}_\nu[X_T], \mathbb{E}_\nu[Y_T] \in [1.489, 1.51], \qquad \mathbb{E}_\nu[X_T + Y_T] \in [2.999, 3.01]$$

*Then,*

$$\mathbb{P}[\mathbf{e} \text{ 2-1-1 happy w.r.t } u] \geq 0.0485\epsilon_{1/2}^2.$$

*Proof.* We break the proof into two stages.

**First we show the bound given $\mathbb{P}_\nu[X_T = Y_T = 1] \geq 0.0498\epsilon_{1/2}$.** Let $\mathcal{E}$ be the event $\{X_T = Y_T = 1\}$ in the conditional measure $\nu$. Note that in $\nu$ we always choose exactly 1 edge from the $\mathbf{e}$ bundle and that is independent of edges in $X, Y$, in particular the above event. Therefore, we can correct the parity of $A, B$ by choosing from $e_A$ or $e_B$. It follows that

$$\mathbb{P}[\mathbf{e} \text{ 2-1-1 happy w.r.t } u] \geq \mathbb{P}_\nu[\mathcal{E}](1.99\epsilon_{1/2})0.49 \geq 0.0485\epsilon_{1/2}^2,$$

where we used that the probability $C_T = 0$ and $u, v, u \cup v$ are subtrees with probability at least 0.49, $\mathbb{E}_\nu[\mathbf{e}(A)_T] \geq 1.99\epsilon_{1/2}$, and $\mathbb{E}_\nu[\mathbf{e}(B)_T] \geq 1.99\epsilon_{1/2}$. To see why these latter inequalities hold, observe that conditioned on $u, v$ trees, we always sample at most one edge between $u, v$. Therefore, since under $\nu$ we choose exactly one edge between $u, v$, the probability of choosing from $e(A)$ (and similarly choosing from $\mathbf{e}(B)$) is at least

$$\frac{\mathbb{E}\mathbf{e}(A)_T | u, v \text{ trees}, C_T = 0}{\mathbb{P}[\mathbf{e} | u, v \text{ trees}, C_T = 0]} \geq \frac{x_{\mathbf{e}(A)} - 2\epsilon_\eta}{x_{\mathbf{e}} + 3\epsilon_{1/1}} \geq \frac{\epsilon_{1/2} - 2\epsilon_\eta}{1/2 + 1.3\epsilon_{1/2}} \geq 1.99\epsilon_{1/2}$$

as desired.

**Second we show $\mathbb{P}_\nu[X_T = Y_T = 1] \geq 0.0498\epsilon_{1/2}$.** If this is true by the lemma's assumption, we are done, otherwise we have the bounds on the expectations:

$$\mathbb{E}_\nu[X_T], \mathbb{E}_\nu[Y_T] \in [1.489, 1.51], \qquad \mathbb{E}_\nu[X_T + Y_T] \in [2.999, 3.01]$$

First suppose that $\mathbb{P}_\nu[X_T + Y_T = 2] \geq 0.001 \geq \epsilon_{1/2}$. Applying Corollary A.11 (and using Theorem A.9 to bound the probability $X_T, Y_T$ are at least 1 and at most 1), we obtain $\mathbb{P}_\nu[X_T = 1 | X_T + Y_T = 2] \geq \frac{1}{2}(0.245)(0.77) \geq 0.094$ and therefore again our desired bound $\mathbb{P}_\nu[X_T = Y_T = 1] \geq 0.06\epsilon_{1/2}$ would hold. Therefore we may assume that $\mathbb{P}_\nu[X_T + Y_T = 2] \leq 0.001$. By log concavity and the fact that $\mathbb{E}_\nu[X_T + Y_T = 3]$, it must be that $\mathbb{P}_\nu[X_T + Y_T = 3] \geq 0.995$.

We use this information to improve the bounds on $\mathbb{P}_\nu[X_T \geq 1], \mathbb{P}_\nu[Y_T \geq 1]$ and $\mathbb{P}_\nu[X_T \leq 1], \mathbb{P}_\nu[Y_T \leq 1]$. First, condition on $X_T + Y_T = 3$, this occurs with probability at least 0.995. Call the resulting measure $\nu'$. By log concavity,

$$1.48 \leq \mathbb{E}_{\nu'}[X_T], \mathbb{E}_{\nu'}[Y_T] \leq 1.52$$

Finally we apply Theorem A.9 and use the fact that there are at most 3 Bernoullis. This demonstrates that

$$\mathbb{P}_\nu[X_T \geq 1], \mathbb{P}_\nu[Y_T \geq 1] \geq 0.864$$
$$\mathbb{P}_\nu[X_T \leq 1], \mathbb{P}_\nu[Y_T \leq 1] \geq 0.419$$

