

Online Appendix to: Inapproximability for Antiferromagnetic Spin Systems in the Tree Nonuniqueness Region

ANDREAS GALANIS, University of Oxford
DANIEL ŠTEFANKOVIČ, University of Rochester
ERIC VIGODA, Georgia Institute of Technology

This online appendix contains proofs and supporting material to our article.

A. THE SMALL SUBGRAPH CONDITIONING METHOD

In this appendix, we prove Lemma 6.12 by applying the small subgraph conditioning method.

A.1. Overview

The small subgraph conditioning method was introduced by Robinson and Wormald [1994] to prove that a random Δ -regular contains asymptotically almost surely (a.a.s.) a Hamilton cycle. Roughly speaking, the method provides a way to get a.a.s. results when the second moment method fails, in the particular case (though common in the random regular setting) where the ratio of the second moment of a variable to the first moment squared converges to a constant strictly greater than 1.

The method was first used to analyze spin models on random regular graphs in Mossel et al. [2009] and was subsequently used in Sly [2010] and Galanis et al. [2012]. In our setting, applying the small subgraph conditioning method of Robinson and Wormald [1994] as in the previous works [Mossel et al. 2009; Sly 2010; Galanis et al. 2012] would not be sufficient, since it guarantees a polynomial multiplicative deviation from the expectation, which is weak in the setting of Lemma 6.9. We instead use an extension of the method given by Janson [1995].

More generally, the method of Robinson and Wormald [1994] is sufficient when the interest is in proving concentration of a variable within a polynomial factor from its expectation. Janson's refinement of the method gives the distributional limit of the variable and explicitly attributes the fluctuations from the expectation to the presence of specific subgraph structures. For the convenience of the reader, we include both versions of the method in Theorem A.1, which is a concatenated version of the respective Theorems in Robinson and Wormald [1994] and Janson [1995]. The theorem can be extrapolated from Janson [1995], after combining Janson [1995, Lemma 1, Remark 4, Remark 9]. The notation $[X]_m$ refers to the m th order falling factorial of the variable X . We shall discuss the theorem statement afterwards.

THEOREM A.1. *Let S be a set of finite cardinality. For $s \in S$ and $i = 1, 2, \dots$, let $\mu_i > 0$ and $\delta_i^{(s)} > -1$ be constants and assume that for each n there are random variables X_{in} , $i = 1, 2, \dots$, and $Y_n^{(s)}$, $s \in S$, all defined on the same probability space $\mathcal{G} = \mathcal{G}_n$ such that X_{in} is nonnegative integer valued, $Y_n^{(s)} \geq 0$ and $\mathbf{E}[Y_n^{(s)}] > 0$ (for n sufficiently large). Furthermore, for every $s \in S$, the following hold:*

(A1) $X_{in} \xrightarrow{d} Z_i$ as $n \rightarrow \infty$, jointly for all i , where $Z_i \sim \text{Po}(\mu_i)$ are independent Poisson random variables;

(A2) for every finite sequence j_1, \dots, j_m of nonnegative integers,

$$\frac{\mathbf{E}_{\mathcal{G}}[Y_n^{(s)}[X_{1n}]_{j_1} \cdots [X_{mn}]_{j_m}]}{\mathbf{E}_{\mathcal{G}}[Y_n^{(s)}]} \rightarrow \prod_{i=1}^m (\mu_i (1 + \delta_i^{(s)}))^i \quad \text{as } n \rightarrow \infty; \quad (150)$$

(A3) $\sum_i \mu_i (\delta_i^{(s)})^2 < \infty$;

(A4) $\mathbf{E}_{\mathcal{G}}[(Y_n^{(s)})^2]/(\mathbf{E}_{\mathcal{G}}[Y_n^{(s)}])^2 \leq \exp(\sum_i \mu_i (\delta_i^{(s)})^2) + o(1)$ as $n \rightarrow \infty$;

Then, the following conclusions hold:

(C1) Let $r(n)$ be a function such that $r(n) \rightarrow 0$ as $n \rightarrow \infty$. For each $s \in S$, it holds that $Y_n^{(s)} > r(n)\mathbf{E}_{\mathcal{G}}[Y_n^{(s)}]$ asymptotically almost surely.

(C2) For $s \in S$,

$$\frac{Y_n^{(s)}}{\mathbf{E}_{\mathcal{G}}[Y_n^{(s)}]} \xrightarrow{d} W^{(s)} = \prod_i [1 + \delta_i^{(s)}]^{Z_i} \exp(-\mu_i \delta_i^{(s)}). \quad (151)$$

This and the convergence in A.1 hold jointly. The infinite product defining $W^{(s)}$ converges almost surely (a.s.) and in L^2 , with

$$\mathbf{E}[W^{(s)}] = 1 \text{ and } \mathbf{E}[(W^{(s)})^2] = \lim_{n \rightarrow \infty} \mathbf{E}_{\mathcal{G}}[(Y_n^{(s)})^2]/(\mathbf{E}_{\mathcal{G}}[Y_n^{(s)}])^2.$$

Moreover, $W^{(s)} > 0$ a.s. iff $\delta_i^{(s)} > -1$ for all i ;

The random variables $Y_n^{(s)}$ in Theorem A.1 are those we are interested in obtaining ‘‘concentration’’ type results, where s is simply an index allowing us to treat simultaneously more than one variables. In our setting, for $G \sim \mathcal{G}_n^r$, $Y_n^{(s)}$ are going to be the variables $Z_G^{\mathbf{p}}(\eta)$ for phases $\mathbf{p} \in \mathcal{Q}$ and configurations η on W . The random variables X_{in} , for graphs with no small multicyclic components, correspond to cycles of length i . For example, in our setting and because the graph G is bipartite, X_{in} is the number of cycles of length i in G where i is even.

The conclusion (C1) of Theorem A.1 is essentially due to Robinson and Wormald [1994], while the conclusion (C2) is an extension of conclusion (C1) due to Janson [1995]. At this point, to obtain Lemma 6.12 (which was the important part to prove Lemma 6.9), we will not explicitly use either of (C1) or (C2) but rather the following variant. The variant was observed in Janson [1995, p. 5], who discusses it without proof in a specific setting, and is also implicit in Robinson and Wormald [1994]. As such, we write and prove a formal statement in the setup of Theorem A.1. The proof follows Janson’s proof of Theorem A.1 but uses a different finish.

LEMMA A.2. Assume that the conditions in Theorem A.1 hold. For an integer $m > 0$ and $s \in S$, let

$$W_{mn}^{(s)} = \prod_{i=1}^m (1 + \delta_i^{(s)})^{X_{in}} \exp(-\mu_i \delta_i^{(s)}).$$

Then, for every $\varepsilon > 0$, it holds that

$$\lim_{m \rightarrow \infty} \limsup_{n \rightarrow \infty} \Pr_{\mathcal{G}_n} \left(\bigcup_{s \in S} \left[\left| \frac{Y_n^{(s)}}{\mathbf{E}_{\mathcal{G}_n}[Y_n^{(s)}]} - W_{mn}^{(s)} \right| > \varepsilon \right] \right) = 0. \quad (152)$$

PROOF OF LEMMA A.2. We prove the statement for a fixed $s \in S$, the extension of the argument to prove (152) is straightforward (e.g., by a union bound) and is omitted. To lighten notation, we will drop s from the notation and without loss of generality, we

also assume $\mathbf{E}_{\mathcal{G}_n}[Y_n] = 1$. We will prove that

$$\limsup_{n \rightarrow \infty} \Pr_{\mathcal{G}_n}(|Y_n - W_{mn}| > \varepsilon) \leq \frac{1}{4} \varepsilon^{-2} \left[\exp\left(\sum_{i=1}^{\infty} \mu_i \delta_i^2\right) - \exp\left(\sum_{i=1}^m \mu_i \delta_i^2\right) \right]. \quad (153)$$

This clearly gives the statement of the lemma, since by assumption (A3) of Theorem A.1, the left-hand side is finite and goes to 0 as $m \rightarrow \infty$. To prove (153), we follow Janson [1995, Proof of Theorem 1] up to a certain point but avoid the use of Skorokhod's theorem in the argument. Janson's proof goes as follows. For a positive integer m , define the functions

$$\begin{aligned} f_n(x_1, \dots, x_m) &= \mathbf{E}_{\mathcal{G}_n}[Y_n | X_{1n} = x_1, \dots, X_{mn} = x_m], \\ f_{\infty}(x_1, \dots, x_m) &= \lim_{n \rightarrow \infty} f_n(x_1, \dots, x_m) = \prod_{i=1}^m (1 + \delta_i)^{x_i} e^{-\mu_i \delta_i}. \end{aligned} \quad (154)$$

The second equality follows by assumption (A2) of Theorem A.1 and Janson [1995, Lemma 1]. Define also the random variable

$$Y_n^{(m)} = \mathbf{E}_{\mathcal{G}_n}[Y_n | X_{1n}, \dots, X_{mn}].$$

Using assumptions (A1) and (A2), Fatou's Lemma and that $Y_n^{(m)}$ is a conditional expectation of Y_n , one obtains

$$\limsup_{n \rightarrow \infty} \mathbf{E}_{\mathcal{G}_n}[|Y_n - Y_n^{(m)}|^2] \leq \exp\left(\sum_{i=1}^{\infty} \mu_i \delta_i^2\right) - \exp\left(\sum_{i=1}^m \mu_i \delta_i^2\right),$$

see Janson [1995, Eq. (5.2)] for details. We now give the main deviation point from Janson's proof, which amounts to proving that for fixed m , we have

$$\lim_{n \rightarrow \infty} \Pr_{\mathcal{G}_n}(|Y_n^{(m)} - W_{mn}| > \varepsilon) = 0. \quad (155)$$

as $n \rightarrow \infty$. Fix $M > 0$. By (154), there is N such that for $n \geq N$ it holds that

$$|f_n(x_1, \dots, x_m) - f_{\infty}(x_1, \dots, x_m)| < \varepsilon \text{ for all integer } x_1, \dots, x_m \in [0, M].$$

It follows that, for $n \geq N$, we have

$$\Pr_{\mathcal{G}_n}(|Y_n^{(m)} - W_{mn}| > \varepsilon) \leq \Pr_{\mathcal{G}_n}\left(\bigcup_{i=1}^m [X_{in} > M]\right)$$

Note that as $n \rightarrow \infty$, the right-hand side by assumption (A1) converges to $\Pr(\bigcup_{i=1}^m [Z_i > M])$. The latter can be made arbitrarily small by letting $M \rightarrow \infty$. This proves (155).

