

# Zeros of Holant problems: locations and algorithms \*

Heng Guo <sup>†</sup>      Chao Liao <sup>‡</sup>      Pinyan Lu <sup>§</sup>      Chihao Zhang <sup>¶</sup>

## Abstract

We present fully polynomial-time (deterministic or randomised) approximation schemes for Holant problems, defined by a non-negative constraint function satisfying a generalised second order recurrence modulo a couple of exceptional cases. As a consequence, any non-negative Holant problem on cubic graphs has an efficient approximation algorithm unless the problem is equivalent to approximately counting perfect matchings, a central open problem in the area. This is in sharp contrast to the computational phase transition shown by 2-state spin systems on cubic graphs. Our main technique is the recently established connection between zeros of graph polynomials and approximate counting. We also use the “winding” technique to deduce the second result on cubic graphs.

## 1 Introduction

Great progress has been made recently in the classification of counting problems. One major achievement is the full dichotomy for counting constraint satisfaction problems (CSPs) [Bul13, DR13], even with complex weights [CC17]. However, such a classification is for exact counting, and for approximation, even to move beyond some rather modest model seems quite difficult.

Holant problems [CLX11] are a framework of expressing counting problems motivated by Valiant’s holographic algorithms [Val08]. The “Holant” is a partition function on graphs where edges are variables and vertices are constraint functions. The benefit of this choice is the ability to express problems like perfect matchings, which are provably not expressible in certain CSP-like

vertex models [FLS07, DGL<sup>+</sup>12, Sch13]. We parameterise Holant problems by the set of constraint functions that can be put on vertices. Similar to the success of classifying counting CSPs, exact classifications have been obtained for Holant problems defined by any set of complex-weighted symmetric Boolean functions [CGW16], and various progresses have been made to go beyond [CLX18, LW18, Bac18].

In this paper, we make progress towards understanding the complexity of approximating symmetric Boolean Holant problems with non-negative weights. Let  $G = (V, E)$  be a graph,  $\pi : V \rightarrow \mathcal{F}$  be an assignment from the set of vertices  $V$  to a set of functions  $\mathcal{F}$ , and  $f_v = \pi(v)$  is the constraint function  $\{0, 1\}^{\deg(v)} \rightarrow \mathbb{C}$  associated with the vertex  $v$ . The “Holant” is defined as follows:

$$(1.1) \quad Z(G; \pi) := \sum_{\sigma \in \{0, 1\}^E} \prod_{v \in V} f_v(\sigma|_{E(v)}),$$

where  $E(v)$  is the set of adjacent edges of  $v$ , and  $\sigma|_{E(v)}$  is the restriction of  $\sigma$  on  $E(v)$ . We use the shorthand  $Z(G)$  or  $Z$  when  $G$  and  $\pi$  are clear from the context.

We call a Boolean constraint function  $f$  *symmetric*, if  $f(\mathbf{x})$  depends only on the hamming weight  $|\mathbf{x}|$  and is invariant under permutations of the indices. For a symmetric  $f$  of arity  $d$ , we associate it with a *signature*  $[f_0, f_1, \dots, f_d]$ , where  $f_i = f(\mathbf{x})$  if  $|\mathbf{x}| = i$ . We may use the term “constraint function” and “signature” interchangeably. For example, if  $f$  is the “exact-one” function, namely  $f = [0, 1, 0, \dots, 0]$ , then  $Z(G)$  counts the number of perfect matchings in  $G$ ; and if  $f$  is the Boolean OR function, namely  $f = [0, 1, 1, \dots, 1]$ , then  $Z(G)$  counts the number of edge covers in  $G$ .

We focus on a fairly expressive family of symmetric functions satisfying generalised second-order recurrences. More precisely, we say  $f = [f_0, f_1, \dots, f_d]$  satisfies a generalised second-order recurrence, if there exist real constants  $(a, b, c) \neq (0, 0, 0)$  such that  $af_k + bf_{k+1} + cf_{k+2} = 0$  for all  $0 \leq k \leq d-2$ . Denote by  $\text{Holant}(f)$  the computational problem of evaluating  $Z(G)$  where every vertex is associated with the signature  $f$ . In particular, the input to  $\text{Holant}(f)$  must be a  $d$ -regular graph, where  $d$  is the arity of  $f$ . Our main theorem is the following.

**THEOREM 1.1.** *Let  $f = [f_0, f_1, \dots, f_d]$  be a symmetric constraint function of arity  $d \geq 3$  satisfying generalised*

\*The third author was supported by National NSF of China (No.61741209) and the Fundamental Research Funds for the Central Universities.

<sup>†</sup>School of Informatics, University of Edinburgh, Informatics Forum, Edinburgh, EH8 9AB, United Kingdom. [hguo@inf.ed.ac.uk](mailto:hguo@inf.ed.ac.uk)

<sup>‡</sup>Department of Computer Science and Engineering, Shanghai Jiao Tong University, No.800 Dongchuan Road, Minhang District, Shanghai, China. [chao.liao.95@gmail.com](mailto:chao.liao.95@gmail.com)

<sup>§</sup>ITCS, Shanghai University of Finance and Economics, No.100 Wudong Road, Yangpu District, Shanghai, China. [lu.pinyan@mail.shufe.edu.cn](mailto:lu.pinyan@mail.shufe.edu.cn)

<sup>¶</sup>John Hopcroft Center for Computer Science, Shanghai Jiao Tong University, No.800 Dongchuan Road, Minhang District, Shanghai, China. [chihao@cs.sjtu.edu.cn](mailto:chihao@cs.sjtu.edu.cn)

second-order recurrences, and  $f_i \geq 0$  for all  $0 \leq i \leq d$ . There is a fully polynomial-time (deterministic or randomised) approximation algorithm for  $\text{Holant}(f)$ , unless, up to a non-zero factor,  $f$  or its reversal is in one of the following form:

- $[0, \lambda \sin \frac{\pi}{d}, \lambda^2 \sin \frac{2\pi}{d}, \dots, \lambda^i \sin \frac{i\pi}{d}, \dots, 0]$  for some  $\lambda > 0$ ;
- $[0, 1, 0, \lambda, 0, \dots, 0, \lambda^{\frac{d-2}{2}}, 0]$  if  $d$  is even, or  $[0, 1, 0, \lambda, 0, \dots, 0, \lambda^{\frac{d-1}{2}}]$  if  $d$  is odd, for some  $0 \leq \lambda < 1$ .

Moreover, in the latter case, approximating  $\text{Holant}(f)$  is equivalent to approximately counting perfect matchings in general graphs.

Understanding the complexity of signatures with second-order recurrences is the cornerstone in the exact counting classifications. Since satisfying first-order recurrences implies that the function is degenerate, these constraint functions are the first class satisfying a recurrence relation with non-trivial complexity. More concretely, this family includes many interesting special cases:

- Matchings and perfect matchings, whose constraint functions are  $[1, 1, 0, 0, \dots, 0]$  and  $[0, 1, 0, 0, \dots, 0]$ , respectively, with  $(a, b, c) = (0, 0, 1)$ .
- Even subgraphs, whose constraint functions are  $[1, 0, 1, 0, \dots]$  with  $(a, b, c) = (1, 0, -1)$ . More generally, we may put weights on even and odd degree vertices, and the constraint functions become  $[x, y, x, y, \dots]$  for some  $x, y \geq 0$ .
- Edge covers, whose constraint functions are  $[0, 1, 1, \dots, 1]$  with  $(a, b, c) = (0, 1, -1)$ .
- Fibonacci gates, namely  $f$  of arity  $d$  such that  $f_{i+2} = bf_{i+1} + f_i$  for all  $0 \leq i \leq d - 2$ .
- All ternary symmetric functions.

For approximate counting, polynomial-time approximation algorithms are known only for a few special cases, such as counting matchings [JS89], weighted even subgraphs [JS93], counting edge covers [LLL14], and a weighted version of Fibonacci gates [LWZ14]. However, neither the Markov chain Monte Carlo approach [JS89, JS93] (including its “winding” extension [McQ13, HLZ16]), nor the correlation decay approach [LWZ14, LLL14], appears to be powerful enough to handle all functions in this family. On the other hand, Theorem 1.1 covers almost all problems in this family, and most of the exceptional cases are equivalent to counting

perfect matchings, a central open problem in approximate counting (see, for example, [DJM17, ŠVW18] on partial progress and barriers). Efficient approximate counting algorithm for perfect matchings is only known in the bipartite case [JSV04].

As a consequence, we have an algorithm for all non-negative Boolean Holant on cubic graphs, unless the problem is equivalent to counting perfect matchings.

**THEOREM 1.2.** *Let  $f = [f_0, f_1, f_2, f_3]$  be a symmetric constraint function of arity 3 where  $f_i \geq 0$  for all  $0 \leq i \leq 3$ .  $\text{Holant}(f)$  has a fully polynomial-time (deterministic or randomised) approximation algorithm, unless  $f$  or its reversal, up to a non-zero factor, is  $[0, 1, 0, \lambda]$  for some  $0 \leq \lambda < 1$ . In the exceptional case, approximating  $\text{Holant}(f)$  is equivalent to approximately counting perfect matchings in general graphs.*

We remark that Theorem 1.2 is in sharp contrast to the computational phase transition phenomenon, as demonstrated by 2-state spin systems on cubic graphs [GJP03, SS14, GŠV16, LLY13, SST14], even without external fields. For spin systems, a clear and sharp threshold between approximable and hard to approximate is established, whereas for Holant problems on cubic graphs, there seems to be no such transition.

**1.1 Our techniques** Our algorithm combines a number of ingredients:

- Barvinok’s approach to approximate partition functions via Taylor expansions [Bar16]. This approach was sharpened by Patel and Regts [PR17a] to run within polynomial-time.
- In order to apply Barvinok’s approach, one has to have some rather precise knowledge of the zeros of the corresponding graph polynomials. For Holant problems, Ruelle [Rue71, Rue99a, Rue99b] has developed a systematic approach of bounding the zeros of the partition function via analysing polynomials associated locally with vertices, under the disguise of “graph-counting polynomials”.
- On top of combining Ruelle’s and Barvinok’s approaches, we also employ holographic transformations a la Valiant [Val08], which is necessary to cover all cases in Theorem 1.1.

Although none of these ingredients is new, the main contribution of our work is to combine them together (with reworks if necessary), and a thorough analysis of the zeros of functions with generalised second-order recurrence. To be more specific, for a symmetric

signature  $f = [f_0, \dots, f_d]$  of arity  $d$ , define the “local” polynomial of  $f$  as

$$(1.2) \quad P_f(z) := \sum_{i=0}^d \binom{d}{i} f_i \cdot z^i.$$

We may also view  $P_f(z)$  as the polynomial for a single vertex with  $d$  dangling edges. For some  $\varepsilon > 0$ , we call a polynomial  $P(z)$   $H_\varepsilon$ -stable, if  $P(z) \neq 0$  as long as  $\Re z > -\varepsilon$ . Then one of our main technical tool (see Theorem 3.1) says that if  $P_f(z)$  is  $H_\varepsilon$ -stable for some  $\varepsilon > 0$ , then a polynomial-time approximation algorithm exists for  $\text{Holant}(f)$ .

In general, to apply Barvinok’s method to approximate counting, one needs to deal with the zeros of the whole partition function, which is usually not an easy task. Previous applications appeal to some powerful tools such as the Lee-Yang theorem from statistical physics [LSS17], or the resolution of a long-standing conjecture [PR17b]. In contrast, our approach requires only analysing some low degree polynomials and is much easier to apply.

To go from Theorem 1.1 to Theorem 1.2, we also need to deal with cases not covered by Theorem 1.1, which cannot be solved using zeros of Holant problems. These exceptional cases are handled by the “winding” technique [McQ13, HLZ16] with Markov chains.

## 2 Ruelle’s method on zeros of Holant problems

Ruelle [Rue71, Rue99a, Rue99b] (building upon the “Asano contraction” [Asa70]) has developed a systematic approach to bound zeros of the so-called “graph-counting polynomials”. As we will see later, these polynomials coincide with unweighted Holant problems.

With a little abuse of notation, let  $Z(G; f)$  be the partition function defined by (1.1) where  $f_v = f$  for all  $v \in V$ , and stratify  $Z(G; f)$  by the number of edges chosen as follows:

$$(2.3) \quad Z_k(G; f) := \sum_{\sigma \in \{0,1\}^E \text{ and } |\sigma|=k} \prod_{v \in V} f(\sigma|_{E(v)}).$$

Define  $Z_k(G; \pi)$  similarly, and again,  $G$  and  $f$  may be omitted when they are clear from the context.

Let  $|E| = m$ . Then  $Z = Z(G; f)$  can be rewritten as the evaluation of the polynomial

$$(2.4) \quad P_G(z) := \sum_{i=0}^m Z_i \cdot z^i$$

at  $z = 1$ . Namely  $Z = P_G(1)$ . When  $f$  is a symmetric 0/1 function, then (2.4) is the same as the “graph-counting” polynomial defined by Ruelle [Rue99b].

Ruelle’s method has two main ingredients. Firstly we want to relate zeros of a univariate polynomial with those of its polar form. For a polynomial  $P(z) = \sum_{i=0}^{d'} a_i z^i$  of degree  $d' \leq d$ , its  $d$ th polar form with variables  $\mathbf{z} = (z_1, \dots, z_d)$  is

$$\widehat{P}(\mathbf{z}) := \sum_{I \subseteq [d]} \frac{a_{|I|}}{\binom{d}{|I|}} z_I,$$

where  $a_i = 0$  if  $i > d'$ ,  $[d]$  denotes  $\{1, 2, \dots, d\}$ , and for an index set  $I$ ,  $z_I = \prod_{i \in I} z_i$ . For example, the polar form of  $P_f(z)$  (recall (1.2)) is,

$$\widehat{P}_f(\mathbf{z}) := \sum_{I \subseteq [d]} f_{|I|} z_I.$$

The polar form  $\widehat{P}(\mathbf{z})$  is the unique multi-linear symmetric polynomial of degree at most  $d'$  such that  $\widehat{P}(z, z, \dots, z) = P(z)$ . When  $d' < d$ , we view  $P(z)$  as a degenerate case, and it has zeros at  $\infty$  with multiplicity  $d - d'$ .