Where we use that in $\nu'$ the quantities are bounded by 0.869 and 0.422 respectively. We now turn to the fact that $\mathbf{e}$ is 2-2 good to bound the probability the sum is 2. By Lemma 5.15 from [KKO21] and stochastic dominance,

$$\mathbb{P}_\nu[X_T + Y_T \leq 2] \geq \mathbb{P}[(\delta(u)_{-\mathbf{e}})_T + Y_T \leq 2] - \mathbb{P}[C_T = 0] \geq 0.4\epsilon_{1/2} - 2\epsilon_{1/1} - \epsilon_\eta \geq 0.22\epsilon_{1/2},$$

where we used $\epsilon_{1/1} < 12\epsilon_{1/2}$. It follows by log-concavity of $X_T + Y_T$ that $\mathbb{P}_\nu[X_T + Y_T = 2] \geq 0.2\epsilon_{1/2}$. Finally, we obtain $\mathbb{P}_\nu[X_T = 1 \mid X_T + Y_T = 2] \geq 0.249$. This is because of the following. WLOG, suppose $\mathbb{E}_\nu[X_T \mid X_T + Y_T = 2] \leq 1$ (otherwise we bound $\mathbb{P}_\nu[Y_T = 1 \mid X_T + Y_T = 2]$). Now by Lemma A.10, we have that $\mathbb{P}_\nu[X_T \geq 1 \mid X_T + Y_T = 2] \geq 0.362$. If $\mathbb{P}_\nu[X_T = 1 \mid X_T + Y_T = 2] \geq 0.25$ we are done already, therefore $\mathbb{P}_\nu[X_T = 2 \mid X_T + Y_T = 2] \geq 0.112$. This implies that $\mathbb{E}_\nu[X_T \mid X_T + Y_T = 1] \geq 0.474$. Now applying Theorem A.9 with at most two Bernoullis, we obtain the bound.

Therefore, $\mathbb{P}_\nu[X_T = Y_T = 1] \geq (0.2\epsilon_{1/2})(0.249) = 0.0498\epsilon_{1/2}$ as desired. □

## A.4 A Parameterized TSP Bound

The following statement is an immediate consequence of [KKO21; KKO22]:

**Lemma 4.5.** *Let $p$ be a lower bound on the probabilities guaranteed by (1) - (6) for $\epsilon_{1/2} \leq 0.0002$, and suppose $p \leq 10^{-4}$. Then given $x \in P_{\text{Sub}}$, the max entropy algorithm returns a solution of expected cost at most $(\frac{3}{2} - 9.7p^2 \cdot 10^{-17}) \cdot c(x)$.*

To see this, one can examine the proof of the "main payment theorem," Theorem 4.33, of [KKO21]. As it is a lengthy statement with many definitions not given here, we refer the reader to [KKO21]. In the proof of this theorem, there is a constant $\epsilon_P$ is set to be $1.56 \cdot 10^{-6}p$, where $p$ is the minimum probability guaranteed by Corollary 5.9, Lemma 5.25, and Lemma 5.28. Then, Lemma 5.25 and Lemma 5.28 are simply aggregations of the bounds in Lemmas 5.21, 5.22, 5.23, 5.24, and 5.27, which are the lemmas we improve in this paper. It is also required that $p \leq 3\epsilon_{1/2}$ and $\epsilon_{1/2} \leq 0.0002$, as this the threshold for the definition of a good edge from Definition 5.13, which is related to Lemma 5.16 and Lemma 5.17. Thus it is sufficient that $p \leq 10^{-4}$ if one does not modify these lemmas.

Therefore, given the assumptions of the lemma, Theorem 4.33 holds without any modification of the proof for $\epsilon_P = 1.56 \cdot 10^{-6}p$. We can then plug the improved version of the payment theorem into Theorem 6.1 from [KKO22], where one sets $\eta = \min\{10^{-12}, \frac{\epsilon_P}{750}\}$, $\beta = \frac{\eta}{4+2\eta}$ and achieves an approximation ratio (and integrality gap) of

$$\frac{3}{2} - \frac{\epsilon_P}{6}\beta + \frac{\epsilon_P \eta}{100} \geq \frac{3}{2} - \frac{\epsilon_P^2}{25000} \geq \frac{3}{2} - 9.7p^2 \cdot 10^{-17}$$

As long as $\epsilon_P/750 < 10^{-12}$ which it is for $p \leq 10^{-4}$.



ECCC ISSN 1433-8092
https://eccc.weizmann.ac.il