The final step is to bound

$$\begin{aligned} & \limsup_{n \rightarrow \infty} \Pr_{\mathcal{G}_n}(|Y_n - W_{mn}| > \varepsilon) \\ & \leq \limsup_{n \rightarrow \infty} \Pr_{\mathcal{G}_n}(|Y_n - Y_n^{(m)}| > \varepsilon/2) + \limsup_{n \rightarrow \infty} \Pr_{\mathcal{G}_n}(|Y_n^{(m)} - W_{mn}| > \varepsilon/2) \\ & \leq \frac{1}{4} \varepsilon^{-2} \left[\exp\left(\sum_{i=1}^{\infty} \mu_i \delta_i^2\right) - \exp\left(\sum_{i=1}^m \mu_i \delta_i^2\right) \right] + 0, \end{aligned}$$

which finishes the proof of (153). \square

A.2. Application of the Small Subgraph Conditioning Method

The application of Theorem A.1, and similarly Lemma A.2, requires a verification of its assumptions. This check is routine for the most part, but it is nevertheless technically arduous, mainly because of assumption (A3), which requires precise calculation of the moments' asymptotics. We suppress the verification in the following lemma whose proof is given later in this section. The lemma includes some details on a few quantities which will be relevant for the proof of Lemma 6.12.

LEMMA A.3. *Let $G \sim \mathcal{G}_n^r$ and X_{i_n} be the number of cycles of even length i appearing in G , $i = 2, 4, \dots$. Let $S = \{(\mathbf{p}, \eta) \mid \mathbf{p} \in \mathcal{Q}, \eta : W \rightarrow [q]\}$ and for $s \in S$ with $s = (\mathbf{p}, \eta)$, set $Y_n^{(s)} = Z_G^{\mathbf{p}}(\eta)$. In the setting of Theorem 1.5, the assumptions of Theorem A.1 hold.*

Further, for $s \in S$ with $s = (\mathbf{p}, \eta)$ and all even $i \geq 2$, $\delta_i^{(s)}$ satisfies (i) $\delta_i^{(s)} > 0$, (ii) $\delta_i^{(s)}$ depends on \mathbf{p} but not on η , (iii) $\sum_i \mu_i \delta_i^{(s)} < \infty$, (iv) if the phases are permutation symmetric, $\delta_i^{(s)}$ depends on the spin model but not on the particular phase \mathbf{p} .

Using Lemmas A.2 and A.3, we are ready to prove Lemma 6.12.

PROOF OF LEMMA 6.12. To see (104), note that the $W_{mn}^{(s)}$ of Lemma A.2 depend on the particular s only through the $\delta_i^{(s)}$'s. By Item (ii) of Lemma A.3, these depend only on \mathbf{p} in general and specifically for the permutation symmetric case, only on the spin model by Item (iv).

It remains to prove that $W_{mn}^{\mathbf{p}}$ are lower bounded uniformly in \mathbf{p} by a positive constant. Since the number of phases \mathbf{p} is bounded by a constant depending only on the spin model, it suffices to show that this is the case for a fixed phase \mathbf{p} . Using Item (i) of Lemma A.3 and that the random variables X_{i_n} are nonnegative integer valued, we have everywhere the bound

$$W_{mn}^{\mathbf{p}} = \prod_{i=1}^m (1 + \delta_i^{\mathbf{p}})^{X_{i_n}} \exp(-\mu_i \delta_i^{\mathbf{p}}) \geq \prod_{i=1}^m \exp(-\mu_i \delta_i^{\mathbf{p}}) > \prod_{i=1}^{\infty} \exp(-\mu_i \delta_i^{\mathbf{p}}).$$

Note that we have identified the $\delta_i^{(s)}$'s with the respective $\delta_i^{\mathbf{p}}$'s, this is justified by Item (ii) of Lemma A.3. The last quantity is finite and positive by Item (iii) in Lemma A.3. \square

We next prove Lemma A.3 which amounts to checking the validity of the assumptions (A1)-(A4) of Theorem A.1 for $Z_G^{\mathbf{p}}(\eta)$ for $\mathbf{p} \in \mathcal{Q}$ and $\eta : W \rightarrow [q]$.

Let us fix first some notation. Recall that a phase $\mathbf{p} \in \mathcal{Q}$ corresponds to a global maximum (α, β) of Ψ_1 . Let $\mathbf{x} = (x_{ij})_{i,j \in [q]}$ be as in Lemma 4.3, that is, the unique vector which maximizes $\Upsilon_1(\alpha, \beta, \mathbf{x})$ when α, β are fixed. In the setting of Theorem 1.5, we may assume that (α, β) is a Hessian local maximum of Ψ_1 . The following lemma puts together some relevant quantities and information which we derived in Section 4.2 in the course of proving Theorem 4.2.

LEMMA A.4. *Suppose that (α, β) is a Hessian local maximum of Ψ_1 . Define the vector $\mathbf{x} = (x_{ij})_{i,j \in [q]}$ as in Lemma 4.3.*

Let \mathbf{J} be the matrix $[\mathbf{L}_{\mathbf{r}}^0 \ \mathbf{0}; \ \mathbf{0} \ \mathbf{L}]$, where \mathbf{L} is the $q \times q$ matrix whose ij -entry is given by $x_{ij}/\sqrt{\alpha_i \beta_j}$. Then, the spectrum of \mathbf{J} is

$$\pm 1, \pm \lambda_1, \dots, \pm \lambda_{q-1},$$

for some positive λ_i which satisfy $\max_i \lambda_i < \frac{1}{\Delta-1}$. Relevant to Lemma A.3, observe that if the phases \mathbf{p} are permutation symmetric, then the λ_i 's are common for all phases.

Let $G \sim \mathcal{G}_n^r$ and $X_i := X_{i,n}$ be the number of cycles in G of even length i . Let $\mathbf{p} = (\alpha, \beta) \in \mathcal{Q}$. We next verify the assumptions of Theorem A.1 for the random variables $Z_G^{\mathbf{p}}(\eta)$, $\eta : W \rightarrow [q]$. We have the following lemmas:

LEMMA A.5 (LEMMA 7.3 IN MOSSEL ET AL. [2009]). *Assumption (A1) of Theorem A.1 holds for even i with*

$$\mu_i = \frac{r(\Delta, i)}{i} = \frac{(\Delta - 1)^i + (-1)^i(\Delta - 1)}{i},$$

where $r(\Delta, i)$ is the number of ways to properly edge color a cycle of length i with Δ colors.

The proof of Lemma A.5 is given in Mossel et al. [2009] and is omitted.

LEMMA A.6. *Let λ_j , $j \in [q - 1]$ be as in Lemma A.4. Then, for all even $i \geq 2$, it holds that*

$$\frac{\mathbf{E}_{\mathcal{G}_n^r}[Z_G^{\mathbf{p}}(\eta)X_i]}{\mathbf{E}_{\mathcal{G}_n^r}[Z_G^{\mathbf{p}}(\eta)]} \rightarrow \mu_i(1 + \delta_i) \text{ as } n \rightarrow \infty, \text{ where } \delta_i := \sum_{j=1}^{q-1} \lambda_j^i.$$

In particular, δ_i is positive for every even $i \geq 2$.

The proof of Lemma A.6 is given in Section A.3.

LEMMA A.7. *Let δ_i , $i = 2, 4, \dots$ be as in Lemma A.6. For every finite sequence m_1, \dots, m_k of nonnegative integers, it holds that*

$$\frac{\mathbf{E}_{\mathcal{G}_n^r}[Z_G^{\mathbf{p}}(\eta)[X_2]_{m_1} \dots [X_{2k}]_{m_k}]}{\mathbf{E}_{\mathcal{G}_n^r}[Z_G^{\mathbf{p}}]} \rightarrow \prod_{i=1}^k (\mu_i(1 + \delta_i))^{m_i} \text{ as } n \rightarrow \infty.$$

Once we give the proof of Lemma A.6, the proof of Lemma A.7 is identical to Mossel et al. [2009, Proof of Lemma 7.5] and is omitted.

LEMMA A.8. *In the notation and setting of Lemma A.4, it holds that*

$$\exp\left(\sum_{\text{even } i \geq 2} \mu_i \delta_i^2\right) = \prod_{i=1}^{q-1} \prod_{j=1}^{q-1} (1 - (\Delta - 1)^2 \lambda_i^2 \lambda_j^2)^{-1/2} \prod_{i=1}^{q-1} \prod_{j=1}^{q-1} (1 - \lambda_i^2 \lambda_j^2)^{-(\Delta-1)/2}.$$

Moreover, $\sum_{\text{even } i \geq 2} \mu_i \delta_i < \infty$.

The proof of Lemma A.8 is given in Section A.3.

Finally, we find the asymptotics of the second moment over the first moment squared.

LEMMA A.9. *In the notation and setting of Lemma A.4, it holds that*

$$\lim_{n \rightarrow \infty} \frac{\mathbf{E}_{\mathcal{G}_n^r}[(Z_G^{\mathbf{p}}(\eta))^2]}{(\mathbf{E}_{\mathcal{G}_n^r}[Z_G^{\mathbf{p}}(\eta)])^2} = C, \text{ where } C := \prod_{i=1}^{q-1} \prod_{j=1}^{q-1} (1 - (\Delta - 1)^2 \lambda_i^2 \lambda_j^2)^{-1/2} \prod_{i=1}^{q-1} \prod_{j=1}^{q-1} (1 - \lambda_i^2 \lambda_j^2)^{-(\Delta-1)/2}.$$

The proof of Lemma A.9 is quite extensive. In Appendix B, we first reduce the lemma to the case $r = 0$. This part of the proof is standard and is analogous to the proof of Lemma 6.11. Then, we compute the asymptotics in terms of determinants of relevant Hessian matrices. These determinants are computed in Appendix B.1.2, where also the proof of Lemma A.9 is completed.

With Lemmas A.5-A.9 at hand, the proof of Lemma A.3 is immediate.