Let  $H$  be a set in  $\mathbb{C}$  and we use  $\overline{H} = \{z \in \mathbb{C} \mid z \notin H\}$  to denote its complement. We say a polynomial  $P(\mathbf{z})$  in  $d \geq 1$  variables is  $H$ -stable if  $P(\mathbf{z}) \neq 0$  whenever  $z_1, \dots, z_d \in H$ . We will be particularly interested in  $H_\varepsilon$ -stability where  $H_\varepsilon$  is the half-plane:

$$H_\varepsilon = \{z \in \mathbb{C} \mid \Re z > -\varepsilon\},$$

and  $\varepsilon > 0$ . The Grace-Szegő-Walsh coincidence theorem [Gra02, Sze22, Wal22] has the following immediate consequence.

**PROPOSITION 2.1.** *A univariate polynomial  $P(z)$  is  $H_\varepsilon$ -stable if and only if its polar form  $\widehat{P}(\mathbf{z})$  is  $H_\varepsilon$ -stable.*

Proposition 2.1 actually applies to an arbitrary circular domain in  $\mathbb{C}$ , but we will only need it for  $H_\varepsilon$ .

The next ingredient is the Asano contraction [Asa70], as extended by Ruelle [Rue71].

**PROPOSITION 2.2.** *Let  $K_1$  and  $K_2$  be closed subsets of the complex plane  $\mathbb{C}$ , which do not contain 0. If the complex polynomial*

$$\alpha + \beta z_1 + \gamma z_2 + \delta z_1 z_2$$

*can vanish only when  $z_1 \in K_1$  or  $z_2 \in K_2$ , then*

$$\alpha + \delta z$$

*can vanish only when  $z \in -K_1 \cdot K_2$ .*

We refer interested readers to [Rue71] for a very elegant proof of Proposition 2.2.

Let the  $\delta$ -strip of  $[0, 1]$  be

$$\{z \in \mathbb{C} \mid |\Im z| \leq \delta \text{ and } -\delta \leq \Re z \leq 1 + \delta\}.$$

LEMMA 2.1. For any  $\varepsilon > 0$ , the complement of  $-\overline{H_\varepsilon} \cdot \overline{H_\varepsilon}$  contains a  $\delta$ -strip of  $[0, 1]$  for some  $\delta > 0$  depending only on  $\varepsilon$ .

*Proof.* An equivalent way to write  $\overline{H_\varepsilon}$  is

$$\overline{H_\varepsilon} = \left\{ \rho e^{i\theta} \mid \rho \geq -\frac{\varepsilon}{\cos \theta} \text{ for } \theta \in \left( \frac{\pi}{2}, \frac{3\pi}{2} \right) \right\}.$$

Thus,

$$\begin{aligned} -\overline{H_\varepsilon} \cdot \overline{H_\varepsilon} &= \left\{ \rho_1 \rho_2 e^{i(\theta_1 + \theta_2 + \pi)} \mid \rho_i \geq -\frac{\varepsilon}{\cos \theta_i} \right. \\ &\quad \left. \text{for } \theta_i \in \left( \frac{\pi}{2}, \frac{3\pi}{2} \right) \text{ and } i \in \{1, 2\} \right\} \\ &= \left\{ \rho e^{i(\theta_1 + \theta_2 + \pi)} \mid \rho \geq \frac{\varepsilon^2}{\cos \theta_1 \cos \theta_2} \right. \\ &\quad \left. \text{for } \theta_1, \theta_2 \in \left( \frac{\pi}{2}, \frac{3\pi}{2} \right) \right\} \\ &= \left\{ \rho e^{i\theta} \mid \rho \geq \frac{\varepsilon^2}{(\cos \frac{\theta - \pi}{2})^2} \text{ for } \theta \in (0, 2\pi) \right\} \\ &= \left\{ \rho e^{i\theta} \mid \rho \geq \frac{2\varepsilon^2}{1 - \cos \theta} \text{ for } \theta \in (0, 2\pi) \right\}, \end{aligned}$$

where the third line is because  $\cos \theta_1 \cos \theta_2$  is maximised at  $\theta_1 = \theta_2$  if their sum is fixed. Thus the complement of  $-\overline{H_\varepsilon} \cdot \overline{H_\varepsilon}$  is

$$-\overline{H_\varepsilon} \cdot \overline{H_\varepsilon} = \left\{ \rho e^{i\theta} \mid \rho < \frac{2\varepsilon^2}{1 - \cos \theta} \text{ for } \theta \in (0, 2\pi) \right\}.$$

It is easy to check that  $\delta = \varepsilon^2/2$  suffices for the claim.

Now we are ready to state a very useful lemma.

LEMMA 2.2. Let  $f$  be a symmetric signature of arity  $\Delta$ . If the local polynomial  $P_f(z)$  is  $H_\varepsilon$ -stable for some  $\varepsilon > 0$ , then the global polynomial  $P_G(z)$  has no zero in the  $\delta$ -strip of  $[0, 1]$ , where  $\delta$  is a constant depending only on  $\varepsilon$ .

*Proof.* We construct  $G = (V, E)$  as follows. Start with a collection of vertices  $v \in V$ , each with  $\Delta$  dangling half-edges  $(e_i^v)_{i \in [\Delta]}$ . Call this graph  $G_0$ , and connect dangling half-edges  $e_i^v$  and  $e_j^u$  sequentially for each edge  $(u, v) \in E$ . This gives a sequence of graphs  $G_1, \dots, G_{|E|} = G$ . The polynomial of  $G_0$  is  $P_{G_0}(z) = \prod_{v \in V} P_v(z)$ , where  $P_v = P_f$ , and consider the multivariate version  $\widehat{P}_{G_0}(\mathbf{z}) = \prod_{v \in V} \widehat{P}_v(\mathbf{z}^v)$ , where  $\widehat{P}_v = \widehat{P}_f$  and  $\mathbf{z}^v$  is the local variables corresponding to  $v$ . Since  $P_f(z)$  is  $H_\varepsilon$ -stable, by Proposition 2.1,  $\widehat{P}_f(\mathbf{z})$  is as well, and so is  $\widehat{P}_{G_0}(\mathbf{z})$ . Suppose from  $G_i$  to  $G_{i+1}$ ,  $e_i^v$  is connected with  $e_j^u$ . Then the transformation from

$\widehat{P}_{G_i}$  to  $\widehat{P}_{G_{i+1}}$  is exactly the Asano contraction as in Proposition 2.2 applied to  $z_i^v$  and  $z_j^u$ . At the end of this procedure we obtain  $G$  and the polynomial  $\widehat{P}_G(\mathbf{z})$  does not vanish on the complement of  $-\overline{H_\varepsilon} \cdot \overline{H_\varepsilon}$ . It implies that the same is true for the univariate  $P_G(z)$ . By Lemma 2.1, the complement of  $-\overline{H_\varepsilon} \cdot \overline{H_\varepsilon}$  contains a  $\delta$ -strip of  $[0, 1]$ , and this  $\delta$  depends only on  $\varepsilon$ .

We note that it is necessary to have some slack  $\varepsilon$  in Lemma 2.2. One example is counting even subgraphs, namely the constraint  $f$  is  $[1, 0, 1, 0, \dots]$ . Although all zeros of  $P_f$  lie on the imaginary axis, the zeros of  $P_G(z)$  can in fact be dense on the unit circle. To see this, let  $G$  be a cycle of length  $n$ . Then  $P_G(z) = 1 + z^n$  as there are only two even subgraphs. The zeros thereof are dense on the unit circle as  $n$  varies.

Lemma 2.2 can be generalised to a set of functions by following the proof of Lemma 2.2, if there is an  $\varepsilon > 0$  such that all of the local polynomials are  $H_\varepsilon$ -stable. A univariate polynomial is called *Hurwitz stable* if all of its zeros are in the open left half-plane. For a fixed  $f$ , clearly if  $P_f(z)$  is Hurwitz stable, then there is some  $\varepsilon > 0$  such that  $P_f(z)$  is  $H_\varepsilon$ -stable. However, Hurwitz stability is not enough to derive the same conclusion of Lemma 2.2 for an infinite set of functions.

### 3 Barvinok's algorithm

Our interest in Ruelle's method is due to the algorithmic approach developed by Barvinok [Bar16, Section 2]. It roughly states that if a polynomial  $P(z) = \sum_{i=1}^n c_i z^i$  of degree  $n$  is zero-free in a strip containing  $[0, 1]$ , then  $P(1)$  can be  $(1 \pm \varepsilon)$ -approximated using  $c_0, \dots, c_k$  for some  $k = O(\log \frac{n}{\varepsilon})$ .

The basic idea is to truncate the Taylor expansion of  $\log P(z)$  at  $z = 0$ . Let  $g(z) := \log P(z)$  and for  $k \geq 0$ ,

$$T_k(g)(z) := \sum_{i=0}^k \frac{g^{(i)}(0)}{i!} z^i,$$

where  $g^{(i)}$  is the  $i$ -th derivative of  $g$ . In other words,  $T_k(g)(z)$  is the first  $k+1$  terms of the Taylor expansion of  $g(z)$  at the origin. Then [Bar16, Lemma 2.2.1] states the following.

PROPOSITION 3.1. Let  $P(z) = \sum_{i=0}^n c_i z^i$  be a polynomial such that for some  $\beta > 1$ ,  $P(z)$  is zero-free in the (closed) disk of radius  $\beta$  centered at the origin. Then there exists a constant  $C_\beta$  such that for any  $0 < \varepsilon < 1$ ,

$$\left| \frac{\exp(T_k(g)(1))}{P(1)} - 1 \right| \leq \varepsilon,$$

where  $k = C_\beta \log \frac{n}{\varepsilon}$ .

This result states that we can approximately evaluate  $P(1)$  using the first  $O(\log \frac{n}{\varepsilon})$  terms of the Taylor expansion of  $\log P(x)$  at the origin, when the polynomial is zero-free in the disk of radius  $\beta > 1$ . If our polynomial  $P_G(x)$  is zero-free in the  $\delta$ -strip of  $[0, 1]$ , then we can apply a transformation, [Bar16, Lemma 2.2.3], to transform it into a polynomial that is zero-free in the disk of radius  $> 1$ .

The following lemma describe the construction.

**LEMMA 3.1.** *Let  $0 < \delta < 1$  be a constant and  $\beta = 1 + \frac{\exp(-\frac{1}{\delta})}{2 - 2 \exp(-\frac{1}{\delta})} > 1$ . There exists a polynomial  $\phi_\delta(z)$  of degree  $\exp(O(\frac{1}{\delta}))$  such that*

- (1)  $\phi_\delta(0) = 0$  and  $\phi_\delta(1) = 1$ ;
- (2) for every  $z \in \mathbb{C}$  with  $|z| \leq \beta$ , the value  $\phi_\delta(z)$  is within the  $2\delta$ -strip of  $[0, 1]$ .

*Proof.* The idea to construct the polynomial  $\phi_\delta$  is to start with the function  $\log(z)$  (the principal logarithm) by noting that the logarithm function maps a circle centered at zero to an interval orthogonal to the real axis. We can then scale and shift the function to restrict the interval to some desired region. Finally, we construct the polynomial  $\phi_\delta$  to approximate it.

To this end, we let  $h(z) := \delta \log \frac{1}{1-\alpha z}$  where  $\alpha$  is a parameter to be set. The condition  $h(0) = 0$  is automatically satisfied. To satisfy  $h(1) = 1$ , we set  $\alpha = 1 - \exp(-\frac{1}{\delta})$ . Then  $\beta = 1 + \frac{\exp(-\frac{1}{\delta})}{2 - 2 \exp(-\frac{1}{\delta})} = \frac{1+\alpha}{2\alpha}$ . Note that  $\beta < \frac{1}{\alpha}$ , so  $h(z)$  is well-defined over the disk of radius  $\beta$  centered at the origin. It is easy to verify that for every  $z \in \mathbb{C}$  with  $|z| \leq \beta$ , it holds that

$$-\delta \log 2 \leq \Re h(z) \leq 1 + \delta \log 2,$$

and

$$|\Im h(z)| \leq \frac{\pi}{2} \cdot \delta.$$

We use a polynomial, namely the Taylor expansion of  $h(z)$  at the origin to approximate  $h(z)$ . For every  $k \geq 0$ , the first  $k$  terms of the Taylor expansion of  $h$  at the origin is

$$T_k(h)(z) = \delta \sum_{i=1}^k \frac{\alpha^i}{i} \cdot z^i.$$

Then for  $m = \frac{\log(10(1+\alpha)) - \log(1-\alpha)}{\log 2 - \log(1+\alpha)} = \exp(O(\frac{1}{\delta}))$ , we have

$$\begin{aligned} |h(z) - T_k(z)| &= \left| \delta \sum_{i=m+1}^{\infty} \frac{\alpha^i}{i} \cdot z^i \right| \\ &\leq \frac{2\delta}{(1-\alpha)(m+1)} \left( \frac{1+\alpha}{2} \right)^{m+1} \\ &\leq \frac{\delta}{10}. \end{aligned}$$

In particular, we have

$$|T_m(h)(1) - 1| = |T_m(h)(1) - h(1)| \leq \frac{\delta}{10}.$$

Finally, we define

$$\phi_\delta(z) = \frac{T_m(h)(z)}{T_m(h)(1)}$$

to force  $\phi_\delta(1) = 1$ . This finishes the construction.