PROOF OF LEMMA A.3. We prove the first part of the lemma by verifying the assumptions of Theorem A.1. Lemma A.5 verifies assumption (A1), Lemma A.7 verifies assumption (A2) and Lemmas A.8 and A.9 verify assumptions (A3) and (A4). This proves the first part of the lemma.

For the second part, just use the second parts in Lemmas A.4, A.6, A.8 to establish Items (i)-(iv). \square

A.3. Proofs of Lemmas A.6 and A.8

In this section, we give the proofs of Lemmas A.6 and A.8.

PROOF OF LEMMA A.6. The proof is close to [Mossel et al. 2009, Proof of Lemma 7.4], the approach only needs a few modifications to account for the q -spin setting. We make the minor notation change from X_i to X_ℓ .

We will do the computations for the case of $G \sim \mathcal{G}_n$ and the random variables $Z_G^{\alpha,\beta}$, the extension to the case $G \sim \mathcal{G}_n^r$ and the random variables $Z_G^{\mathbf{p}}(\eta)$ has slightly more complicated expressions, but otherwise the derivation is completely analogous (this was first displayed by Sly [2010]). We give the details of the more complicated calculations in Section D.2 and here we focus on the main ideas in the setting previously specified. (All probabilities and expectations in the following argument will refer to the graph distribution \mathcal{G}_n .)

We will show that

$$\frac{\mathbf{E}[Z_G^{\alpha,\beta} X_\ell]}{\mathbf{E}[Z_G^{\alpha,\beta}]} \rightarrow \mu_\ell(1 + \delta_\ell) \text{ as } n \rightarrow \infty,$$

where μ_ℓ, δ_ℓ are as in the statement of the lemma.

Let $\mathcal{S} = \{S_1, \dots, S_q\}$ and $\mathcal{T} = \{T_1, \dots, T_q\}$ be partitions of V_1 and V_2 respectively such that $|S_i| = \alpha_i n$ and $|T_j| = \beta_j n$ for all $i, j \in [q]$. Denote by $Y_{\mathcal{S},\mathcal{T}}$ the weight of the configuration σ that \mathcal{S}, \mathcal{T} induce, that is, for a vertex $v \in V_1$, $\sigma(v) = i$ iff $v \in S_i$ and similarly for vertices in V_2 .

Fix a specific pair of \mathcal{S}, \mathcal{T} . By symmetry,

$$\frac{\mathbf{E}[Z_G^{\alpha,\beta} X_\ell]}{\mathbf{E}[Z_G^{\alpha,\beta}]} = \frac{\mathbf{E}[Y_{\mathcal{S},\mathcal{T}} X_\ell]}{\mathbf{E}[Y_{\mathcal{S},\mathcal{T}}]}. \quad (156)$$

We now decompose X_ℓ as follows.

— ξ will denote a proper Δ -edge colored, rooted, and oriented ℓ -cycle ($r(\Delta, \ell)$ possibilities), in which the vertices are colored with $\{Y_1, \dots, Y_q, G_1, \dots, G_q\}$ and edges are colored with $\{1, \dots, \Delta\}$.

A vertex colored with Y_i (respectively, G_i) for some $i \in [q]$ will be loosely called yellow (respectively, green) and signifies that the vertex belongs to S_i (respectively, T_i). Since a yellow vertex belongs to V_1 , and a green vertex belongs to V_2 , a vertex coloring is consistent with the bipartiteness of the random graph if adjacent vertices of the cycle are not both yellow or green, that is, the vertex assignments which are prohibited for neighboring vertices in the cycle are (Y_i, Y_j) and (G_i, G_j) , $\forall (i, j) \in [q]^2$.

Note here that we do not explicitly prohibit assignments (Y_i, G_j) in the presence of a hard constraint $B_{ij} = 0$; this will be accounted otherwise. The color of the edges will prescribe which of the Δ perfect matchings an edge of a (potential) cycle will belong to.

—Given ξ , we use ζ to denote the position that an ℓ -cycle can be, that is, the exact vertices it traverses in order, such that the prescription of the vertex colors of ξ is satisfied.

— $\mathbf{1}_{\xi,\zeta}$ is the indicator function whether a cycle specified by ξ, ζ is present in G .

Note that each possible cycle corresponds to exactly 2ℓ different configurations ξ (the number of ways to root and orient the cycle). For each of those ξ , the respective sets of configurations ζ are the same. Hence, we may write

$$X_\ell = \frac{1}{2\ell} \sum_{\xi} \sum_{\zeta} \mathbf{1}_{\xi, \zeta}. \quad (157)$$

Let $p_1 := \Pr[\mathbf{1}_{\xi, \zeta} = 1]$. We have

$$\mathbf{E}[Y_{S, \mathcal{T}} X_\ell] = \frac{1}{2\ell} \sum_{\xi} \sum_{\zeta} p_1 \cdot \mathbf{E}[Y_{S, \mathcal{T}} | \mathbf{1}_{\xi, \zeta} = 1].$$

In light of (156), we need to study the ratio $\mathbf{E}[Y_{S, \mathcal{T}} | \mathbf{1}_{\xi, \zeta} = 1] / \mathbf{E}_G[Y_{S, \mathcal{T}}]$. At this point, to simplify notation, we may assume that ξ, ζ are fixed.

We have shown in Section 2 that

$$\mathbf{E}[Y_{S, \mathcal{T}}] = \left(\sum_{\mathbf{x}} \frac{\prod_i \binom{\alpha_i n}{x_{i1}n, \dots, x_{iq}n} \prod_j \binom{\beta_j n}{x_{1j}n, \dots, x_{qj}n}}{\binom{n}{x_{11}n, \dots, x_{qq}n}} \prod_{i,j} B_{ij}^{nx_{ij}} \right)^\Delta, \quad (158)$$

where the variables $\mathbf{x} = (x_{11}, \dots, x_{qq})$ correspond to the number of edges between S, \mathcal{T} in one matching (precisely, nx_{ij} is the number of edges between the sets S_i and T_j).

For the calculation of $\mathbf{E}[Y_{S, \mathcal{T}} | \mathbf{1}_{\xi, \zeta} = 1]$, we need some notation. For colors $c_1, c_2 \in \{Y_1, \dots, Y_q, G_1, \dots, G_q\}$, we say that an edge is of type $\{c_1, c_2\}$ if its endpoints have colors c_1, c_2 . Let y_i, g_j denote the number of vertices colored with Y_i, G_j respectively. For $k = 1, \dots, \Delta$, let $a_{ij}(k)$ denote the number of edges of color k and type $\{Y_i, G_j\}$. Finally, for $i, j \in [q]$ let $a_{ij} = \sum_k a_{ij}(k)$. By considering the sum of the degrees of vertices colored Y_i , the sum of the degrees of vertices colored G_j and the total number of edges of the cycle, we obtain the following equalities.

$$\sum_j a_{ij} = 2y_i, \quad \sum_i a_{ij} = 2g_j, \quad \sum_{i,j} a_{ij} = 2\ell. \quad (159)$$

We are almost set to compute $\mathbf{E}[Y_{S, \mathcal{T}} | \mathbf{1}_{\xi, \zeta} = 1]$. We denote by \mathbf{x}_k the same set of variables as in (158) but for the k th matching. Namely, $nx_{i,j,k}$ is the number of edges between the sets S_i and T_j in the k th matching. This number includes the $a_{ij}(k)$ edges prescribed by ξ, ζ . To simplify the following formulas, let $nx'_{i,j,k} := nx_{i,j,k} - a_{ij}(k)$. We have

$$\mathbf{E}[Y_{S, \mathcal{T}} | \mathbf{1}_{\xi, \zeta} = 1] = \prod_{k=1}^{\Delta} \left[\sum_{\mathbf{x}_k} \frac{\prod_i \binom{\alpha_i n - \sum_j a_{ij}(k)}{nx'_{i1,k}, \dots, nx'_{iq,k}} \prod_j \binom{\beta_j n - \sum_i a_{ij}(k)}{nx'_{1j,k}, \dots, nx'_{qj,k}}}{\binom{n - \sum_{i,j} a_{ij}(k)}{x'_{11,k}n, \dots, x'_{qq,k}n}} \prod_{i,j} B_{ij}^{nx_{i,j,k}} \right].$$

In these sums, for any $\varepsilon > 0$ and all sufficiently large n , terms whose \mathbf{x}_k 's are ε -far from the optimal value of \mathbf{x} given in Lemma 4.3 have exponentially small contribution and may be ignored. Standard approximations of binomial coefficients, see, for example, Galanis et al. [2012, Lemma 28], give

$$\begin{aligned} \frac{\binom{\alpha_i n - \sum_j a_{ij}(k)}{x'_{i1,k}n, \dots, x'_{iq,k}n}}{\binom{\alpha_i n}{x_{i1,k}n, \dots, x_{iq,k}n}} &\sim \frac{\prod_j (x_{ij,k})^{a_{ij}(k)}}{\alpha_i^{\sum_j a_{ij}(k)}}, & \frac{\binom{\beta_j n - \sum_i a_{ij}(k)}{x'_{1j,k}n, \dots, x'_{qj,k}n}}{\binom{\beta_j n}{x_{1j,k}n, \dots, x_{qj,k}n}} &\sim \frac{\prod_i (x_{ij,k})^{a_{ij}(k)}}{\beta_j^{\sum_i a_{ij}(k)}}, \\ & & \frac{\binom{n - \sum_{i,j} a_{ij}(k)}{x'_{11,k}n, \dots, x'_{qq,k}n}}{\binom{n}{x_{11,k}n, \dots, x_{qq,k}n}} &\sim \prod_{i,j} (x_{ij,k})^{a_{ij}(k)}, \end{aligned}$$

where $f(n) \sim g(n)$ stands for $f(n)/g(n) \rightarrow 1$ as $n \rightarrow \infty$. Thus, we obtain

$$\frac{\mathbf{E}[Y_{S,T} | \mathbf{1}_{\xi,\zeta} = \mathbf{1}]}{\mathbf{E}[Y_{S,T}]} \sim \frac{\prod_{i,j} (x_{ij})^{a_{ij}}}{\prod_i \alpha_i^{\sum_j a_{ij}} \prod_j \beta_j^{\sum_i a_{ij}}}.$$

We have $p_1 \sim n^{-\ell}$ and for given ξ , the number of possible ζ is asymptotic to $n^\ell \prod_i \alpha_i^{y_i} \prod_j \beta_j^{g_j}$. Thus, for the given ξ , we have

$$\frac{\sum_\zeta p_1 \mathbf{E}[Y_{S,T} | \mathbf{1}_{\xi,\zeta} = \mathbf{1}]}{\mathbf{E}[Y_{S,T}]} \sim \frac{\prod_i \alpha_i^{y_i} \prod_j \beta_j^{g_j} \prod_{i,j} (x_{ij})^{a_{ij}}}{\prod_i \alpha_i^{\sum_j a_{ij}} \prod_j \beta_j^{\sum_i a_{ij}}} = \prod_{i,j} \left(\frac{x_{ij}}{\sqrt{\alpha_i \beta_j}} \right)^{a_{ij}}$$

Note that the rhs evaluates to 0 whenever there exist i, j such that $B_{ij} = 0$ but $a_{ij} \neq 0$, since then we have $x_{ij} = 0$. This is in complete accordance with the fact that the configuration induced by the partition $\{S, T\}$ has zero weight. Thus, by (156), we have

$$\frac{\mathbf{E}[Z_G^{\alpha,\beta} X_\ell]}{\mathbf{E}[Z_G^{\alpha,\beta}]} \sim \frac{r(\Delta, \ell)}{2^\ell} \cdot \sum_{\mathbf{a}} N_{\mathbf{a}} \prod_{i,j} \left(\frac{x_{ij}}{\sqrt{\alpha_i \beta_j}} \right)^{a_{ij}}, \quad (160)$$

where $\mathbf{a} = \{a_{11}, \dots, a_{qq}\}$ and $N_{\mathbf{a}}$ is the number of possible ξ with a_{ij} edges having assignment (Y_i, G_j) . To analyze this sum, we employ a technique given in Janson [1995]. The idea is to define a weighted transition matrix and view it as the (weighted) adjacency matrix of a weighted graph. The powers of the matrix count the (multiplicative) weight of walks in the graph and a closed walk in this graph will correspond to a specification ξ . By defining the weights appropriately, one can also ensure that each closed walk will correctly capture the weight of the specification ξ .

In our setting, the transition matrix is simply the matrix \mathbf{J} of Lemma A.4. The first q rows and q columns correspond to the colors Y_i and the remaining rows and columns to colors G_j . The total weight of closed walks of length ℓ is given by $\text{Tr}(\mathbf{J}^\ell)$. Using the description of the eigenvalues given in Lemma A.4, we obtain that for even ℓ , $\text{Tr}(\mathbf{J}^\ell) = 2(1 + \sum_{i=1}^{q-1} \lambda_i^\ell)$. This concludes the proof. \square

PROOF OF LEMMA A.8. Using Lemma A.5, we have

$$\sum_{\text{even } i \geq 2} \mu_i \delta_i^2 = \sum_{\text{even } i \geq 2} \frac{r(\Delta, i)}{i} \cdot \left(\sum_{j=1}^{q-1} \lambda_j^i \right)^2 = \sum_{\text{even } i \geq 2} \frac{(\Delta - 1)^i + (\Delta - 1)}{i} \cdot \left(\sum_{j=1}^{q-1} \sum_{j'=1}^{q-1} \lambda_j^i \lambda_{j'}^i \right).$$

Observe that $\sum_{j \geq 1} \frac{x^{2j}}{2^j} = -\frac{1}{2} \ln(1 - x^2)$ for all $|x| < 1$. By Lemma A.4, the λ_j 's satisfy $(\Delta - 1)\lambda_j < 1$ for all j , so that $(\Delta - 1)\lambda_j \lambda_{j'} < 1$ for all j, j' . It follows that

$$\sum_{\text{even } i \geq 2} \mu_i \delta_i^2 = -\frac{1}{2} \left(\sum_{i,j} \ln(1 - (\Delta - 1)^2 \lambda_i^2 \lambda_j^2) + (\Delta - 1) \sum_{i,j} \ln(1 - \lambda_i^2 \lambda_j^2) \right),$$

thus proving the first part of the lemma. The proof of $\sum_i \mu_i \delta_i < \infty$ is completely analogous. \square

B. MOMENT ASYMPTOTICS

In this section, we prove Lemma A.9. For the purposes of this section, we will identify \mathcal{Q} with the dominant phases of a random Δ -regular bipartite graph. Thus, we will use $\mathbf{p} \in \mathcal{Q}$ to denote a dominant phase (α, β) .

We first recall some relevant definitions from Section 6.4. For $r \geq 0$, let $G \sim \mathcal{G}_n^r$ and $\sigma : U \cup W \rightarrow [q]$ be a configuration on G . The *footprint* of σ is a pair of q -dimensional vectors $\alpha_\sigma, \beta_\sigma$ whose i th entries are equal to $|\sigma^{-1}(i) \cap U^+|/n, |\sigma^{-1}(i) \cap U^-|/n$, respectively. The *phase* $Y(\sigma)$ of σ is the dominant phase (α, β) , which is closest to $(\alpha_\sigma, \beta_\sigma)$, precisely:

$$Y(\sigma) = \arg \max_{\mathbf{p}=(\alpha, \beta) \in \mathcal{Q}} (\|\alpha - \alpha_\sigma\|_2^2 + \|\beta - \beta_\sigma\|_2^2)^{1/2}. \quad (161)$$

Finally, recall that for $\mathbf{p} \in \mathcal{Q}$, $Z_G^{\mathbf{p}}$ is the partition function “conditioned on the phase \mathbf{p} ”, that is, the contribution to the partition function of G from configurations σ with $Y(\sigma) = \mathbf{p}$, and for $\eta : W \rightarrow [q]$, $Z_G^{\mathbf{p}}(\eta)$ is the contribution to the partition function of G from configurations σ with $Y(\sigma) = \mathbf{p}$ and $\sigma_W = \eta$, see (100) for more details.

To prove Lemma A.9, for $\mathbf{p} \in \mathcal{Q}$ and a configuration $\eta : W \rightarrow [q]$, we need to compute the asymptotics of $\mathbf{E}_{\mathcal{G}_n^r}[(Z_G^{\mathbf{p}}(\eta))^2]/(\mathbf{E}_{\mathcal{G}_n^r}[Z_G^{\mathbf{p}}(\eta)])^2$. The following lemma reduces the computation to the case $r = 0$. Note that for $r = 0$, the set of vertices W is empty and the distribution \mathcal{G}_n^r coincides with the distribution $\mathcal{G} := \mathcal{G}_n$ on random Δ -regular bipartite graphs from Section 2.

LEMMA B.1. *Let $\mathbf{p} = (\alpha, \beta) \in \mathcal{Q}$ be a Hessian dominant phase. Then, for every fixed $r > 0$, for every $\eta : W \rightarrow [q]$ it holds that*

$$\lim_{n \rightarrow \infty} \frac{\mathbf{E}_{\mathcal{G}_n^r}[(Z_G^{\mathbf{p}}(\eta))^2]}{(\mathbf{E}_{\mathcal{G}_n^r}[Z_G^{\mathbf{p}}(\eta)])^2} = \lim_{n \rightarrow \infty} \frac{\mathbf{E}_{\mathcal{G}_n}[(Z_G^{\mathbf{p}})^2]}{(\mathbf{E}_{\mathcal{G}_n}[Z_G^{\mathbf{p}}])^2}.$$

We will prove Lemma B.1 in Appendix D.1. In light of Lemma B.1, we need to compute the limiting ratio of $\mathbf{E}_{\mathcal{G}_n}[(Z_G^{\mathbf{p}})^2]/(\mathbf{E}_{\mathcal{G}_n}[Z_G^{\mathbf{p}}])^2$ for $\mathbf{p} \in \mathcal{Q}$. We do this by computing separately the asymptotics of $\mathbf{E}_{\mathcal{G}_n}[Z_G^{\mathbf{p}}]$ and $\mathbf{E}_{\mathcal{G}_n}[(Z_G^{\mathbf{p}})^2]$. We begin with an observation that will allow us to deduce the asymptotics of $\mathbf{E}_{\mathcal{G}}[(Z_G^{\mathbf{p}})^2]$ from the asymptotics of $\mathbf{E}_{\mathcal{G}}[Z_G^{\mathbf{p}}]$ applied to the spin system with interaction matrix $\mathbf{B} \otimes \mathbf{B}$.

LEMMA B.2. *Let $\mathbf{p} = (\alpha, \beta) \in \mathcal{Q}$ be a dominant phase for the spin system with interaction matrix \mathbf{B} . Then, $\mathbf{p}' = (\alpha \otimes \alpha, \beta \otimes \beta)$ is a dominant phase for the spin system with interaction matrix $\mathbf{B} \otimes \mathbf{B}$.*

For $G \sim \mathcal{G}_n$, let $Z_G^{\mathbf{p}'}$ be the contribution to the partition function of G for the spin system with interaction matrix $\mathbf{B} \otimes \mathbf{B}$ from configurations σ with $Y(\sigma) = \mathbf{p}'$. Then, $\lim_{n \rightarrow \infty} \frac{\mathbf{E}_{\mathcal{G}_n}[(Z_G^{\mathbf{p}'})^2]}{\mathbf{E}_{\mathcal{G}_n}[Z_G^{\mathbf{p}'}]} = 1$.

PROOF OF LEMMA B.2. The first part of the lemma is an immediate consequence of Lemma 3.2. In fact, the proof of Lemma 3.2 shows the stronger fact that the set of dominant phases for the spin system with interaction matrix $\mathbf{B} \otimes \mathbf{B}$ is given by $\mathcal{Q}^{\otimes 2} := \{(\alpha \otimes \alpha', \beta \otimes \beta') \mid (\alpha, \beta), (\alpha', \beta') \in \mathcal{Q}\}$.