Therefore, for a polynomial  $P(z)$  that is zero-free in the  $\delta$ -strip of  $[0, 1]$ , we can use Proposition 3.1 to approximately evaluate  $P_\phi(z) := P(\phi_{\frac{\delta}{2}}(z))$ , which is zero-free in the disk of radius  $\beta$  at the origin for the value  $\beta$  defined in Lemma 3.1. Note that  $P(\phi_{\frac{\delta}{2}}(1)) = P(1)$ .

**PROPOSITION 3.2.** *Let  $P(z)$  be a polynomial of degree  $n$  such that for some  $\delta > 0$ ,  $P(z)$  is zero-free in the  $\delta$ -strip of  $[0, 1]$ . Then there exists a constant  $C_\delta$  such that for any  $0 < \varepsilon < 1$ ,*

$$\left| \frac{\exp(T_k(\log P_\phi)(1))}{P(1)} - 1 \right| \leq \varepsilon,$$

where  $k = C_\delta \log \frac{n}{\varepsilon}$ .

At last, we show the Taylor expansion  $T_k(\log P_\phi)(1)$  can be computed efficiently from the coefficients of  $P$ .

**PROPOSITION 3.3.** *Let  $P(z)$  be a polynomial of degree  $n$  such that for some constant  $\delta > 0$ ,  $P(z)$  is zero-free in the  $\delta$ -strip of  $[0, 1]$ . For every  $0 \leq k \leq n$ , assume that that we have oracle access to the first  $k$  coefficients of  $P(z)$ , we can compute*

$$T_k(\log P_\phi)(1)$$

in time  $O(k^2)$ .

Since the degree of  $\phi_{\frac{\delta}{2}}(z)$  is  $\exp(O(\frac{1}{\delta}))$ , we can write  $P_\phi(z) = \sum_{i=1}^m c_i z^i$  where  $m = n + C_\delta$  for some constant  $C_\delta$  depending only on  $\delta$ . It is easy to compute the coefficients  $c_k$  given the coefficients of  $P(z)$  of degree at most  $k$  in  $O(k)$  time. Let  $g_\phi := \log P_\phi$ , we now show how to compute  $T_k(g_\phi)$  using  $(c_i)_{i \leq k}$ .

Let  $z_1, \dots, z_m$  be the zeros of a polynomial  $P_\phi(z)$  and for  $0 \leq k \leq m$ , let  $p_k := \sum_{i=1}^m z_i^{-k}$  be the  $k$ -th inverse power sum of the zeros of  $P_\phi(z)$ .

Newton's identities state the relation between  $(p_k)_k$  and the coefficients  $(c_i)_i$ .

**PROPOSITION 3.4. (NEWTON'S IDENTITY)** *For every  $1 \leq k \leq m$ , it holds that*

$$k \cdot c_k = - \sum_{i=0}^{k-1} c_i \cdot p_{k-i}$$

Newton's identities essentially provide a way to compute all  $p_k$  consecutively using  $(c_i)_i$ , and vice versa. To be specific,

$$p_0 = m;$$

$$p_k = -c_0^{-1} \cdot \left( \sum_{i=1}^{k-1} p_i \cdot c_{k-i} + k \cdot c_k \right) \text{ for } 1 \leq k \leq m.$$

Therefore, it costs  $O(k^2)$  time to compute  $p_k$  using above recurrence.

On the other hand, we can write  $P_\phi(z) = c_m \prod_{i=1}^m (z - z_i)$ . Recall that  $g_\phi(z) = \log P_\phi(z) = \log c_m + \sum_{i=1}^m \log(z - z_i)$ .

It is easy to calculate that for any  $i \geq 1$ ,

$$g_\phi^{(i)}(0) = -(i-1)! \sum_{j=1}^m z_j^{-i} = -(i-1)! p_i.$$

Therefore,

$$(3.5) \quad T_k(g_\phi)(z) := \log c_0 - \sum_{i=1}^k \frac{p_i}{i} z^i.$$

This proves Proposition 3.3.

**3.1 Computing the inverse power sums** Given Proposition 3.1 and (3.5), the main task then reduces to compute the first  $k$  inverse power sums  $(p_i)_{i \leq k}$ . We follow the method of Patel and Regts [PR17a].

We need some notations first. Let  $\mathcal{G}$  be a family of all graphs, and  $\mathcal{G}_k$  be all graphs with at most  $k$  vertices. We call a function  $g : \mathcal{G} \rightarrow \mathbb{C}$  a *graph invariant* if  $g(G) = g(H)$  whenever  $G \simeq H$ . A graph *polynomial* is a graph invariant  $Q : \mathcal{G} \rightarrow \mathbb{C}[z]$ , where  $\mathbb{C}[z]$  is the polynomial ring over  $\mathbb{C}$ . We call a graph invariant  $g(\cdot)$  *additive* if for any two graphs  $G$  and  $H$ , it holds that  $g(G \sqcup H) = g(G) + g(H)$ , where  $G \sqcup H$  is the graph consisting of disjoint copies of  $G$  and  $H$ . Similarly, we call it *multiplicative* if for every two graphs  $G$  and  $H$ , it holds that  $g(G \sqcup H) = g(G) \cdot g(H)$ . For graphs  $H$  and  $G$ , we use  $\#\text{Ind}(H, G)$  to denote the number of *induced* subgraphs of  $G$  isomorphic to  $H$ . Then  $\#\text{Ind}(H, \cdot)$  is a graph invariant for a fixed graph  $H$ . By convention let  $\#\text{Ind}(\emptyset, G) = 1$  for any  $G$ .

**DEFINITION 3.1.** Let  $Q(G)(z) = \sum_{i=1}^{d(G)} a_i(G) z^i$  be a *multiplicative graph polynomial of degree  $d(G)$*  such that  $Q(G)(0) = 1$  for any  $G$ . We call  $Q(\cdot)$  a *bounded induced graph counting polynomial (BIGCP)* if there are constants  $\alpha, \beta \in \mathbb{N}$  such that the following holds:

- for every graph  $G$ , there exist  $\lambda_{H,i} \in \mathbb{C}$  such that

$$(3.6) \quad a_i(G) = \sum_{H \in \mathcal{G}_{\alpha i}} \lambda_{H,i} \cdot \#\text{Ind}(H, G);$$

- for every  $H \in \mathcal{G}_{\alpha i}$ ,  $\lambda_{H,i}$  can be computed in time  $\exp(\beta \cdot |V(H)|)$ , where  $V(H)$  is the set of vertices of  $H$ .

Patel and Regts [PR17a, Theorem 3.2] has shown that the inverse power sums can be computed for BIGCP in single exponential time.

**PROPOSITION 3.5.** Let  $\Delta \in \mathbb{N}$ ,  $G$  be a graph with maximum degree  $\Delta$  and  $Q(G)(\cdot)$  be a BIGCP. There is a deterministic  $\exp(C\Delta k)$ -time algorithm, which computes the inverse power sums  $(p_i)_{i \leq k}$  of  $Q(G)(\cdot)$ , for some constant  $C > 0$ .

To our need, we just need to verify that  $P_G(\cdot)$  from (2.4) is a BIGCP, whenever  $f_0 = 1$ .

**LEMMA 3.2.** Let  $G = (V, E)$  be a  $\Delta$ -regular graph and  $f = [f_0, f_1, \dots, f_\Delta]$  be a signature. If  $f_0 = 1$ , then the Holant polynomial  $P_G(\cdot)$  is a BIGCP with  $\alpha = 2$  and  $\beta = C\Delta$  for some constant  $C > 0$ .

*Proof.* Clearly  $P_G(0) = Z_0(G) = f_0^{|V|} = 1$ . We would like to define  $\lambda_{H,i}$  so that for every  $1 \leq i \leq n$ ,

$$(3.7) \quad Z_i(G) = \sum_{H \in \mathcal{G}_{2i}} \lambda_{H,i} \cdot \#\text{Ind}(H, G).$$

For any  $\sigma \in \{0, 1\}^E$ , let  $G[\sigma]$  be the subgraph induced by the set of vertices with at least 1 adjacent edges under  $\sigma$ . Let  $S_i$  be the set of subgraphs induced by assignments of Hamming weight  $i$ , namely  $S_i := \{G[\sigma] : \sigma \in \{0, 1\}^E \text{ and } |\sigma| = i\}$ . The equivalence relation of graph isomorphisms induces a partition of  $S_i$ . We choose one graph from each equivalence class and denote this family of graphs by  $\mathcal{H}_i$ . Therefore, for every two distinct graphs  $H_1, H_2 \in \mathcal{H}_i$ , they are not isomorphic. Moreover, as  $G[\sigma]$  has at most  $2i$  vertices,  $\mathcal{H}_i \subseteq \mathcal{G}_{2i}$ .

For every  $H \in \mathcal{H}_i$ , consider an assignment  $\pi$  of signatures, where  $v \in V$  of degree  $d \leq \Delta$  is assigned  $[f_0, f_1, \dots, f_d]$ , a truncated  $f$ . Let

$$\lambda_{H,i} := Z_i(H; \pi).$$

To verify (3.7), we rewrite

$$\begin{aligned}
 Z_i(G) &= \sum_{\sigma \in \{0,1\}^E \text{ and } |\sigma|=i} \prod_{v \in V} f(\sigma|_{E(v)}) \\
 &= \sum_{H \in \mathcal{G}_{2i}} \sum_{\substack{\sigma \in \{0,1\}^E \\ |\sigma|=i \text{ and } G[\sigma] \simeq H}} \prod_{v \in V} f(\sigma|_{E(v)}) \\
 &= \sum_{H \in \mathcal{G}_{2i}} \sum_{\substack{G' \text{ is an induced subgraph of } G \\ G' \simeq H}} \sum_{\substack{\sigma \in \{0,1\}^E \\ |\sigma|=i \text{ and } G[\sigma]=G'}} \prod_{v \in V} f(\sigma|_{E(v)}) \\
 &= \sum_{H \in \mathcal{G}_{2i}} \sum_{\substack{G' \text{ is an induced subgraph of } G \\ G' \simeq H}} Z_i(G'; \pi) \cdot f_0^{|V \setminus V(H)|} \\
 &= \sum_{H \in \mathcal{G}_{2i}} Z_i(H; \pi) \cdot \#\text{Ind}(H, G),
 \end{aligned}$$

since  $Z_i(G'; \pi) = Z_i(H; \pi)$  whenever  $G' \simeq H$ . Thus (3.7) holds.

Since  $\mathcal{H}_i \subseteq \mathcal{G}_{2i}$ , we have that  $\alpha = 2$ . Moreover,  $H$  contains at most  $\Delta |V(H)|$  edges. As a consequence,  $Z_i(H; \pi)$  can be computed in time  $2^{O(\Delta |H|)}$ . Thus, we can take  $\beta = C\Delta$  for some constant  $C > 0$ .

Gathering what we have seen so far, we have the following theorem.

**THEOREM 3.1.** *Let  $f$  be a symmetric signature of arity  $\Delta$ . If the local polynomial  $P_f(x)$  is  $H_\varepsilon$ -stable for some  $\varepsilon > 0$ , then there is an FPTAS for  $\text{Holant}(f)$ .*

*Proof.* Since  $P_f(x)$  is  $H_\varepsilon$ -stable,  $f_0 \neq 0$ . We may thus normalize  $f$  so that  $f_0 = 1$ . By Lemma 2.2,  $P_f(x)$  being  $H_\varepsilon$ -stable implies that for any  $\Delta$ -regular  $G = (V, E)$ ,  $P_G(x)$  is zero-free in a  $\delta$ -strip containing  $[0, 1]$ . Recall that  $Z(G; f) = P_G(1)$ . By Proposition 3.2, we can  $(1 \pm \varepsilon)$ -approximate  $P_G(1)$  using  $\exp(T_k(\log P_G)(x))$  for some  $k = O(\log \frac{m}{\varepsilon})$ , where  $m = |E|$ . In order to compute  $T_k(\log P_G)(x)$ , we use Proposition 3.5 and Lemma 3.2 to compute the inverse power sums  $(p_i)$  of  $P_G(x)$ , and then apply Proposition 3.4 to get the first  $k$  coefficients of  $P_G(x)$ . The theorem then follows from Proposition 3.3.

**REMARK 3.1.** *Theorem 3.1 is a sufficient but not necessary condition for a Holant problem to be approximable. To see this, once again, consider the problem of counting even subgraphs.*

## 4 Holographic transformations

Theorem 3.1 implies an FPTAS for  $\text{Holant}(f)$  if  $f$  is  $H_\varepsilon$ -stable. However, an FPTAS may still exist even if  $f$  is not  $H_\varepsilon$ -stable. One way to extend the reach of this approach is via Valiant's holographic transformation [Val08], which changes  $f$  but preserves the partition function. We remark that even with holographic transformations, this approach is not exhaustive. An example is the problem of counting even subgraphs.

We use  $\text{Holant}(f | g)$  to denote the Holant problem where the input is a bipartite graph  $H = (U, V, E)$ . Each vertex in  $U$  or  $V$  is assigned the signature  $f$  or  $g$ , respectively. Call this assignment  $\pi$ , namely  $\pi(u) = f$  for any  $u \in U$  and  $\pi(v) = g$  for any  $v \in V$ . Recall (1.1), and  $Z(H; \pi)$  is the output of the computational problem  $\text{Holant}(f | g)$ . The signature  $f$  is considered as a row vector (or covariant tensor), whereas the signature  $g$  is considered as a column vector (or contravariant tensor).

Let  $T$  be an invertible 2-by-2 matrix. Let  $d_1 = \text{arity}(f)$  and  $d_2 = \text{arity}(g)$ . Define  $f' = f \cdot T^{\otimes d_1}$  and  $g' = (T^{-1})^{\otimes d_2} g$ . Let  $\pi'$  be the assignment such that  $\pi'(u) = f'$  for any  $u \in U$  and  $\pi'(v) = g'$  for any  $v \in V$ .

**PROPOSITION 4.1.** ([VAL08]) *If  $T \in \mathbb{C}^{2 \times 2}$  is an invertible matrix, then for any bipartite graph  $H$ ,  $Z(H; \pi) = Z(H; \pi')$ , where  $\pi'$  is defined above.*

Therefore, an invertible holographic transformation does not change the complexity of the Holant problem in the bipartite setting. For a (non-bipartite) Holant problem, we can always view the edge as a binary equality function  $=_2$ . Thus,  $\text{Holant}(f)$  is the same as  $\text{Holant}(f | =_2)$ . Let  $\mathbf{O}_2(\mathbb{C})$  be the set of 2-by-2 orthogonal matrices, namely  $\mathbf{O}_2(\mathbb{C}) = \{T \in \mathbb{C}^{2 \times 2} \mid TT^T = I_2\}$ . As orthogonal transformations preserve the binary equality, the following result will become handy in the standard setting.

**PROPOSITION 4.2.** ([CLX11]) *If  $T \in \mathbf{O}_2(\mathbb{C})$  is an orthogonal matrix then for any  $d$ -regular graph  $G$  and a signature  $f$  of arity  $d$ ,  $Z(G; f) = Z(G; f \cdot T^{\otimes d})$ .*

As a particular consequence of Proposition 4.2, under the transformation  $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ , the complexity of  $\text{Holant}(f)$  is equivalent to  $\text{Holant}(\bar{f})$  where  $\bar{f} = [f_d, f_{d-1}, \dots, f_0]$ . We will use this fact in the following without explicitly mentioning it.

## 5 Second-order recurrences

The aim of this section is to study the locations of zeros of local polynomials of signatures satisfying generalised second-order recurrences in order to apply Theorem 3.1. Specifically, we identify the family of signatures whose

local polynomials are  $H_\varepsilon$ -stable for some  $\varepsilon > 0$ , under some suitable holographic transformations.

For a tuple of reals  $(a, b, c) \neq (0, 0, 0)$ , define

$$\mathcal{F}_{a,b,c} := \{[f_0, f_1, \dots, f_d] : af_k + bf_{k+1} + cf_{k+2} = 0, \forall 0 \leq k \leq d-2, \text{ and } f_k \geq 0, \forall 0 \leq k \leq d\}.$$

The family  $\mathcal{F}_{a,b,c}$  consists of signatures with non-negative entries satisfying second-order linear recurrence relation parameterized by  $(a, b, c)$ . Whenever  $\mathcal{F}_{a,b,c}$  appears, we always assume that  $(a, b, c) \neq (0, 0, 0)$ .

The following proposition states the general form of a function satisfying a generalised second-order recurrence.

**PROPOSITION 5.1.** *Let  $f = [f_0, \dots, f_d] \in \mathcal{F}_{a,b,c}$  be a signature and  $c \neq 0$ . There are two cases:*

- if  $b^2 \neq 4ac$ , then

$$f_k = x\phi_1^k + y\phi_2^k,$$

where  $\phi_1, \phi_2$  are the two roots of the polynomial  $cz^2 + bz + a = 0$  and  $x, y$  are two constants determined by  $f_0$  and  $f_1$ ;

- if  $b^2 = 4ac$ , then

$$f_k = x\phi^k + yk\phi^{k-1},$$

where  $\phi$  is the unique root of the polynomial  $cz^2 + bz + a = 0$  and  $x, y$  are two constants determined by  $f_0$  and  $f_1$ . In case of  $\phi = 0$ , we follow the convention that  $0 \cdot 0^{-1} = 0$ .

In this section, we assume that all signatures (or their reversals) in consideration have nonzero leading term, i.e.,  $f_0 \neq 0$  or  $f_d \neq 0$ . We will discuss the case of  $f_0 = f_d = 0$  in Section 6.

We will use  $\mathcal{F}_{a,b,c}^*$  to denote the subset family of  $\mathcal{F}_{a,b,c}$  with this additional property  $f_0 > 0$ . It turns out that the behaviour of signatures in  $\mathcal{F}_{a,b,c}^*$  is closely related to the sign of the value  $b^2 - 4ac$ , namely the discriminant of the characteristic polynomial  $cz^2 + bz + a$ . Therefore, our discussion is divided into three parts.

**5.1  $b^2 - 4ac > 0$**  In this case, the characteristic polynomial of signatures in  $\mathcal{F}_{a,b,c}^*$  has two distinct real roots. We first single out a special case.

**LEMMA 5.1.** *Let  $f$  be a symmetric signature of arity  $d \geq 3$ , where  $d$  is an odd integer,  $f_i \geq 0$  for all  $i = 0, 1, \dots, d$ , and  $f$  is not identically zero. If there exist  $p, q, s, t \in \mathbb{R}$  such that  $p^2 + q^2 = s^2 + t^2$ ,  $ps + qt < 0$ , and  $f = (p, q)^{\otimes d} + (s, t)^{\otimes d}$ , then up to a non-zero scalar,  $f$  or  $\bar{f}$  is  $[1, 0, \lambda^2, 0, \dots, \lambda^{d-1}, 0]$  for some  $\lambda > 1$ , where  $\bar{f} = [f_d, f_{d-1}, \dots, f_0]$*

*Proof.* Since  $f = (p, q)^{\otimes d} + (s, t)^{\otimes d}$ , we have  $f_i = q^i p^{d-i} + t^i s^{d-i}$ . We discuss the sign of  $qt$ .

First assume  $qt \geq 0$ . The fact  $f_1 \geq 0$  yields

$$qp^{d-1} + ts^{d-1} \geq 0.$$

Since  $d$  is odd, then  $q$  and  $t$  must be both non-negative. Let  $t = \sqrt{p^2 + q^2 - s^2} \geq 0$ . It follows from  $ps + qt < 0$  that  $ps < 0$ . We can assume without loss of generality that  $p > 0, s < 0$  and  $|p| \geq |s|$  (a consequence of  $f_0 \geq 0$ ). To ease the presentation, let  $s' = -s > 0$ . Then

$$\begin{aligned} ps + qt < 0 &\iff qt < ps' \\ &\iff q^2(p^2 + q^2 - s'^2) < p^2 s'^2 \\ &\iff |q| < |s'|. \end{aligned}$$

We then consider the requirement  $f_{d-1} \geq 0$ . This is equivalent to

$$\begin{aligned} q^{d-1}p + t^{d-1}s &\geq 0 \\ \iff q^{d-1}p &\geq t^{d-1}s' \\ \iff q^2 p^{\frac{2}{d-1}} &\geq (p^2 + q^2 - s'^2) s'^{\frac{2}{d-1}} \\ \iff q^2 (p^{\frac{2}{d-1}} - s'^{\frac{2}{d-1}}) &\geq (p^2 - s'^2) s'^{\frac{2}{d-1}}. \end{aligned}$$

We apply  $|q| < |s'|$  and obtain

$$\begin{aligned} (p^2 - s'^2) s'^{\frac{2}{d-1}} &\leq s'^2 (p^{\frac{2}{d-1}} - s'^{\frac{2}{d-1}}) \\ \iff \frac{p^2}{s'^2} - 1 &\leq \frac{p^{\frac{2}{d-1}}}{s'^{\frac{2}{d-1}}} - 1 \\ \iff |s| &\geq |p|. \end{aligned}$$

Therefore, it must hold that  $p = -s, q = t$  and we have  $f = (p, q)^{\otimes d} + (-p, q)^{\otimes d}$ . Moreover,  $ps + qt < 0$  implies that  $p > q$ . If  $q = t = 0$ , then  $f$  is identically zero, a contradiction. Otherwise  $q > 0$ , and we can choose  $\lambda = \frac{p}{q} > 1$  and  $\bar{f}$  is  $[1, 0, \lambda, 0, \lambda^2, 0, \dots]$  up to a non-zero scalar.

Now we assume  $qt < 0$ , and without loss of generality further assume that  $q > 0$  and  $t < 0$ . Let  $t = -\sqrt{p^2 + q^2 - s^2}$ . We distinguish between  $ps \geq 0$  and  $ps < 0$ .

- (i) If  $ps \geq 0$ , then

$$\begin{aligned} ps < -qt &\implies p^2 s^2 < q^2 (p^2 + q^2 - s^2) \\ (5.8) \quad &\implies |s| < |q|. \end{aligned}$$

Again,  $f_1 \geq 0$  implies that  $qp^{d-1} + ts^{d-1} \geq 0$ . This is equivalent to

$$\begin{aligned} qp^{d-1} &\geq \sqrt{p^2 + q^2 - s^2} \cdot s^{d-1} \\ \iff q^2 p^{2d-2} &\geq (p^2 + q^2 - s^2) \cdot s^{2d-2} \\ \iff q^2 (p^{2d-2} - s^{2d-2}) &\geq s^{2d-2} (p^2 - s^2). \end{aligned}$$



Since  $f_d \geq 0$ , we have  $|q| \geq |t|$ . Together with  $p^2 + q^2 = s^2 + t^2$ , it implies that  $p^2 \leq s^2$ . Thus we have either  $|p| = |s|$  or

$$q^2 \leq \frac{s^{2d-2}(s^2 - p^2)}{s^{2d-2} - p^{2d-2}}.$$

If  $|p| = |s|$ , then  $p = s \geq 0$  and  $t = -q$ . In this case,  $f = (p, q)^{\otimes d} + (p, -q)^{\otimes d}$ . If  $p = 0$ , then  $f$  is identically zero, a contradiction. Thus  $p > 0$ , and we can choose  $\lambda = \frac{q}{p}$ , and  $\lambda > 1$  because  $0 > ps + qt = p^2 - q^2$ .

Otherwise, since

$$\begin{aligned} s^{2d-2} - p^{2d-2} &= (s^2 - p^2) \left( \sum_{i=0}^{d-2} s^{2i} p^{2(d-2-i)} \right) \\ &\geq s^{2d-4} (s^2 - p^2), \end{aligned}$$

we have  $q^2 \leq s^2$ . This contradicts to (5.8).

- (ii) If  $ps < 0$ , we first assume that  $p < 0$  and  $s > 0$ . In this case, we let  $p' = -p$  and  $t' = -t$ . Then  $f_0, f_1, f_2 \geq 0$  implies

$$s^d \geq p'^d; \quad t' s^{d-1} \leq q p'^{d-1}; \quad t'^2 s^{d-2} \geq q^2 p'^{d-2},$$

where  $p', q, t', s$  above are all positive. The first two imply that  $t'p' \leq qs$ , and the last two imply that  $t'p' \geq qs$ . Thus  $t'p' = qs$ . This is further equivalent to  $s^2 q^2 = p'^2 (p^2 + q^2 - s^2)$ , or  $(p^2 + q^2)(p^2 - s^2) = 0$ . It implies that either  $p = q = 0$  or  $p = -s$ . In both cases,  $f$  is identically zero, a contradiction.

Finally, consider the case when  $p > 0$  and  $s < 0$ . Then  $f_0 = p^d + s^d \geq 0$  implies  $|p| \geq |s|$ . On the other hand,  $f_d = q^d + t^d \geq 0$  is equivalent to  $|q| \geq |t|$ . However  $p^2 + q^2 = s^2 + t^2$ . Thus we have  $p = -s$  and  $q = -t$ . This means that  $f$  is identically zero, also a contradiction.

Let  $=_d$  be the equality function of arity  $d$ , namely the function  $[1, 0, \dots, 0, 1]$ . We call the problem  $\text{Holant}(=_d | [\beta, 1, \beta])$  a *ferromagnetic Ising model without external fields*, if  $\beta > 1$ . An FPRAS for this problem has been given by Jerrum and Sinclair [JS93]. Then we have the following lemma.

LEMMA 5.2. *Let  $f = [f_0, f_1, \dots, f_d] \in \mathcal{F}_{a,b,c}^*$  with  $b^2 - 4ac > 0$ . Then one of the following holds:*

- *Holant( $f$ ) can be solved exactly in polynomial-time; or*
- *there is an invertible matrix  $M \in \mathbb{C}^{2 \times 2}$  such that  $\text{Holant}(f \cdot M^{\otimes d} | (M^{-1})^{\otimes 2} \cdot (=)_2)$  is a ferromagnetic Ising model without external fields; or*

- *there is an orthogonal matrix  $M \in \mathbf{O}_2(\mathbb{C})$  such that either  $P_{f \cdot M^{\otimes d}}(z)$  or  $P_{\bar{f} \cdot M^{\otimes d}}(z)$  is  $H_\varepsilon$ -stable for some  $\varepsilon > 0$ , where  $\bar{f} = [f_d, f_{d-1}, \dots, f_0]$ ; or*
- *$f$  or  $\bar{f}$  is  $[1, 0, \lambda^2, 0, \lambda^4, 0, \dots, \lambda^{d-1}, 0]$  for some  $\lambda > 1$  and has an odd arity  $d$ .*

*Proof.* If  $c = 0$ , then  $af_k + bf_{k+1} = 0$  for all  $k \leq d - 2$ . Thus,  $f_0, \dots, f_{d-1}$  form a geometric sequence with some ratio  $\phi \in \mathbb{R}$ , and  $f$  can be written as  $f = x(1, \phi)^{\otimes d} + y(0, 1)^{\otimes d}$ , where  $x, y, \phi \in \mathbb{R}$ . Pulling  $x$  and  $y$  into the tensor power, there exist  $p, q, s, t \in \mathbb{R}$  and  $r = 1$  or  $-1$  such that  $f$  is a non-zero multiple of  $(p, q)^{\otimes d} + r(s, t)^{\otimes d}$ .