For the second part, let

$$\Sigma_1 = \left\{ (\alpha', \beta') \mid \alpha', \beta' \in \Delta_q, \mathbf{p} = \arg \min_{\mathbf{p}^*=(\alpha^*, \beta^*) \in \mathcal{Q}} (\|\alpha' - \alpha^*\|^2 + \|\beta' - \beta^*\|^2)^{1/2} \right\}, \quad (162)$$

so that

$$Z_G^{\mathbf{p}} = \sum_{(\alpha', \beta') \in \Sigma_1} Z_G^{\alpha', \beta'} \quad \text{and} \quad (Z_G^{\mathbf{p}})^2 = \sum_{(\alpha', \beta'), (\alpha'', \beta'') \in \Sigma_1} Z_G^{\alpha', \beta'} Z_G^{\alpha'', \beta''}.$$

It follows that $\mathbf{E}_G[(Z_G^{\mathbf{p}})^2]$ is given by the sum in the right-hand side in (6), but now the sum is over $\boldsymbol{\gamma}, \boldsymbol{\delta}$ which satisfy

$$\begin{aligned} \sum_k \gamma_{ik} &= \alpha'_i \quad (\forall i \in [q]), & \sum_l \delta_{jl} &= \beta'_j \quad (\forall j \in [q]), \\ \sum_i \gamma_{ik} &= \alpha''_k \quad (\forall k \in [q]), & \sum_j \delta_{jl} &= \beta''_l \quad (\forall l \in [q]), \end{aligned}$$

and $(\alpha', \beta'), (\alpha'', \beta'')$ range over Σ_1 . Note that by the definition of Σ_1 , $\mathbf{p} = (\alpha, \beta)$ is the unique dominant phase (of the spin system with interaction matrix \mathbf{B}) contained in Σ_1 . By Lemma 3.2, for any $\varepsilon > 0$ and all sufficiently large n , terms in the sum with $(\|\boldsymbol{\gamma} - \alpha \otimes \alpha\|_2^2 + \|\boldsymbol{\delta} - \beta \otimes \beta\|_2^2)^{1/2} \geq \varepsilon$ have exponentially small contribution and hence may be ignored. Similarly, for the spin system with interaction matrix $\mathbf{B} \otimes \mathbf{B}$, $\mathbf{E}_G[Z_G^{\mathbf{p}'}]$ is given by the sum in the right-hand side in (6), where now the sum is over $\boldsymbol{\gamma}, \boldsymbol{\delta}$ with $(\boldsymbol{\gamma}, \boldsymbol{\delta}) \in \Sigma_2$, where

$$\Sigma_2 := \left\{ (\boldsymbol{\gamma}', \boldsymbol{\delta}') \mid \boldsymbol{\gamma}', \boldsymbol{\delta}' \in \Delta_{q^2}, \mathbf{p}' = \arg \min_{\mathbf{p}^* = (\boldsymbol{\gamma}^*, \boldsymbol{\delta}^*) \in Q^{\otimes 2}} (\|\boldsymbol{\gamma}' - \boldsymbol{\gamma}^*\|^2 + \|\boldsymbol{\delta}' - \boldsymbol{\delta}^*\|^2)^{1/2} \right\}.$$

Once again, by Lemma 3.2, for any $\varepsilon > 0$ and all sufficiently large n , terms in the sum with $(\|\boldsymbol{\gamma} - \alpha \otimes \alpha\|_2^2 + \|\boldsymbol{\delta} - \beta \otimes \beta\|_2^2)^{1/2} \geq \varepsilon$ have exponentially small contribution and hence may be ignored. It follows that for all sufficiently small $\varepsilon > 0$, the remaining terms in the two sums are identical which completes the proof. \square

As a consequence of Lemma B.2, we may focus on the asymptotics of the first moment $\mathbf{E}_G[Z_G^{\mathbf{p}}]$. Let

$$P_1 = \{(i, j) \in [q]^2 \mid B_{ij} > 0\}. \quad (163)$$

In the presence of a hard constraint $B_{ij} = 0$, edge assignments (i, j) yield a zero-weight configuration. In the maximization of Υ_1 (see (8)), the hard constraint $B_{ij} = 0$ was not directly relevant, since for $x_{ij} > 0$ the function Υ_1 evaluates to $-\infty$. Indeed, we found that the optimal x_{ij} is of the form $B_{ij}R_iC_j$ and hence zero. However, the asymptotics of $\mathbf{E}_G[Z_G^{\mathbf{p}}]$ include products of the optimal values of the x_{ij} and to correctly capture them, we need to explicitly account for the zero entries of the matrix \mathbf{B} .

To do so, in the formulation (5), we hard-code $x_{ij} = 0$ for a pair $(i, j) \notin P_1$ and hence the variables $\alpha, \beta, \mathbf{x}$ are restricted to the space

$$\begin{aligned} \sum_i \alpha_i &= 1, & \sum_j \beta_j &= 1, \\ \sum_j x_{ij} &= \alpha_i \quad (\forall i \in [q]), & \sum_i x_{ij} &= \beta_j \quad (\forall j \in [q]), \\ x_{ij} &= 0 \quad (\forall (i, j) \in [q]^2 \setminus P_1), & x_{ij} &\geq 0 \quad (\forall (i, j) \in P_1). \end{aligned} \quad (164)$$

We will also need to have a set of affinely independent variables which describe the polytope (164). Note that the dimension of the polytope (164) is $(q^2 + 2q) - (2q + 1) - (q^2 - |P_1|) = |P_1| - 1$. To get affinely independent variables $\alpha, \beta, \mathbf{x}$, we use the equalities in (164) and substitute an appropriate set of $(q + 1)^2 - |P_1|$ variables. We will not need to understand these substitutions till Appendix B.1.1, yet in the integrations which follow it is preferable to have integration variables rather than integrate over subspaces.

After this process, we are going to have $|P_1| - 1$ variables lying in a full dimensional space. We refer to this set of variables as the full dimensional representation of (164). For simplicity, we will still use $\alpha, \beta, \mathbf{x}$ for these variables and refer, for example, to x_{ij} even if x_{ij} is not in the full dimensional representation of (164), under the understanding that this is just a shorthand for the substituted expression. Using these conventions, we may view $\Upsilon_1(\alpha, \beta, \mathbf{x})$ as a function of the full dimensional representation of (164), and we will refer to this setup as the full dimensional representation of Υ_1 .

The following lemma expresses the asymptotics of $\mathbf{E}_G[Z_G^{\mathbf{p}}]$ in terms of determinants of suitable matrices corresponding to the Hessian matrix of the function Υ_1 evaluated at a global maximum $(\alpha^*, \beta^*, \mathbf{x}^*)$ (the relevant maximum is specified by the phase \mathbf{p}). To prove the lemma, we will need to perform integrations in a full dimensional space and hence we will work with the full dimensional representation of Υ_1 . Roughly, the asymptotics of $\mathbf{E}_G[Z_G^{\mathbf{p}}]$ are obtained by a Gaussian integration first with respect to the variables \mathbf{x} (with α, β fixed) and then with respect to the variables α, β . The Hessian matrix of the full dimensional representation of Υ_1 will thus be relevant, denoted by \mathbf{H}_1^f (for convenience, we scale \mathbf{H}_1^f by $1/\Delta$), as well as the square submatrix of \mathbf{H}_1^f corresponding to the variables \mathbf{x} , denoted by $\mathbf{H}_{1,\mathbf{x}}^f$. Note, by our earlier discussion, \mathbf{H}_1^f has dimension $(|P_1|-1) \times (|P_1|-1)$ while $\mathbf{H}_{1,\mathbf{x}}^f$ has dimension $(|P_1|-(2q-1)) \times (|P_1|-(2q-1))$. A more thorough exposition of these Hessian matrices is given in Section B.1.1.

LEMMA B.3. *Let $\mathbf{p} = (\alpha^*, \beta^*)$ be a Hessian dominant phase, that is, (α^*, β^*) maximizes $\Upsilon_1(\alpha, \beta)$. Let \mathbf{x}^* be the (unique) maximizer of $\Upsilon_1(\alpha^*, \beta^*, \mathbf{x})$ (given in Lemma 4.3). Denote by \mathbf{H}_1^f be the Hessian of the full dimensional representation of $\Upsilon_1(\alpha, \beta, \mathbf{x})$ scaled by $1/\Delta$ (evaluated at $\alpha^*, \beta^*, \mathbf{x}^*$) and by $\mathbf{H}_{1,\mathbf{x}}^f$ the square submatrix of \mathbf{H}_1^f corresponding to rows and columns indexed by \mathbf{x} . Then*

$$\lim_{n \rightarrow \infty} \frac{\mathbf{E}_G[Z_G^{\mathbf{p}}]}{e^{n\Upsilon_1(\alpha^*, \beta^*, \mathbf{x}^*)}} = \frac{\left(\prod_i \alpha_i^* \prod_j \beta_j^*\right)^{(\Delta-1)/2} \left(\prod_{(i,j) \in P_1} x_{ij}^*\right)^{-\Delta/2}}{\Delta^{q-1} (\text{Det}(-\mathbf{H}_1^f))^{1/2} (\text{Det}(-\mathbf{H}_{1,\mathbf{x}}^f))^{(\Delta-1)/2}}.$$

The computation of the determinants in Lemma B.3 is given in Section B.1.2, where also the proof of Lemma A.9 is completed.

PROOF OF LEMMA B.3. We assume that $\alpha, \beta, \mathbf{x}$ is a full dimensional representation of (164). We denote by $(\alpha^*, \beta^*, \mathbf{x}^*)$ the optimal vector which maximizes the full dimensional representation of $\Upsilon_1(\alpha, \beta, \mathbf{x})$. We have that $\alpha_i^*, \beta_j^* > 0$ for all i, j and $x_{ij}^* > 0$ for $(i, j) \in P_1$. Pick δ sufficiently small such that

$$\|(\alpha, \beta, \mathbf{x}) - (\alpha^*, \beta^*, \mathbf{x}^*)\|_2 \leq \delta \text{ implies } \alpha_i, \beta_j > 0 \text{ for all } i, j \text{ and } x_{ij} > 0 \text{ for } (i, j) \in P_1.$$

Since Υ_1 has the unique global maximum $(\alpha^*, \beta^*, \mathbf{x}^*)$ at the intersection of the spaces (162) and (164), standard compactness arguments imply that there exists $\varepsilon(\delta) > 0$ such that $\|(\alpha, \beta, \mathbf{x}) - (\alpha^*, \beta^*, \mathbf{x}^*)\|_2 \geq \delta$ implies $\Upsilon_1(\alpha^*, \beta^*, \mathbf{x}^*) - \Upsilon_1(\alpha, \beta, \mathbf{x}) \geq \varepsilon$. It follows that the contribution of terms with $\|(\alpha, \beta, \mathbf{x}) - (\alpha^*, \beta^*, \mathbf{x}^*)\|_2 \geq \delta$ to $\mathbf{E}_G[Z_G^{\mathbf{p}}]$ is exponentially small and may be ignored. Hence, we may restrict our attention to $\alpha, \beta, \mathbf{x}$ satisfying $\|(\alpha, \beta, \mathbf{x}) - (\alpha^*, \beta^*, \mathbf{x}^*)\|_2 < \delta$. Moreover, using Taylor's expansion, and since (α^*, β^*) is a Hessian dominant phase, we may choose δ small enough such that Υ_1 decays quadratically in a δ -ball around $(\alpha^*, \beta^*, \mathbf{x}^*)$.