Otherwise  $c \neq 0$ . It follows from Proposition 5.1 that we can rewrite  $f = x(1, \phi_1)^{\otimes d} + y(1, \phi_2)^{\otimes d}$ , where  $\phi_1, \phi_2 \in \mathbb{R}$  and  $\phi_1 \neq \phi_2$ . Since  $f$  has non-negative weights, it implies that  $x, y \in \mathbb{R}$  as well. Thus, similar to the case above, there exist  $p, q, s, t \in \mathbb{R}$  and  $r = 1$  or  $-1$  such that  $f$  is a non-zero multiple of  $(p, q)^{\otimes d} + r(s, t)^{\otimes d}$ .

The four possibilities of the lemma come from the values these reals might take. If  $pt = qs$ , then  $f$  is degenerate and the partition function can be computed in polynomial time (see e.g. [CC17, Chapter 2]). Thus we assume  $pt - qs \neq 0$  in the following.

First we consider the case that  $p^2 + q^2 = s^2 + t^2$ . We claim that we can always write  $f = (p, q)^{\otimes d} + (s, t)^{\otimes d}$  without loss of generality. To see this, we distinguish between the parity of  $d$ . If  $d$  is odd, then  $(p, q)^{\otimes d} - (s, t)^{\otimes d} = (p, q)^{\otimes d} + (-s, -t)^{\otimes d}$ . If  $d$  is even, we know from  $f = (p, q)^{\otimes d} - (s, t)^{\otimes d}$  that  $f_0 = p^d - s^d$  and  $f_d = q^d - t^d$ . Therefore,  $f_0 > 0$  and  $f_d \geq 0$  imply  $p^2 > s^2$  and  $q^2 \geq t^2$ , which contradicts  $p^2 + q^2 = s^2 + t^2$ .

We write  $\text{Holant}(f)$  as  $\text{Holant}(f | =_2)$ . Let  $M' = \begin{bmatrix} p & q \\ s & t \end{bmatrix}$  be an invertible matrix due to  $pt - qs \neq 0$  and  $M = M'^{-1}$ . It follows from Proposition 4.1 that  $\text{Holant}(f | =_2)$  is equivalent to  $\text{Holant}(f \cdot M^{\otimes d} | (M^{-1})^{\otimes 2} \cdot (=)_2)$ . We verify that this particular Holant problem is either solvable in polynomial-time or equivalent to a ferromagnetic Ising model without external fields. We have

$$\begin{aligned} f \cdot M^{\otimes d} &= \left( (1, 0)^{\otimes d} + (0, 1)^{\otimes d} \right) M'^{\otimes d} M^{\otimes d} \\ &= (1, 0)^{\otimes d} + (0, 1)^{\otimes d}, \end{aligned}$$

and

$$\begin{aligned} (M^{-1})^{\otimes 2} \cdot (=)_2 &= M'^{\otimes 2} \cdot (=)_2 \\ &= (p^2 + q^2, ps + qt, ps + qt, s^2 + t^2)^T. \end{aligned}$$

If  $ps + qt = 0$ , clearly it is solvable in polynomial-time since the edges in every component of the instance must be assigned with the same value in order to contribute

a non-zero weight to the partition function. If  $ps + qt > 0$ , we have that  $(p^2 + q^2)(s^2 + t^2) - (ps + qt)^2 = (pt - qs)^2 > 0$ , and it is a ferromagnetic Ising model without external fields. If  $ps + qt < 0$  and  $d$  is even, then a further transformation  $\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$  makes the middle term positive, and it is a ferromagnetic Ising model again. Otherwise, Lemma 5.1 applies, and we are in the last case of the lemma.

The remaining case is that  $pt \neq qs$  and  $p^2 + q^2 \neq s^2 + t^2$ . If  $|q| = |t|$ , then  $|p| \neq |s|$  and we replace  $(p, q, s, t)$  by  $(q, p, t, s)$ . This is equivalent to work with  $\bar{f}$ . So from now on we also assume that  $|q| \neq |t|$ . Let  $M' = \begin{bmatrix} w & 1 \\ 1 & -w \end{bmatrix}$  where  $w \in \mathbb{R}$  is a parameter to be set later. Then  $f \cdot M'^{\otimes d}$  is  $(q + pw, p - qw)^{\otimes d} + r(t + sw, s - tw)^{\otimes d}$  and

$$P_{f \cdot M'^{\otimes d}}(z) = (q + pw + (p - qw)z)^d + r(t + sw + (s - tw)z)^d.$$

Recall that  $r = 1$  or  $-1$ , so the zeros of this polynomial must satisfy

$$(5.9) \quad |q + pw + (p - qw)z| = |t + sw + (s - tw)z|.$$

We show that by choosing appropriate  $w$  the roots to this equation are in the open left half-plane.

If  $p = q = 0$  and  $s - tw \neq 0$ , the roots to the equation (5.9) must be  $-\frac{t+sw}{s-tw}$ . Since  $p^2 + q^2 \neq s^2 + t^2$ , it holds that  $(s, t) \neq (0, 0)$ . There are four cases.

- If  $t = 0$ , let  $w = 1$ . It holds that  $s - tw = s \neq 0$  and  $-\frac{t+sw}{s-tw} = -w < 0$ .
- If  $s = 0$ , let  $w = -1$ . It holds that  $s - tw = t \neq 0$  and  $-\frac{t+sw}{s-tw} = \frac{1}{w} < 0$ .
- If  $st < 0$ , let  $w = \frac{2s}{t} < 0$ . It holds that  $s - tw = -s \neq 0$  and  $-\frac{t+sw}{s-tw} = \frac{t}{s} + w < 0$ .
- If  $st > 0$ , let  $w = 0$ . It holds that  $s - tw = s \neq 0$  and  $-\frac{t+sw}{s-tw} = -\frac{t}{s} < 0$ .

The case of  $s = t = 0$  is completely analogous.

Now we can make the further assumption that  $(p, q) \neq (0, 0)$  and  $(s, t) \neq (0, 0)$ . Let  $\alpha = -\frac{p-qw}{s-tw} \in \mathbb{R}$  be another parameter, which eventually will be set to 1 or  $-1$ . As  $w = \frac{\alpha s + p}{\alpha t + q}$  and  $|q| \neq |t|$ , the value of the parameter  $w$  will be determined when the sign of  $\alpha$  is chosen. Since  $p - qw = \frac{\alpha(pt - qs)}{\alpha t + q} \neq 0$ , we let  $z_1 = -\frac{q+pw}{p-qw}$  which is well-defined. Similarly it holds that  $s - tw = \frac{qs - pt}{\alpha t + q} \neq 0$ , and we let  $z_2 = -\frac{t+sw}{s-tw}$ . The equation (5.9) is equivalent to

$$(5.10) \quad |\alpha| \cdot |z - z_1| = |z - z_2|.$$

Since  $|\alpha| = 1$ , in order to make the roots to the equation (5.10) in the open left half-plane, it suffices to make sure that

$$(5.11) \quad z_1 + z_2 = \frac{(p^2 + q^2) - (s^2 + t^2)}{\alpha(qs - pt)} < 0.$$

Since  $p^2 + q^2 \neq s^2 + t^2$ , we can let  $\alpha = -1$  if  $\frac{(p^2 + q^2) - (s^2 + t^2)}{qs - pt} > 0$ , or let  $\alpha = 1$  otherwise.

We have showed that there is a matrix  $M' \in \mathbb{C}^{2 \times 2}$  such that the zeros of  $P_{f \cdot M'^{\otimes d}}(z)$  are in the open left half-plane. Since a polynomial has only a finite number of zeros, there is a constant  $\varepsilon > 0$  that  $P_{f \cdot M'^{\otimes d}}(z)$  is  $H_\varepsilon$ -stable. It holds that  $M'(M')^T = \begin{bmatrix} 1+w^2 & 0 \\ 0 & 1+w^2 \end{bmatrix} = (1+w^2)I_2$  where  $1+w^2 > 0$  as  $w \in \mathbb{R}$ . Let  $M = \frac{1}{\sqrt{1+w^2}}M'$ . Clearly  $MM^T = I_2$  and  $M \in \mathbf{O}_2(\mathbb{C})$ . Since  $P_{f \cdot M'^{\otimes d}}(z) = (1+w^2)^{d/2}P_{f \cdot M^{\otimes d}}(z)$ ,  $P_{f \cdot M^{\otimes d}}(z)$  has the same set of zeros as  $P_{f \cdot M'^{\otimes d}}(z)$ . So  $P_{f \cdot M^{\otimes d}}$  is also  $H_\varepsilon$ -stable for some  $\varepsilon > 0$ .

**5.2**  $b^2 - 4ac = 0$  When the characteristic polynomial of  $f$  has only one real root of multiplicity two, we show that there always exists an orthogonal transformation to reduce  $f$  to a function whose local polynomial is  $H_\varepsilon$ -stable.

**LEMMA 5.3.** *Let  $f = [f_0, f_1, \dots, f_d] \in \mathcal{F}_{a,b,c}^*$  with  $b^2 - 4ac = 0$ , then there is an orthogonal matrix  $M \in \mathbf{O}_2(\mathbb{C})$  such that  $P_{f \cdot M^{\otimes d}}(z)$  is  $H_\varepsilon$ -stable for some  $\varepsilon > 0$ .*

*Proof.* If  $c = 0$ , then  $b = 0$  since  $b^2 - 4ac = 0$ . This cannot happen because if so  $f_0$  would be zero. If  $b = 0$ , then  $a = 0$  since  $c \neq 0$  and  $b^2 - 4ac = 0$ . In this case,  $f$  is of form  $[f_0, f_1, 0, \dots, 0]$  and we can simply pick  $M = I_2$ . Clearly  $P_{f \cdot M^{\otimes d}}(z) = f_0 + df_1z$ , which is  $H_\varepsilon$ -stable for some  $\varepsilon > 0$  since  $f_0 > 0$  and  $f_1 \geq 0$ .

So now we assume that  $b \neq 0$ . Since  $c \neq 0$  and  $b^2 - 4ac = 0$ , the equation  $cz^2 + bz + a = 0$  has one real root with multiplicity two and we denote it by  $\phi$ . Note that  $\phi = -\frac{b}{2c} \neq 0$  since  $b \neq 0$ . It follows from Proposition 5.1 that  $f_k = x\phi^k + y'k\phi^{k-1}$  for  $0 \leq k \leq d$  and some  $x, y' \in \mathbb{R}$ . Since  $\phi \neq 0$ , to ease the presentation, we let  $y = \frac{y'}{\phi}$  and rewrite  $f_k = x\phi^k + yk\phi^k$ . Clearly  $x = f_0 > 0$ . We can write  $f$  as

$$x(1, \phi)^{\otimes d} + y \sum_{k=1}^d (1, \phi)^{\otimes(k-1)} \otimes (0, \phi) \otimes (1, \phi)^{\otimes(d-k)}.$$

Let  $M' = \begin{bmatrix} 1 & w \\ -w & 1 \end{bmatrix}$  where  $w \in \mathbb{R}$  is a parameter to be set

later. Then

$$f \cdot M'^{\otimes d} = x(1 - \phi w, \phi + w)^{\otimes d} + y \sum_{k=1}^d (1 - \phi w, \phi + w)^{\otimes(k-1)} \otimes (-\phi w, \phi) \otimes (1 - \phi w, \phi + w)^{\otimes(d-k)},$$

and

$$P_{f \cdot M'^{\otimes d}}(z) = x(1 - \phi w + (\phi + w)z)^d + yd(1 - \phi w + (\phi + w)z)^{d-1}(-\phi w + \phi z).$$

The zeros of this polynomial must satisfy

$$(5.12) \quad (1 - \phi w + (\phi + w)z)^{d-1}$$

$$(5.13) \quad (x - (x + yd)\phi w + (xw + (x + yd)\phi)z) = 0.$$

If  $\phi + w \neq 0$  and  $xw + (x + yd)\phi \neq 0$ , then the roots of this equation must be of the form  $-\frac{1-\phi w}{\phi+w}$  or  $-\frac{x-(x+yd)\phi w}{xw+(x+yd)\phi}$ . We choose appropriate  $w$  and check that these two roots are negative,  $\phi + w \neq 0$  and  $xw + (x + yd)\phi \neq 0$ . Recall that  $\phi \neq 0$  and  $x = f_0 > 0$ . We discuss various cases depending on the sign of  $\phi$  and  $x + yd$ .

- If  $x + yd = 0$ , then the roots of the equation (5.13) are  $-\frac{1-\phi w}{\phi+w}$  and  $-\frac{1}{w}$ . If  $\phi < 0$ , let  $w = -2\phi > 0$  and  $-\frac{1-\phi w}{\phi+w} = \frac{1+2\phi^2}{\phi} < 0$ . If  $\phi > 0$ , let  $w = \frac{1}{2\phi} > 0$  and  $-\frac{1-\phi w}{\phi+w} = -\frac{1}{2\phi+\frac{1}{\phi}} < 0$ . Clearly  $\phi + w \neq 0$  and  $xw + (x + yd)\phi \neq 0$  in both cases.
- If  $\phi > 0$  and  $x + yd > 0$ , then let  $w = \min\left\{\frac{1}{2\phi}, \frac{x}{2(x+yd)\phi}\right\} > 0$ . It holds that

$$-\frac{1 - \phi w}{\phi + w} \leq -\frac{1}{2(\phi + w)} < 0, \\ -\frac{x - (x + yd)\phi w}{xw + (x + yd)\phi} \leq -\frac{x}{2(xw + (x + yd)\phi)} < 0.$$

Whatever  $w = \frac{1}{2\phi}$  or  $w = \frac{x}{2(x+yd)\phi}$ , it is clear that  $\phi + w \neq 0$  and  $xw + (x + yd)\phi \neq 0$ .

- If  $\phi > 0$  and  $x + yd < 0$ , then  $f_d = \phi^d(x + yd) < 0$ . This contradicts to  $f_d \geq 0$ .
- If  $\phi < 0$  and  $x + yd > 0$ , then consider  $f_d = \phi^d(x + yd)$ . If  $d$  is odd, then  $f_d < 0$ . Contradiction. Thus  $d$  must be even. Then  $\phi^{d-1} < 0$ . Since  $f_{d-1} = \phi^{d-1}(x + y(d-1)) \geq 0$ , it holds that  $x + y(d-1) \leq 0$ . As  $x > 0$ ,  $y$  must be negative, and then it contradicts to  $x + yd > 0$ .

- If  $\phi < 0$  and  $x + yd < 0$ , then consider  $f_d = \phi^d(x + yd)$ . If  $d$  is even, then  $f_d < 0$ . But  $f_d$  must be non-negative, so  $d$  must be odd. Then  $\phi^{d-1} > 0$ . Since  $f_{d-1} = \phi^{d-1}(x + y(d-1)) \geq 0$ , it holds that  $x + y(d-1) \geq 0$ . Since  $d > 1$ , we can similarly deduce that  $x + y(d-2) \leq 0$ . This contradicts that  $x > 0$  and  $x + y(d-1) \geq 0$ .

We have showed that there is a matrix  $M' \in \mathbb{C}^{2 \times 2}$  such that the zeros of  $P_{f \cdot M'^{\otimes d}}(z)$  are in the open left half-plane. Since a polynomial has only a finite number of zeros, there is a constant  $\varepsilon > 0$  that  $P_{f \cdot M'^{\otimes d}}(z)$  is  $H_\varepsilon$ -stable. It holds that  $M'M'^T = \begin{bmatrix} 1+w^2 & 0 \\ 0 & 1+w^2 \end{bmatrix} = (1+w^2)I_2$  where  $1+w^2 > 0$  as  $w \in \mathbb{R}$ . Let  $M = \frac{1}{\sqrt{1+w^2}}M'$ , and clearly  $M \in \mathbf{O}_2(\mathbb{C})$ . Since  $P_{f \cdot M'^{\otimes d}}(z) = (1+w^2)^{d/2}P_{f \cdot M^{\otimes d}}(z)$ ,  $P_{f \cdot M^{\otimes d}}(z)$  has the same set of zeros as  $P_{f \cdot M'^{\otimes d}}(z)$ . So  $P_{f \cdot M^{\otimes d}}$  is also  $H_\varepsilon$ -stable for some  $\varepsilon > 0$ .

**5.3**  $b^2 - 4ac < 0$  When the characteristic polynomial of  $f$  has two distinct complex roots, we show that the local polynomial of  $f$  itself is  $H_\varepsilon$ -stable.

LEMMA 5.4. Let  $f = [f_0, f_1, \dots, f_d] \in \mathcal{F}_{a,b,c}^*$  with  $b^2 - 4ac < 0$ , then  $P_f(z)$  is  $H_\varepsilon$ -stable for some  $\varepsilon > 0$ .

*Proof.* It holds that  $c \neq 0$  since otherwise  $b^2 - 4ac \geq 0$ . Since  $c \neq 0$  and  $b^2 - 4ac < 0$ , it follows from Proposition 5.1 that  $f_k = x\phi^k + y\bar{\phi}^k$  for  $0 \leq k \leq d$ , where  $\phi, \bar{\phi}$  are the two conjugate roots of the polynomial  $cz^2 + bz + a = 0$  and  $x, y \in \mathbb{R}$  are constants. Clearly  $x + y = f_0$  and  $x\phi + y\bar{\phi} = f_1$ . Since  $f_0$  is real, it holds that  $\Im(y) = -\Im(x)$ . Since  $f_1$  is real and  $f_1 = x\phi + y\bar{\phi} = (x + y)\Re(\phi) + i(x - y)\Im(\phi)$ , it holds that  $\Re(x) = \Re(y)$ . Thus  $y = \bar{x}$  and  $f_k = x\phi^k + \bar{x}\bar{\phi}^k$  for  $0 \leq k \leq d$ . We write  $f = x(1, \phi)^{\otimes d} + \bar{x}(1, \bar{\phi})^{\otimes d}$  and

$$P_f(z) = x(1 + \phi z)^d + \bar{x}(1 + \bar{\phi} z)^d.$$

The zeros of  $P_f(z)$  must satisfy

$$(5.14) \quad |x| \cdot |1 + \phi z|^d = |\bar{x}| \cdot |1 + \bar{\phi} z|^d.$$

Note that  $\phi \neq 0$ , and  $x \neq 0$  since otherwise  $\bar{x} = 0$  and  $f$  would be  $[0, 0, \dots, 0]$ . So the equation (5.14) is equivalent to

$$\left| z - \left(-\frac{1}{\phi}\right) \right| = \left| z - \left(-\frac{1}{\bar{\phi}}\right) \right|.$$

Since  $-\frac{1}{\phi}$  and  $-\frac{1}{\bar{\phi}}$  are the complex conjugates of each other, the roots of this equation and thus the zeros of

$P_f(z)$  must lie on the real axis. On the other hand, if  $z \geq 0$

$$P_f(z) = \sum_{k=0}^d \binom{n}{k} f_k \cdot z^k > 0,$$

since  $f_0 > 0$ . Thus the zeros of  $P_f(z)$  are negative reals. Since a polynomial has only a finite number of zeros, there is a constant  $\varepsilon > 0$  such that  $P_f(x)$  is  $H_\varepsilon$ -stable.

## 6 Exceptional cases

Section 5 covered all signatures in  $\mathcal{F}_{a,b,c}$  unless  $f_0 = f_d = 0$ . We discuss the remaining cases in this section. We will classify all of them, but the approximation complexity in one case is still open.

Let  $b \in \mathbb{R}$ , and define  $\mathcal{A}_b$  to be the following class

$$\left\{ [f_0, f_1, \dots, f_d] \mid \forall 0 \leq k \leq d-2, \right. \\ \left. \frac{b^2}{4 \cos^2 \frac{\pi}{d}} f_k + b f_{k+1} + f_{k+2} = 0, f_0 = 0 \text{ and } f_1 > 0 \right\}.$$

Notice that  $\mathcal{A}_b$  is a special case of  $\mathcal{F}_{a,b,c}$  except that the parameter  $a$  depends on the arity  $d$ . In fact, if  $f \in \mathcal{A}_b$ , then we can scale  $f$  so that  $f$  has the following form

$$\left[ 0, \lambda \sin \frac{\pi}{d}, \lambda^2 \sin \frac{2\pi}{d}, \dots, \lambda^{d-1} \sin \frac{(d-1)\pi}{d}, 0 \right],$$

for  $\lambda = -\frac{b}{2 \cos \frac{\pi}{d}} > 0$ . (Recall that  $b < 0$ .) Namely,  $f_i = \lambda^i \sin \frac{i\pi}{d}$ .

**LEMMA 6.1.** *Let  $f = [f_0, f_1, \dots, f_d] \in \mathcal{F}_{a,b,c}$  for some  $d \geq 3$ . If  $f_0 = f_d = 0$ , then there are three possibilities:*

- I.  $f \in \mathcal{A}_b$  for some  $b < 0$ ;
- II.  $[f_0, f_1, \dots, f_d]$  is of form  $[0, *, 0, 0, \dots, 0]$  or its reversal  $[0, 0, \dots, 0, *, 0]$ ;
- III.  $[f_0, f_1, \dots, f_d]$  is of form  $\lambda [0, 1, 0, \mu, 0, \mu^2, \dots, 0, \mu^{\frac{d-2}{2}}, 0]$  for some  $\lambda, \mu > 0$  and even  $d$ .

*Proof.* We start by considering the case  $c = 0$ . Then  $a f_k + b f_{k+1} = 0$  for every  $0 \leq k \leq d-2$ . It is easy to verify that  $f$  is identically 0 as  $f_0 = f_d = 0$ , which belongs to type II. Thus, we may assume that  $c \neq 0$  and normalise  $c$  to 1 in the following. There are two further cases depending on whether  $b^2 - 4a = 0$ .

The first case is when  $b^2 - 4a \neq 0$ . It follows from Proposition 5.1 that  $f_0 = x + y = 0$  and  $f_d = x\phi_1^d + y\phi_2^d = 0$ . These two identities together imply

$$x(\phi_1^d - \phi_2^d) = 0,$$

which further implies either  $x = y = 0$  (and therefore  $f_k = 0$  for all  $k$ ) or  $\phi_1^d = \phi_2^d$ . We only need to discuss the case when  $\phi_1^d = \phi_2^d$  and  $x \neq 0$ . There are two possibilities.

- (1) If  $\frac{\phi_1}{\phi_2} \in \mathbb{R}$ , then  $\phi_1 = -\phi_2$  as  $b^2 \neq 4a$ . It implies that  $d$  is even. This is type III.
- (2) Otherwise,  $\frac{\phi_1}{\phi_2} \notin \mathbb{R}$ . In this case,  $b^2 - 4a < 0$  and  $\phi_1$  and  $\phi_2$  are conjugate of each other. By swapping  $\phi_1$  and  $\phi_2$  if necessary, we may assume that  $0 < \arg \phi_1 < \pi$ . Then there exists some integer  $0 < t < d$ ,  $t \neq d/2$ , so that  $\arg \phi_1 = \frac{t\pi}{d}$  and  $\frac{\phi_1}{\phi_2} = e^{\frac{2t\pi}{d}i} \notin \mathbb{R}$ . Since  $a > b^2/4 \geq 0$ ,  $|\phi_1| = |\phi_2| = \sqrt{a}$ , and

$$f_k = x(\phi_1^k - \phi_2^k) = 2x \cdot a^{\frac{k}{2}} \left( \sin \frac{tk\pi}{d} \right) i.$$

Recall that we have the further requirement  $f_k \geq 0$  for every  $0 \leq k \leq d$ . For  $k = 1$ , as  $0 < t < d$ ,  $\sin \frac{t\pi}{d} > 0$ , and thus  $x$  must lie on the negative imaginary axis. Then, it must be that  $\sin \frac{tk\pi}{d} \geq 0$  for all  $0 \leq k \leq d$ . If  $t > 1$ , then taking  $k = \lfloor \frac{d}{t} \rfloor + 1 \leq d$  implies a contradiction. Thus  $t = 1$ .

The assumption  $0 < \arg \phi_1 < \pi$  implies that  $\cos \frac{\pi}{d} = \frac{-b}{2\sqrt{a}} > 0$ . Thus,  $b < 0$  and  $a = \frac{b^2}{4 \cos^2 \frac{\pi}{d}}$ . This verifies that  $f$  is of type I.

At last we turn to the case that  $b^2 - 4a = 0$ . It follows from Proposition 5.1 that  $f_k = x\phi^k + yk\phi^{k-1}$  where  $\phi = -b/2$ . Then  $f_0 = 0$  means that  $x = 0$ , and  $f_d = 0$  means that  $y\phi^{d-1} = 0$ . Thus either  $y = 0$  or  $\phi = 0$ , and any of the two cases implies that  $f$  is of type II.

Next we show that type II and type III signatures are equivalent to approximately counting perfect matchings in general graphs. Denote by  $\text{EXACTONE}_d$  the function  $[0, 1, 0, \dots, 0]$  of arity  $d$ , and by  $\mathcal{EO}$  the (infinite) set  $\{\text{EXACTONE}_d \mid d \in \mathbb{N}^+\}$ . Then  $\text{Holant}(\mathcal{EO})$  is the problem of counting perfect matchings in a graph, denoted  $\#PM$ . (There is only one function for each degree/arity. So the mapping from vertices to functions is obvious for the infinite set  $\mathcal{EO}$ .)

For type III signatures, since multiplying by a constant does not change the complexity, we may assume that  $\lambda = \sqrt{\mu}$ . Then  $f = [0, \lambda, 0, \lambda^3, 0, \dots, \lambda^{d-1}, 0]$  with  $\lambda > 0$ . We will assume  $\lambda < 1$ . This is because that if  $\lambda = 1$ , then the problem is tractable exactly, (see, for example, [CGW16]) and if  $\lambda > 1$ , then taking its reversal makes  $\lambda < 1$ . We adopt the approximation-preserving reduction  $\leq_{AP}$  from [DGGJ04], and use  $\leq_G$  to denote gadget reductions, which is a special form of  $\leq_{AP}$ .

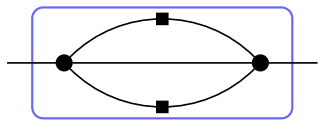


Figure 1: A gadget for type III signatures, where squares are  $[0, 1, 0]$ , and circles are  $[0, 1, 0, 1, 0]$ . All edges are  $[1, 0, \mu]$ .