Utilizing the choice of δ and Stirling's approximation for factorials, we obtain from (4) and (8) that

$$\frac{\mathbf{E}_G[Z_G^{\mathbf{p}}]}{e^{n\Upsilon_1(\alpha^*, \beta^*, \mathbf{x}^*)}} = (1 + O(n^{-1})) \sum_{\alpha, \beta} \left(\frac{1}{\sqrt{2\pi n}}\right)^{2(q-1)} \left(\prod_i \alpha_i \prod_j \beta_j\right)^{(\Delta-1)/2} \left[\sum_{\mathbf{x}} \left(\frac{1}{\sqrt{2\pi n}}\right)^{|P_1|-(2q-1)} \left(\prod_{(i,j) \in P_1} \frac{1}{\sqrt{x_{ij}}}\right) e^{n(\Upsilon_1(\alpha, \beta, \mathbf{x}) - \Upsilon_1(\alpha^*, \beta^*, \mathbf{x}^*)/\Delta)} \right]^\Delta.$$

We now compute

$$L := \lim_{n \rightarrow \infty} \frac{\mathbf{E}_G[Z_G^p]}{e^{n\Upsilon_1(\boldsymbol{\alpha}^*, \boldsymbol{\beta}^*, \mathbf{x}^*)}}.$$

Standard techniques of rewriting sums as integrals and an application of the dominated convergence theorem which utilizes the quadratic decay of Υ_1 around $\boldsymbol{\alpha}^*$, $\boldsymbol{\beta}^*$, \mathbf{x}^* (for details of this argument, see, e.g., Janson et al. [2000, Section 9.4]) ultimately give

$$L = \left(\prod_i \alpha_i^* \prod_j \beta_j^* \right)^{(\Delta-1)/2} \left(\prod_{(i,j) \in P_1} x_{ij}^* \right)^{-\Delta/2} \quad (165)$$

$$\left(\frac{1}{\sqrt{2\pi}} \right)^{2(q-1)} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \left[\left(\frac{1}{\sqrt{2\pi}} \right)^{|P_1|-(2q-1)} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} e^{\frac{1}{2}(\boldsymbol{\alpha}, \boldsymbol{\beta}, \mathbf{x}) \cdot \mathbf{H}(\boldsymbol{\alpha}, \boldsymbol{\beta}, \mathbf{x})^\top} d\mathbf{x}} \right]^\Delta d\boldsymbol{\alpha} d\boldsymbol{\beta},$$

where $\mathbf{H} = \mathbf{H}_1^f$ denotes the Hessian matrix of the (full dimensional representation of) Υ_1 evaluated at $(\boldsymbol{\alpha}^*, \boldsymbol{\beta}^*, \mathbf{x}^*)$ scaled by $1/\Delta$ and the operator \cdot stands for matrix multiplication.

We thus focus on computing the integral in (165). We begin with the inner integration. Let

$$I_1 = \left[\left(\frac{1}{\sqrt{2\pi}} \right)^{|P_1|-(2q-1)} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} e^{\frac{1}{2}(\boldsymbol{\alpha}, \boldsymbol{\beta}, \mathbf{x}) \cdot \mathbf{H}(\boldsymbol{\alpha}, \boldsymbol{\beta}, \mathbf{x})^\top} d\mathbf{x}} \right]^\Delta.$$

To calculate I_1 , we first decompose the exponent to isolate the terms involving \mathbf{x} . We obtain

$$\frac{1}{2}(\boldsymbol{\alpha}, \boldsymbol{\beta}, \mathbf{x}) \cdot \mathbf{H}(\boldsymbol{\alpha}, \boldsymbol{\beta}, \mathbf{x})^\top = \frac{1}{2}(\boldsymbol{\alpha}, \boldsymbol{\beta}) \cdot \mathbf{H}_{\boldsymbol{\alpha}, \boldsymbol{\beta}} \cdot (\boldsymbol{\alpha}, \boldsymbol{\beta})^\top - \frac{1}{2}\mathbf{x} \cdot (-\mathbf{H}_{\mathbf{x}}) \cdot \mathbf{x}^\top + \mathbf{T} \cdot \mathbf{x}^\top,$$

where $\mathbf{H} = \begin{bmatrix} \mathbf{H}_{\boldsymbol{\alpha}, \boldsymbol{\beta}} & \mathbf{H}_{\boldsymbol{\alpha}, \boldsymbol{\beta}, \mathbf{x}} \\ \mathbf{H}_{\boldsymbol{\alpha}, \boldsymbol{\beta}, \mathbf{x}}^\top & \mathbf{H}_{\mathbf{x}} \end{bmatrix}$ and $\mathbf{T} = (\boldsymbol{\alpha}, \boldsymbol{\beta}) \cdot \mathbf{H}_{\boldsymbol{\alpha}, \boldsymbol{\beta}, \mathbf{x}}$. Specifically:

- $\mathbf{H}_{\boldsymbol{\alpha}, \boldsymbol{\beta}}$ is the square submatrix of \mathbf{H} corresponding to the rows indexed by $\boldsymbol{\alpha}, \boldsymbol{\beta}$ and the columns indexed by $\boldsymbol{\alpha}, \boldsymbol{\beta}$;
- $\mathbf{H}_{\mathbf{x}}$ is the square submatrix of \mathbf{H} corresponding to the rows indexed by \mathbf{x} and the columns indexed by \mathbf{x} ;
- $\mathbf{H}_{\boldsymbol{\alpha}, \boldsymbol{\beta}, \mathbf{x}}$ is the submatrix of \mathbf{H} corresponding to the rows indexed by $\boldsymbol{\alpha}, \boldsymbol{\beta}$ and the columns indexed by \mathbf{x} .

Note that $\mathbf{H}_{\mathbf{x}}$ is the Hessian of $g_1(\mathbf{x})$ evaluated at \mathbf{x}^* . Since $g_1(\mathbf{x})$ is concave, we have that $\mathbf{H}_{\mathbf{x}}$ is negative definite. Utilizing this decomposition, we obtain

$$I_1 = e^{\frac{\Delta}{2}(\boldsymbol{\alpha}, \boldsymbol{\beta}) \cdot \mathbf{H}_{\boldsymbol{\alpha}, \boldsymbol{\beta}} \cdot (\boldsymbol{\alpha}, \boldsymbol{\beta})^\top} \left[\left(\frac{1}{\sqrt{2\pi}} \right)^{|P_1|-(2q-1)} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} e^{-\frac{1}{2}\mathbf{x} \cdot (-\mathbf{H}_{\mathbf{x}}) \cdot \mathbf{x}^\top + \mathbf{T} \cdot \mathbf{x}^\top} d\mathbf{x}} \right]^\Delta$$

$$= \frac{1}{(\text{Det}(-\mathbf{H}_{\mathbf{x}}))^{\Delta/2}} e^{\frac{\Delta}{2}(\mathbf{T} \cdot (-\mathbf{H}_{\mathbf{x}})^{-1} \cdot \mathbf{T}^\top + (\boldsymbol{\alpha}, \boldsymbol{\beta}) \cdot \mathbf{H}_{\boldsymbol{\alpha}, \boldsymbol{\beta}} \cdot (\boldsymbol{\alpha}, \boldsymbol{\beta})^\top)},$$

where the Gaussian integral with the linear term can be reduced to a standard Gaussian integral (without a linear term) by completing the square (or rather the quadratic form) in the exponent utilizing the Cholesky decomposition of the positive definite matrix $-\mathbf{H}_{\mathbf{x}}$.

We are left with the task of computing the integral

$$I_2 = \left(\frac{1}{\sqrt{2\pi}} \right)^{2(q-1)} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} e^{\frac{\Delta}{2}(\mathbf{T} \cdot (-\mathbf{H}_{\mathbf{x}})^{-1} \cdot \mathbf{T}^\top + (\boldsymbol{\alpha}, \boldsymbol{\beta}) \cdot \mathbf{H}_{\boldsymbol{\alpha}, \boldsymbol{\beta}} \cdot (\boldsymbol{\alpha}, \boldsymbol{\beta})^\top)} d\boldsymbol{\alpha} d\boldsymbol{\beta}. \quad (166)$$

Using the definition of \mathbf{T} , we have

$$\mathbf{T} \cdot (-\mathbf{H}_x)^{-1} \cdot \mathbf{T}^\top + (\alpha, \beta) \cdot \mathbf{H}_{\alpha, \beta} \cdot (\alpha, \beta)^\top = (\alpha, \beta) \cdot (\mathbf{H}_{\alpha, \beta} - \mathbf{H}_{\alpha, \beta, x} \cdot \mathbf{H}_x^{-1} \cdot \mathbf{H}_{\alpha, \beta, x}^\top) \cdot (\alpha, \beta)^\top.$$

The matrix $\mathbf{M} = \mathbf{H}_{\alpha, \beta} - \mathbf{H}_{\alpha, \beta, x} \cdot \mathbf{H}_x^{-1} \cdot \mathbf{H}_{\alpha, \beta, x}^\top$ is the Schur complement of the block \mathbf{H}_x of \mathbf{H} . We can thus conclude that \mathbf{M} is negative definite (since \mathbf{H} is negative definite). Further, we may also conclude that $\text{Det}(\mathbf{H}) = \text{Det}(\mathbf{H}_x) \text{Det}(\mathbf{M})$. A Gaussian integration then yields

$$I_2 = \left(\frac{1}{\Delta^{2(q-1)} \text{Det}(-\mathbf{M})} \right)^{1/2} = \left(\frac{\text{Det}(-\mathbf{H}_x)}{\Delta^{2(q-1)} \text{Det}(-\mathbf{H})} \right)^{1/2}. \quad (167)$$

Combining Eqs. (165), (166), and (167), we obtain the statement of the lemma. \square

B.1. The Determinants

This section addresses the computation of the determinants of the Hessians in Lemma B.3. The calculations are quite complex since one has to make a choice of free variables, do the substitutions, differentiate, and then hope that the structure of the problem will prevail in the determinants. Pushing this procedure in our setting leads to complications since the choice of free variables takes away much of the combinatorial structure of the problem. We follow a different path, which amongst other things, reveals that the determinants, via the matrix-tree theorem, correspond to counting weighted trees in appropriate graphs.