LEMMA 6.2. *Let  $d \geq 4$  be an even integer, and  $0 < \lambda < 1$ . If  $f = [0, \lambda, 0, \lambda^3, 0, \dots, \lambda^{d-1}, 0]$  of arity  $d$ , then*

$$\text{Holant}(\text{EXACTONE}_4) \leq_{\text{AP}} \text{Holant}(f).$$

*Proof.* Applying a holographic transformation by  $T = \begin{bmatrix} 1 & 0 \\ 0 & \lambda \end{bmatrix}$ , we have that

$$\begin{aligned} \text{Holant}(f) &\equiv \text{Holant}\left(f \cdot (T^{-1})^{\otimes d} \mid T^{\otimes 2} \cdot =_2\right) \\ &\equiv \text{Holant}([0, 1, 0, 1, 0, \dots, 1, 0] \mid [1, 0, \mu]), \end{aligned}$$

where  $0 < \mu = \lambda^2 < 1$ . Thus  $\text{Holant}(f)$  is to count the number of odd subgraphs with edge weight  $\mu$  in a  $d$ -regular graph. Notice that doing a self-loop simply reduces the degree of a vertex by 2, while leaving the constraint on the vertex still requires “odd-degrees”. Thus, with enough self-loops, we may simulate a binary disequality  $[0, 1, 0]$  as well as an arity-4 signature  $[0, 1, 0, 1, 0]$  on the left hand side of the bipartite Holant formulation.

Consider the gadget in Figure 1. Then, it is easy to verify that the effective binary function is  $(2\mu^2 + 2\mu^3)[1, 0, 1]$  on the left hand side. Finally, with  $[1, 0, 1]$  on the left, we can form a path of length  $n$ , and the resulting binary function is  $[1, 0, \mu^n]$  on the right. More formally, we have the following chain of reductions:

$$\begin{aligned} &\text{Holant}([0, 1, 0, 1, 0, \dots, 1, 0] \mid [1, 0, \mu]) \\ &\geq_{\mathbf{G}} \text{Holant}([0, 1, 0], [0, 1, 0, 1, 0] \mid [1, 0, \mu]) \\ &\geq_{\mathbf{G}} \text{Holant}([1, 0, 1], [0, 1, 0, 1, 0] \mid [1, 0, \mu]) \\ &\geq_{\mathbf{G}} \text{Holant}([0, 1, 0, 1, 0] \mid [1, 0, \mu^n]). \end{aligned}$$

The last problem is counting odd subgraphs with  $\mu^n$  edge weights in 4-regular graphs and  $\mu < 1$ . Now, one moment’s reflection realises that odd subgraphs with exponentially small edge weights is approximately perfect matchings, which finishes the reduction.

Similar ideas can also handle the last case in Lemma 5.2, after taking its reversal and renaming  $\lambda$ .

LEMMA 6.3. *Let  $d \geq 3$  be an odd integer, and  $0 < \lambda < 1$ . If  $f = [0, \lambda, 0, \lambda^3, 0, \dots, \lambda^d]$  of arity  $d$ , then*

$$\text{Holant}(\text{EXACTONE}_3) \leq_{\text{AP}} \text{Holant}(f).$$

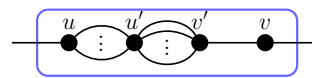


Figure 2: A gadget for weighted equalities. There are  $n_1$  edges between  $u$  and  $u'$ , and  $n_2$  edges between  $u'$  and  $v'$ .

*Proof.* As in the proof of Lemma 6.2, we do the same holographic transformation by  $T = \begin{bmatrix} 1 & 0 \\ 0 & \lambda \end{bmatrix}$ :

$$\text{Holant}(f) \equiv \text{Holant}([0, 1, 0, 1, 0, \dots, 1] \mid [1, 0, \mu]),$$

where  $0 < \mu = \lambda^2 < 1$ . Once again, with sufficiently many self-loops, we get  $[0, 1, 0, 1]$  and  $[0, 1]$  on the left hand side. Connecting  $[0, 1]$  back to  $[0, 1, 0, 1]$  through  $[1, 0, \mu]$  yields  $\mu[1, 0, 1]$  on the left. Thus, similar to the proof of Lemma 6.2, we can simulate  $[1, 0, \mu^n]$  on the right. More formally, we have the following chain of reductions:

$$\begin{aligned} &\text{Holant}([0, 1, 0, 1, 0, \dots, 1] \mid [1, 0, \mu]) \\ &\geq_{\mathbf{G}} \text{Holant}([0, 1], [0, 1, 0, 1] \mid [1, 0, \mu]) \\ &\geq_{\mathbf{G}} \text{Holant}([1, 0, 1], [0, 1, 0, 1] \mid [1, 0, \mu]) \\ &\geq_{\mathbf{G}} \text{Holant}([0, 1, 0, 1] \mid [1, 0, \mu^n]) \\ &\geq_{\text{AP}} \text{Holant}(\text{EXACTONE}_3). \end{aligned}$$

On the other hand, we have the following lemma.

LEMMA 6.4. *Let  $d \geq 3$  be an integer and  $0 < \lambda < 1$ . Let  $f = [0, \lambda, 0, \lambda^3, 0, \dots]$  be a symmetric signature of arity  $d$ . Then*

$$\text{Holant}(f) \leq_{\text{AP}} \#\text{PM}.$$

*Proof.* First, by the same holographic transformations as in the proofs of Lemma 6.2 and Lemma 6.3,

$$\text{Holant}(f) \equiv \text{Holant}([0, 1, 0, 1, 0, \dots] \mid [1, 0, \mu]),$$

where  $\mu = \lambda^2 > 0$ .

Consider the gadget in Figure 2, where all vertices are the “exact one” function, namely  $[0, 1, 0, \dots, 0]$ . It is easy to see that this gadget is equivalent to a weighted equality  $[1, 0, \frac{n_2}{n_1}]$ . Thus we can use it to arbitrarily closely approximate  $[1, 0, \mu]$  by tuning  $n_1$  and  $n_2$  for any  $\mu > 0$ .

In addition, consider the gadget in Figure 3, where, once again, all vertices are  $[0, 1, 0, 0]$ . The resulting signature is  $[0, 1, 0, 1]$ .

A simple calculation verifies that a sequence of  $d$  signatures  $[0, 1, 0, 1]$  connected together, as in Figure 4, yields a signature  $[0, 1, 0, \dots, 1, 0]$  of arity  $d + 2$  if  $d$  is odd, or a signature  $[1, 0, 1, 0, \dots, 1]$  of arity  $d + 2$  if  $d$  is even. In the even case, to get  $[0, 1, 0, 1, \dots, 0]$ , we

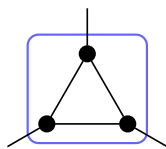


Figure 3: A gadget to create  $[0, 1, 0, 1]$ .

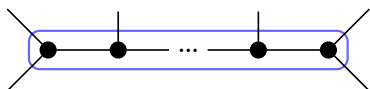


Figure 4: A gadget to create  $[1, 0, 1, 0, \dots, 1]$  or  $[0, 1, 0, 1, \dots, 0]$ .

simply connect one of its dangling edges with  $[0, 1, 0]$ . Formally, we have the following sequence of reductions:

$$\begin{aligned} & \text{Holant}([0, 1, 0, 1, 0, \dots] \mid [1, 0, \mu]) \\ & \leq_{\mathbb{G}} \text{Holant}(\{[0, 1, 0, 1, 0, \dots], [1, 0, \mu]\}) \\ & \leq_{\text{AP}} \#\text{PM}. \end{aligned}$$

Lemma 6.2, Lemma 6.3, and Lemma 6.4 together imply the following:

$$(6.15) \quad \text{Holant}(\text{EXACTONE}_4) \leq_{\text{AP}} \text{Holant}(f) \leq_{\text{AP}} \#\text{PM},$$

if  $d$  is even,

$$(6.16) \quad \text{Holant}(\text{EXACTONE}_3) \leq_{\text{AP}} \text{Holant}(f) \leq_{\text{AP}} \#\text{PM},$$

if  $d$  is odd,

where  $f = [0, \lambda, 0, \lambda^3, 0, \dots]$  for some  $0 < \lambda < 1$  has arity  $d \geq 3$ . Note that  $\text{Holant}(\text{EXACTONE}_3)$  or  $\text{Holant}(\text{EXACTONE}_4)$  is just an alias of counting perfect matchings in 3- or 4-regular graphs, which is equivalent to  $\#\text{PM}$  in approximation. This is a folklore fact, and is shown in the next couple of lemmas.

LEMMA 6.5.  $\text{Holant}(\text{EXACTONE}_3) \leq_{\text{AP}} \text{Holant}(\text{EXACTONE}_4)$ .

*Proof.* Note that a self-loop on  $[0, 1, 0, 0, 0]$  gives  $[0, 1, 0]$ , and connecting it back to  $[0, 1, 0, 0, 0]$  yields  $[1, 0, 0]$ . Thus,

$$\text{Holant}([0, 1, 0, 0, 0], [1, 0, 0]) \leq_{\mathbb{G}} \text{Holant}([0, 1, 0, 0, 0]).$$

Given an instance  $G$  (namely a 3-regular graph) of  $\text{Holant}([0, 1, 0, 0])$ , consider a disjoint union of  $G$  and its copy  $G'$ . We add a new vertex  $u$  for each pair  $v$  and  $v'$ , and connect  $u$  to both  $v$  and  $v'$ . Now all original vertices in  $G$  and  $G'$  have degrees exactly 4. Put

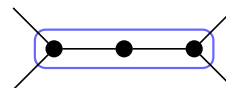


Figure 5: A gadget to create  $\text{EXACTONE}_d$ .

$[0, 1, 0, 0, 0]$  on all these vertices, and  $[1, 0, 0]$  on all new vertices. It is easy to see that the partition function of this new instance is the square of the number of perfect matchings of  $G$ . Thus, we have the following reduction chain:

$$\begin{aligned} \text{Holant}([0, 1, 0, 0]) & \leq_{\text{AP}} \text{Holant}([0, 1, 0, 0, 0], [1, 0, 0]) \\ & \leq_{\mathbb{G}} \text{Holant}([0, 1, 0, 0, 0]). \end{aligned}$$

However, approximate counting perfect matchings in 3-regular graphs is as hard as that in general graphs.

LEMMA 6.6.  $\#\text{PM} \leq_{\text{AP}} \text{Holant}(\text{EXACTONE}_3)$ .

*Proof.* Consider the gadget in Figure 5.

Notice that if we put  $[0, 1, 0, 0]$  on the two degree three vertices, and  $[0, 1, 0]$  on the middle vertex, the resulting signature is  $[0, 1, 0, 0, 0]$ . More generally, if we replace one of the degree three vertex by  $\text{EXACTONE}_d$ , then the resulting signature is  $\text{EXACTONE}_{d+1}$ . Namely, using this gadget, we can simulate the whole set of  $\mathcal{EO}$ , and

$$\#\text{PM} \leq_{\mathbb{G}} \text{Holant}([0, 1, 0, 0], [0, 1, 0]).$$

Moreover, a self-loop on  $[0, 1, 0, 0]$  gives  $[0, 1]$ , and connecting back to it gives  $[1, 0, 0]$ . By using the same squaring trick in Lemma 6.5, we can use  $[1, 0, 0]$  as  $[1, 0]$ . Thus, we have the following reduction chain:

$$\begin{aligned} \#\text{PM} & \leq_{\mathbb{G}} \text{Holant}([0, 1, 0, 0], [0, 1, 0]) \\ & \leq_{\mathbb{G}} \text{Holant}([0, 1, 0, 0], [1, 0]) \\ & \leq_{\text{AP}} \text{Holant}([0, 1, 0, 0], [1, 0, 0]) \\ & \leq_{\mathbb{G}} \text{Holant}([0, 1, 0, 0]). \end{aligned}$$

Holant problems defined by type II signatures are counting perfect matchings in  $d$ -regular graphs. Clearly, by doing sufficiently many self-loops, either  $\text{Holant}(\text{EXACTONE}_3) \leq_{\text{AP}} \text{Holant}(\text{EXACTONE}_d)$  or  $\text{Holant}(\text{EXACTONE}_4) \leq_{\text{AP}} \text{Holant}(\text{EXACTONE}_d)$ , depending on the parity of  $d$ . Thus, combining this fact with Lemma 6.5, Lemma 6.6, (6.15) and (6.16), we have the following result.

LEMMA 6.7. Let  $f = [0, 1, 0, \lambda^2, 0, \dots]$  for some  $0 \leq \lambda < 1$ . Then

$$\text{Holant}(f) \equiv_{\text{AP}} \#\text{PM}.$$

Notice that in Lemma 6.7 we manipulate the form a little bit so that it cover type II and type III in Lemma 6.1, as well as the last case in Lemma 5.2.

## 7 Proof of main theorems

We are now ready to assemble all the ingredients to prove our main theorems. We restate Theorem 1.1 for convenience.

**THEOREM 7.1.** *Let  $f = [f_0, f_1, \dots, f_d]$  be a symmetric constraint function of arity  $d \geq 3$  satisfying generalised second-order recurrences, and  $f_i \geq 0$  for all  $0 \leq i \leq d$ . There is a fully polynomial-time (deterministic or randomised) approximation algorithm for  $\text{Holant}(f)$ , unless, up to a non-zero factor,  $f$  or its reversal is in one of the following form:*

- $[0, \lambda \sin \frac{\pi}{d}, \lambda^2 \sin \frac{2\pi}{d}, \dots, \lambda^i \sin \frac{i\pi}{d}, \dots, 0]$  for some  $\lambda > 0$ ;
- $[0, 1, 0, \lambda, 0, \dots, 0, \lambda^{\frac{d-2}{2}}, 0]$  if  $d$  is even, or  $[0, 1, 0, \lambda, 0, \dots, 0, \lambda^{\frac{d-1}{2}}]$  if  $d$  is odd, for some  $0 \leq \lambda < 1$ .