The proof has two parts. The first part connects different formulations of the Hessian of a constrained maximization in an abstract setting. Essentially, this puts together well known concepts from optimization in a way that will allow to stay as close as possible to the combinatorial structure of the determinants. The second part specialises the work of the first part to compute the required determinants and is unavoidably more computational.

B.1.1. Hessian Formulations for Constrained Problems. The setting of this section is the following: we are given Υ , a function of $\mathbf{z} \in \mathbb{R}^n$, subject to the linear constraints $\mathbf{A}\mathbf{z} = \mathbf{b}$, where $\mathbf{A} \in \mathbb{R}^{m \times n}$. The assumption of linear constraints stems from the setting of Lemma B.3, yet the arguments extend to other constraints as well by considering gradients of these constraints at the point \mathbf{z}_0 and implicit functions. Without loss of generality, we will also assume that $\mathbf{b} = \mathbf{0}$.

We are interested in the Hessian \mathbf{H}^ℓ of a full dimensional representation of Υ . A full dimensional representation of Υ consists essentially of substituting an appropriate subset of the variables \mathbf{z} using the constraints $\mathbf{A}\mathbf{z} = \mathbf{0}$. Note that the representation is not as much tied to Υ as it is tied to the space $\mathbf{A}\mathbf{z} = \mathbf{0}$. Specifically, assume that the row rank of \mathbf{A} is r . In all the relevant constrained functions we consider, the constraints are not linearly independent so such an assumption is necessary. A full dimensional representation of Υ is specified by two submatrices of \mathbf{A} denoted by $(\mathbf{A}_f, \mathbf{A}_{fs})$. The matrix \mathbf{A}_f is a submatrix of \mathbf{A} consisting of r linearly independent rows of \mathbf{A} , so that $\mathbf{A}\mathbf{z} = \mathbf{0}$ iff $\mathbf{A}_f \mathbf{z} = \mathbf{0}$. Then, \mathbf{A}_{fs} is an $r \times r$ submatrix of \mathbf{A}_f which is invertible. The variables corresponding to columns of \mathbf{A}_{fs} are denoted by \mathbf{z}_s . The remaining variables \mathbf{z}_f are called free and \mathbf{A}_{ff} is the submatrix of \mathbf{A}_f induced by the columns indexed by \mathbf{z}_f . Renaming if needed, the equation $\mathbf{A}_f \mathbf{z} = \mathbf{0}$ may be naturally decomposed as

$$[\mathbf{A}_{ff} \ \mathbf{A}_{fs}] \begin{bmatrix} \mathbf{z}_f \\ \mathbf{z}_s \end{bmatrix} = \mathbf{0}, \text{ so that } \mathbf{z} = \begin{bmatrix} \mathbf{z}_f \\ \mathbf{z}_s \end{bmatrix} = \begin{bmatrix} \mathbf{I} \\ -(\mathbf{A}_{fs})^{-1} \mathbf{A}_{ff} \end{bmatrix} \mathbf{z}_f.$$

Thus, we can now think of Υ as a function which is completely determined by the variables \mathbf{z}_f which, in contrast with the variables \mathbf{z} , span a full dimensional space.

Denote by \mathbf{H} the unconstrained Hessian of Υ with respect to the variables \mathbf{z} and by \mathbf{H}^f the Hessian of the full dimensional representation of Υ with respect to the variables \mathbf{z}_f . The Hessians \mathbf{H} , \mathbf{H}^f are connected by the following equation, which follows by straightforward matrix calculus and its proof is omitted.

$$\mathbf{H}^f = \mathbf{S}^\top \mathbf{H} \mathbf{S}, \text{ where } \mathbf{S} = \begin{bmatrix} \mathbf{I} \\ -(\mathbf{A}_{fs})^{-1} \mathbf{A}_{ff} \end{bmatrix}. \quad (168)$$

Note that \mathbf{H}^f is different, though closely related, from the constrained Hessian \mathbf{H}^c of Υ in the subspace $\mathbf{A}\mathbf{z} = \mathbf{0}$, see, for example, Luenberger and Ye [2008, Chapter 11]. The constrained Hessian \mathbf{H}^c has infinitely many matrix representations, all of which correspond to similar matrices, that is, matrices with the same set of eigenvalues. A matrix representation may be obtained by first picking an orthonormal basis of the $(n-r)$ -dimensional space $\{\mathbf{z} \mid \mathbf{A}\mathbf{z} = \mathbf{0}\}$. Let \mathbf{E} denote the $n \times (n-r)$ matrix whose columns are the vectors in the basis. Then, a matrix representation of \mathbf{H}^c is given by

$$\mathbf{H}^c = \mathbf{E}^\top \mathbf{H} \mathbf{E}, \quad (169)$$

where \mathbf{H} is as before the unconstrained Hessian of Υ with respect to the variables \mathbf{z} . We are ready to prove the following. It is useful to recall here that congruent matrices have the same number of negative, zero, and positive eigenvalues.

LEMMA B.4. \mathbf{H}^f is congruent to any matrix representation of \mathbf{H}^c . Moreover, it holds that

$$\text{Det}(\mathbf{H}^f) = \text{Det}(\mathbf{H}^c) \text{Det}(\mathbf{A}_f \mathbf{A}_f^\top) / \text{Det}(\mathbf{A}_{fs})^2.$$

PROOF OF LEMMA B.4. The columns of the matrix \mathbf{S} defined in Eq. (168) form a basis of the space $\{\mathbf{z} \mid \mathbf{A}\mathbf{z} = \mathbf{0}\}$. Indeed, \mathbf{S} has clearly full column rank and also $\mathbf{A}_f \mathbf{S} = \mathbf{0}$, implying $\mathbf{A}\mathbf{S} = \mathbf{0}$ as well. For future use, by a direct evaluation

$$\mathbf{S}^\top \mathbf{S} = \mathbf{I} + \mathbf{A}_{ff}^\top (\mathbf{A}_{fs} \mathbf{A}_{fs}^\top)^{-1} \mathbf{A}_{ff}, \text{ so } \text{Det}(\mathbf{S}^\top \mathbf{S}) = \text{Det}\left(\mathbf{I} + \mathbf{A}_{ff} \mathbf{A}_{ff}^\top (\mathbf{A}_{fs} \mathbf{A}_{fs}^\top)^{-1}\right),$$

where the latter equality uses Sylvester's determinant theorem. This clearly yields

$$\text{Det}(\mathbf{S}^\top \mathbf{S}) = \text{Det}(\mathbf{A}_f \mathbf{A}_f^\top) / \text{Det}(\mathbf{A}_{fs})^2. \quad (170)$$

Comparing (168) and (169), the only difference is that \mathbf{S} does not necessarily encode an orthonormal basis. Nevertheless, there clearly exists an invertible matrix \mathbf{P} such that $\mathbf{S}\mathbf{P}$ consists of orthonormal columns, for example, by the Gram-Schmidt process on the columns of \mathbf{S} . It follows that $\mathbf{P}^\top \mathbf{H}^f \mathbf{P}$ is a matrix representation of \mathbf{H}^c . This proves the first part of the lemma and also gives $\text{Det}(\mathbf{H}^c) = \text{Det}(\mathbf{H}^f) \text{Det}(\mathbf{P})^2$.

For the second part, the selection of \mathbf{P} implies that $(\mathbf{S}\mathbf{P})^\top \mathbf{S}\mathbf{P}$ is the identity matrix and hence $\text{Det}(\mathbf{S}^\top \mathbf{S}) \text{Det}(\mathbf{P})^2 = 1$. The desired equality follows. \square

Lemma B.4 allows us to focus on the determinant of \mathbf{H}^c or equivalently the product of its eigenvalues. The latter may be handled using bordered Hessians. Specifically, let \mathbf{A}_f be any submatrix of \mathbf{A} induced by r linearly independent rows (recall that $\mathbf{A}\mathbf{z} = \mathbf{0}$ iff $\mathbf{A}_f \mathbf{z} = \mathbf{0}$). Then, λ is an eigenvalue of \mathbf{H}^c iff it is a root of the polynomial

$$p(\lambda) = \text{Det}\left(\begin{bmatrix} \mathbf{0} & \mathbf{A}_f \\ -\mathbf{A}_f^\top & \mathbf{H} - \lambda \mathbf{I}_n \end{bmatrix}\right), \quad (171)$$

see, for example, Luenberger and Ye [2008, Chapter 11].

In our case, deleting rows of \mathbf{A} to obtain \mathbf{A}_f would cause undesirable complications. In the following, we circumvent such deletions by adding suitable "perturbations". We will also allow for certain degrees of freedom to select the perturbations which will be exploited in the computations. We first prove the following.

For a polynomial $p(s)$, $[s^t]p(s)$ denotes the coefficient of s^t in $p(s)$.

LEMMA B.5. *Let $\mathbf{M} \in \mathbb{R}^{m \times m}$ be a symmetric matrix with rank r and let $\mu_i, i = 1, \dots, m$ be the eigenvalues of \mathbf{M} . Let $\mathbf{T} \in \mathbb{R}^{m \times m}$ be a symmetric matrix. We can choose orthonormal unit eigenvectors $\mathbf{v}_i, i = 1, \dots, m$ of \mathbf{M} (where \mathbf{v}_i has eigenvalue μ_i) so that*

$$[\varepsilon^{m-r}] \text{Det}(\varepsilon \mathbf{T} + \mathbf{M}) = \prod_{i; \mu_i \neq 0} \mu_i \prod_{i; \mu_i = 0} \mathbf{v}_i^\top \mathbf{T} \mathbf{v}_i, \quad (172)$$

In particular, if \mathbf{T} is positive semidefinite and $[\mathbf{T} \mathbf{M}]$ has full row rank, the right-hand side of (172) is nonzero.