Moreover, in the latter case, approximating  $\text{Holant}(f)$  is equivalent to approximately counting perfect matchings in general graphs.

*Proof.* We apply Lemma 5.2, Lemma 5.3 and Lemma 5.4. Then one of followings must happen

1.  $f_0 = f_d = 0$ ; or
2.  $f$  or  $\bar{f}$  is  $[1, 0, \lambda^2, 0, \lambda^4, 0, \dots]$  for some  $\lambda > 1$  and has an odd arity; or
3.  $\text{Holant}(f)$  can be solved exactly in polynomial-time; or
4. there is an invertible matrix  $M \in \mathbb{C}^{2 \times 2}$  such that  $\text{Holant}(f \cdot M^{\otimes d} \mid (M^{-1})^{\otimes 2} \cdot (=_{=2}))$  is a ferromagnetic two-spin system; or
5. there is an orthogonal matrix  $M \in \mathbf{O}_2(\mathbb{C})$  such that either  $P_{f \cdot M^{\otimes d}}(z)$  or  $P_{\bar{f} \cdot M^{\otimes d}}(z)$  is  $H_\varepsilon$ -stable for some  $\varepsilon > 0$ , where  $\bar{f}$  is the reversal of  $f$ .

We are done in Case (3), as well as in Case (5) by Proposition 4.2 and Theorem 3.1. In Case (4), we invoke the FPRAS by Jerrum and Sinclair [JS93]. In Case (1) and Case (2), we are in the desired form of the theorem by Lemma 6.1. (In case  $\mu > 1$  in Lemma 6.1, we can take its reversal so that  $\mu < 1$ , and if  $\mu = 1$ , then exact counting is tractable [CGW16].) Finally, the approximation complexity of  $[0, 1, 0, \lambda, 0, \lambda^2, 0, \dots]$  signatures is handled in Lemma 6.7.

**REMARK 7.1.** *It is worth noticing that our algorithm applies beyond regular graphs. In fact, for any finite family of signatures  $\mathcal{F}$ , we can define  $\text{Holant}(\mathcal{F})$  as*

*the problem of computing the partition function on a graph  $G$ , where each vertex  $v$  of  $G$  is associated with a function  $f_v \in \mathcal{F}$ . It is straightforward to adapt the algorithm described in the proof of Theorem 1.1 for one to solve  $\text{Holant}(\mathcal{F})$ <sup>1</sup>. It is not hard to see the adaptation provides an efficient approximation algorithm for  $\text{Holant}(\mathcal{F})$  as long as there exists an orthogonal matrix  $M \in \mathbf{O}_2(\mathbb{C})$  and  $\varepsilon > 0$  such that  $P_{f \cdot M^{\otimes d}}$  is  $H_\varepsilon$ -stable for every  $f \in \mathcal{F}$ , where  $d$  is the arity of  $f$ . For example, we can let  $\mathcal{F}$  be the family of signatures for matchings up to arity  $d$ , or the family of signatures for edge covers up to arity  $d$ . Therefore, our algorithm recovers a number of previously known deterministic approximation algorithms for special cases of Holant problems, such as counting matchings [BGK<sup>+</sup>07, PR17a] and counting edge covers in bounded degree graphs [LLL14].*

*On the other hand, even for the same tuple  $(a, b, c)$ , signatures in  $\mathcal{F}_{a,b,c}$  may require different  $M$  to be  $H_\varepsilon$ -stable. It is not clear how to obtain an algorithm in such cases.*

We deduce Theorem 1.2 from Theorem 1.1 by noting that all ternary signatures satisfy generalised second-order recurrence relations. Therefore, we only need to deal with the case where  $f = [0, a, b, 0]$  for some  $a, b > 0$ .

We design an FPRAS for  $\text{Holant}(f)$  using the machinery called “winding” developed in [McQ13, HLZ16]. We sketch the construction here without getting into too much technical details, which is out of the scope of the current paper. We break every edge into two half edges, and then simulate a Markov chain whose state space consists of all consistent edge assignments and assignments with at most two inconsistencies. It has been shown by McQuillan [McQ13] that the Markov chain mixes rapidly as long as the signature  $f$  is *windable*. It is then straightforward to use the algebraic characterization of windable functions in [HLZ16] to verify that every function of the form  $[0, a, b, 0]$  with non-negative  $a, b$  is windable. At last, it is trivial to check that, using the notations in [McQ13], the signature  $[0, a, b, 0]$  is *strictly terraced* when both  $a, b > 0$ . This fact implies that the ratio between the total weight of nearly consistent assignments and that of consistent assignments can be bounded by a polynomial in the size of the instance. Therefore, we obtain an efficient Gibbs sampler for  $\text{Holant}(f)$ , which can be turned into

<sup>1</sup>The main adaptation is to show that  $Z_i(G)$  is still a BIGCP when more than one constraint function are present. Since  $\mathcal{F}$  is finite, we can therefore view functions in  $\mathcal{F}$  as colors and enumerate *vertex colored induced subgraphs* instead of ordinary induced subgraphs in the proof of Lemma 3.2. Similar technique already appeared in [PR17a]

an FPRAS to compute the partition function via self-reduction [JVV86].

The remaining open case in Theorem 1.1 is when  $f \in \mathcal{A}_b$  of arity  $d \geq 3$ . Numerical evidences suggest that these signatures are windable, via the criteria in [HLZ16]. We conjecture that this is indeed the case, which would imply FPRAS for computing the partition functions of type I signatures, since this class is “strictly terraced” in the language of [McQ13].

## Acknowledgements

Part of the work was done while HG, CL, and CZ were visiting the Institute of Theoretical Computer Science, Shanghai University of Finance and Economics, and we would like to thank their hospitality. We thank Leslie Ann Goldberg and Mark Jerrum for pointing out a bug in Lemma 5.2 in an earlier version. We thank anonymous referees for many useful comments.

## References

- [Asa70] Taro Asano. Theorems on the partition functions of the Heisenberg ferromagnets. *J. Phys. Soc. Japan*, 29:350–359, 1970. 3
- [Bac18] Miriam Backens. A complete dichotomy for complex-valued Holant<sup>c</sup>. In *ICALP*, volume 107 of *LIPICs*, pages 12:1–12:14, 2018. 1
- [Bar16] Alexander I. Barvinok. *Combinatorics and Complexity of Partition Functions*, volume 30 of *Algorithms and combinatorics*. Springer, 2016. 2, 4, 5
- [BGK<sup>+</sup>07] Mohsen Bayati, David Gamarnik, Dimitriy A. Katz, Chandra Nair, and Prasad Tetali. Simple deterministic approximation algorithms for counting matchings. In *STOC*, pages 122–127, 2007. 15
- [Bul13] Andrei A. Bulatov. The complexity of the counting constraint satisfaction problem. *J. ACM*, 60(5):34:1–34:41, 2013. 1
- [CC17] Jin-Yi Cai and Xi Chen. Complexity of counting CSP with complex weights. *J. ACM*, 64(3):19:1–19:39, 2017. 1, 9
- [CGW16] Jin-Yi Cai, Heng Guo, and Tyson Williams. A complete dichotomy rises from the capture of vanishing signatures. *SIAM J. Comput.*, 45(5):1671–1728, 2016. 1, 12, 15
- [CLX11] Jin-yi Cai, Pinyan Lu, and Mingji Xia. Computational complexity of Holant problems. *SIAM J. Comput.*, 40(4):1101–1132, 2011. 1, 7
- [CLX18] Jin-Yi Cai, Pinyan Lu, and Mingji Xia. Dichotomy for real Holant<sup>c</sup> problems. In *SODA*, pages 1802–1821, 2018. 1
- [DGGJ04] Martin E. Dyer, Leslie Ann Goldberg, Catherine S. Greenhill, and Mark Jerrum. The relative complexity of approximate counting problems. *Algorithmica*, 38(3):471–500, 2004. 12
- [DGL<sup>+</sup>12] Jan Draisma, Dion C. Gijswijt, László Lovász, Guus Regts, and Alexander Schrijver. Characterizing partition functions of the vertex model. *J. Algebra*, 350:197–206, 2012. 1
- [DJM17] Martin E. Dyer, Mark Jerrum, and Haiko Müller. On the switch Markov chain for perfect matchings. *J. ACM*, 64(2):12:1–12:33, 2017. 2
- [DR13] Martin E. Dyer and David Richerby. An effective dichotomy for the counting constraint satisfaction problem. *SIAM J. Comput.*, 42(3):1245–1274, 2013. 1
- [FLS07] Michael Freedman, László Lovász, and Alexander Schrijver. Reflection positivity, rank connectivity, and homomorphism of graphs. *J. Amer. Math. Soc.*, 20(1):37–51, 2007. 1
- [GJP03] Leslie Ann Goldberg, Mark Jerrum, and Mike Paterson. The computational complexity of two-state spin systems. *Random Struct. Algorithms*, 23(2):133–154, 2003. 2
- [Gra02] John H. Grace. The zeros of a polynomial. *Proc. Camb. Philos. Soc.*, 11:352–357, 1902. 3
- [GŠV16] Andreas Galanis, Daniel Štefankovič, and Eric Vigoda. Inapproximability of the partition function for the antiferromagnetic Ising and hard-core models. *Comb. Probab. Comput.*, 25(4):500–559, 2016. 2
- [HLZ16] Lingxiao Huang, Pinyan Lu, and Chihao Zhang. Canonical paths for MCMC: from art to science. In *SODA*, pages 514–527, 2016. 2, 3, 15, 16
- [JS89] Mark Jerrum and Alistair Sinclair. Approximating the permanent. *SIAM J. Comput.*, 18(6):1149–1178, 1989. 2
- [JS93] Mark Jerrum and Alistair Sinclair. Polynomial-time approximation algorithms for the Ising model. *SIAM J. Comput.*, 22(5):1087–1116, 1993. 2, 9, 15
- [JSV04] Mark Jerrum, Alistair Sinclair, and Eric Vigoda. A polynomial-time approximation algorithm for the permanent of a matrix with nonnegative entries. *J. ACM*, 51(4):671–697, 2004. 2
- [JVV86] Mark Jerrum, Leslie G. Valiant, and Vijay V. Vazirani. Random generation of combinatorial structures from a uniform distribution. *Theor. Comput. Sci.*, 43:169–188, 1986. 16
- [LLL14] Chengyu Lin, Jingcheng Liu, and Pinyan Lu. A simple FPTAS for counting edge covers. In *SODA*, pages 341–348, 2014. 2, 15
- [LLY13] Liang Li, Pinyan Lu, and Yitong Yin. Correlation decay up to uniqueness in spin systems. In *SODA*, pages 67–84, 2013. 2
- [LSS17] Jingcheng Liu, Alistair Sinclair, and Piyush Srivastava. The Ising partition function: Zeros and deterministic approximation. In *FOCS*, pages 986–997, 2017. 3
- [LW18] Jiabao Lin and Hanpin Wang. The complexity of Boolean Holant problems with nonnegative weights. *Siam J. Comput.*, 47(3):798–828, 2018. 1
- [LWZ14] Pinyan Lu, Menghui Wang, and Chihao Zhang. FPTAS for weighted Fibonacci gates and its applications. In *ICALP*, pages 787–799, 2014. 2
- [McQ13] Colin McQuillan. Approximating Holant problems by winding. *CoRR*, abs/1301.2880, 2013. 2, 3, 15, 16
- [PR17a] Viresh Patel and Guus Regts. Deterministic polynomial-time approximation algorithms for parti-



- tion functions and graph polynomials. *SIAM J. Comput.*, 46(6):1893–1919, 2017. [2](#), [6](#), [15](#)
- [PR17b] Han Peters and Guus Regts. On a conjecture of Sokal concerning roots of the independence polynomial. *CoRR*, abs/1701.08049, 2017. To appear in *Mich. Math. J.* [3](#)
- [Rue71] David Ruelle. Extension of the Lee-Yang circle theorem. *Phys. Rev. Lett.*, 26:303–304, 1971. [2](#), [3](#)
- [Rue99a] David Ruelle. Counting unbranched subgraphs. *J. Algebraic Combin.*, 9(2):157–160, 1999. [2](#), [3](#)
- [Rue99b] David Ruelle. Zeros of graph-counting polynomials. *Comm. Math. Phys.*, 200(1):43–56, 1999. [2](#), [3](#)
- [Sch13] Alexander Schrijver. Characterizing partition functions of the spin model by rank growth. *Indag. Math. (N.S.)*, 24(4):1018–1023, 2013. [1](#)
- [SS14] Allan Sly and Nike Sun. The computational hardness of counting in two-spin models on  $d$ -regular graphs. *Ann. Probab.*, 42(6):2383–2416, 2014. [2](#)
- [SST14] Alistair Sinclair, Piyush Srivastava, and Marc Thurley. Approximation algorithms for two-state anti-ferromagnetic spin systems on bounded degree graphs. *J. Stat. Phys.*, 155(4):666–686, 2014. [2](#)
- [ŠVW18] Daniel Štefankovič, Eric Vigoda, and John Wilmes. On counting perfect matchings in general graphs. In *LATIN*, pages 873–885, 2018. [2](#)
- [Sze22] Gábor Szegő. Bemerkungen zu einem Satz von J. H. Grace über die Wurzeln algebraischer Gleichungen. *Math. Z.*, 13(1):28–55, 1922. [3](#)
- [Val08] Leslie G. Valiant. Holographic algorithms. *SIAM J. Comput.*, 37(5):1565–1594, 2008. [1](#), [2](#), [7](#)
- [Wal22] Joseph L. Walsh. On the location of the roots of certain types of polynomials. *Trans. Amer. Math. Soc.*, 24(3):163–180, 1922. [3](#)