PROOF OF LEMMA B.5. Without loss of generality, assume that μ_1, \dots, μ_r are the non-zero eigenvalues of \mathbf{M} . Since \mathbf{M} is symmetric, there exists an orthogonal matrix \mathbf{S} such that $\mathbf{S}^\top \mathbf{M} \mathbf{S} = \mathbf{diag}(\mu_1, \dots, \mu_m)$ where the columns of \mathbf{S} , denoted by $\mathbf{v}'_1, \mathbf{v}'_2, \dots, \mathbf{v}'_m$ henceforth, are eigenvectors of \mathbf{M} . We may decompose \mathbf{S} into the block form $[\mathbf{S}_{\neq 0} \mathbf{S}_0]$ where \mathbf{S}_0 is an $m \times (m - r)$ matrix such that $\mathbf{M} \mathbf{S}_0$ is the zero matrix.

Since \mathbf{T} is symmetric, the matrix $\mathbf{S}_0^\top \mathbf{T} \mathbf{S}_0$ is symmetric and hence there exists an orthogonal matrix \mathbf{R}_0 such that $\mathbf{R}_0^\top (\mathbf{S}_0^\top \mathbf{T} \mathbf{S}_0) \mathbf{R}_0$ is the diagonal matrix whose diagonal entries are the eigenvalues of $\mathbf{S}_0^\top \mathbf{T} \mathbf{S}_0$. Let \mathbf{S}' be the matrix with block form $[\mathbf{S}_{\neq 0} \mathbf{S}_0 \mathbf{R}_0]$. We claim that the columns $\mathbf{v}_1, \dots, \mathbf{v}_m$ of \mathbf{S}' satisfy the statement of the lemma.

Since $\mathbf{v}_i = \mathbf{v}'_i$ for $i \leq r$ and the \mathbf{v}_i 's with $i > r$ are linear combinations of eigenvectors of \mathbf{M} corresponding to the same eigenvalue 0, we have that the \mathbf{v}_i 's are eigenvectors of \mathbf{M} . Orthonormality of the \mathbf{v}_i 's follows from

$$(\mathbf{S}')^\top \mathbf{S}' = \begin{bmatrix} \mathbf{S}_{\neq 0}^\top \mathbf{S}_{\neq 0} & \mathbf{S}_{\neq 0}^\top \mathbf{S}_0 \mathbf{R}_0 \\ \mathbf{R}_0^\top \mathbf{S}_0^\top \mathbf{S}_{\neq 0} & \mathbf{R}_0^\top \mathbf{S}_0^\top \mathbf{S}_0 \mathbf{R}_0 \end{bmatrix},$$

and the fact that the matrices $\mathbf{S}_{\neq 0}^\top \mathbf{S}_{\neq 0}, \mathbf{S}_{\neq 0}^\top \mathbf{S}_0$ are zero matrices, while $\mathbf{S}_0^\top \mathbf{S}_0, \mathbf{S}_{\neq 0}^\top \mathbf{S}_{\neq 0}, \mathbf{R}_0^\top \mathbf{R}_0$ are the identity matrices (by the orthogonality of \mathbf{S}, \mathbf{R}_0).

It remains to prove (172). By the orthogonality of \mathbf{S}' and the fact that the columns of \mathbf{S}' are eigenvectors of \mathbf{M} , we have

$$\text{Det}(\varepsilon \mathbf{T} + \mathbf{M}) = \text{Det}(\varepsilon (\mathbf{S}')^\top \mathbf{T} \mathbf{S}' + \mathbf{diag}(\mu_1, \dots, \mu_m)).$$

Let $\mathbf{W} := \varepsilon (\mathbf{S}')^\top \mathbf{T} \mathbf{S}' + \mathbf{diag}(\mu_1, \dots, \mu_m)$ and denote the (i, j) -entry of \mathbf{W} by $W_{i,j}$. By Leibniz's formula, we have that

$$\text{Det}(\mathbf{W}) = \sum_{\sigma \in S_m} \text{sgn}(\sigma) \prod_{i=1}^m W_{i, \sigma(i)}.$$

Note that the coefficient of ε^{m-r} in $\text{Det}(\mathbf{W})$ is obtained by considering only those σ such that $\sigma(i) = i$ for all $i \in [r]$ (since the only entries of \mathbf{W} which are nonzero for $\varepsilon = 0$ are $W_{i,i}$ for $i = 1, \dots, r$). The contribution from such σ to the coefficient of ε^{m-r} is given by $\text{Det}(\mathbf{T}_0) \prod_{i=1}^r \mu_i$ where $\mathbf{T}_0 = \mathbf{R}_0^\top \mathbf{S}_0^\top \mathbf{T} \mathbf{S}_0 \mathbf{R}_0$. By the choice of \mathbf{R}_0 , we have that \mathbf{T}_0 is a diagonal matrix (whose diagonal entries are eigenvalues of $\mathbf{S}_0^\top \mathbf{T} \mathbf{S}_0$). The diagonal entries of \mathbf{T}_0 are given by $\mathbf{v}_{r+1}^\top \mathbf{T} \mathbf{v}_{r+1}, \dots, \mathbf{v}_m^\top \mathbf{T} \mathbf{v}_m$. Hence, $\text{Det}(\mathbf{T}_0) = \prod_{i=r+1}^m \mathbf{v}_i^\top \mathbf{T} \mathbf{v}_i$, yielding (172).

For the second part, by the positive semidefiniteness of \mathbf{T} , we have $\mathbf{v}_i^\top \mathbf{T} \mathbf{v}_i = 0$ iff $\mathbf{T} \mathbf{v}_i = 0$. Thus, for the right-hand side in (172) to be zero, there must exist $i > r$ such that $\mathbf{v}_i^\top [\mathbf{T} \mathbf{M}] = \mathbf{0}$ (recall that any \mathbf{v}_i with $i > r$ satisfies $\mathbf{M} \mathbf{v}_i = \mathbf{0}$), which is excluded by the assumption that $[\mathbf{T} \mathbf{M}]$ has full row rank, concluding the proof. \square

The following lemma gives the promised extension of (171).

LEMMA B.6. *Suppose that \mathbf{T} is a diagonal positive semidefinite $m \times m$ matrix such that $[\mathbf{T} \mathbf{A}]$ has full row rank. Let \mathbf{H} (respectively \mathbf{H}^c) be the unconstrained (respectively*

constrained) Hessian of Υ evaluated at a point \mathbf{z}_0 . Then, λ is an eigenvalue of \mathbf{H}^c iff it is a root of the polynomial

$$p(\lambda) = [\varepsilon^{m-r}] \text{Det}(\mathbf{H}_\lambda) \text{ where } \mathbf{H}_\lambda = \begin{bmatrix} \varepsilon \mathbf{T} & \mathbf{A} \\ -\mathbf{A}^\top & \mathbf{H} - \lambda \mathbf{I}_n \end{bmatrix}. \quad (173)$$

Further, if \mathbf{H} is invertible, then $\text{Det}(\mathbf{H}^c) = (-1)^r \text{Det}(\mathbf{H}) \frac{[\varepsilon^{m-r}] \text{Det}(\varepsilon \mathbf{T} + \mathbf{A} \mathbf{H}^{-1} \mathbf{A}^\top)}{[\varepsilon^{m-r}] \text{Det}(\varepsilon \mathbf{T} - \mathbf{A} \mathbf{A}^\top)}$.

PROOF OF LEMMA B.6. Let $\mathbf{T} = (t_{i,j})_{i,j \in [m]}$ and $\mathbf{H}_\lambda = (h_{i,j})_{i,j \in [m+n]}$. Let $\mathcal{W} = \binom{[m]}{m-r}$ and for $W \in \mathcal{W}$ let $P_W = \{\sigma \in S_{m+n} \mid \{i \in [m] \mid \sigma(i) = i\} = W\}$. Since \mathbf{T} is diagonal, by Leibniz's formula,

$$p(\lambda) = [\varepsilon^{m-r}] \text{Det}(\mathbf{H}_\lambda) = \sum_{W \in \mathcal{W}} \prod_{i \in W} t_{i,i} \sum_{\sigma \in P_W} \text{sgn}(\sigma) \prod_{i \in [m+n] \setminus W} h_{i,\sigma(i)}. \quad (174)$$

Let $\mathbf{A}_{[m] \setminus W}$ be the $r \times n$ submatrix of \mathbf{A} which is obtain by excluding the rows indexed by W . Identifying permutations in P_W with permutations of $[n+r]$ in the natural way, we obtain

$$\sum_{\sigma \in P_W} \text{sgn}(\sigma) \prod_{i \in [m+n] \setminus W} h_{i,\sigma(i)} = \text{Det} \left(\begin{bmatrix} \mathbf{0} & \mathbf{A}_{[m] \setminus W} \\ -\mathbf{A}_{[m] \setminus W}^\top & \mathbf{H} - \lambda \mathbf{I}_n \end{bmatrix} \right) \equiv q_W(\lambda). \quad (175)$$

If $\mathbf{A}_{[m] \setminus W}$ has row rank $< r$, then $q_W(\lambda)$ is 0. Otherwise, the roots of $q_W(\lambda)$ are the eigenvalues of \mathbf{H}^c , c.f., (171). By (174), this is also the case for $p(\lambda)$, provided it is not identically zero.

To prove that $p(\lambda)$ is nonzero, we prove that the leading coefficient of $p(\lambda)$ is nonzero. Starting from (175) and plugging into (174), the leading coefficient of $p(\lambda)$ can easily be seen to equal

$$[\varepsilon^{m-r}] \text{Det} \left(\begin{bmatrix} \varepsilon \mathbf{T} & \mathbf{A} \\ -\mathbf{A}^\top & -\mathbf{I}_n \end{bmatrix} \right) = [\varepsilon^{m-r}] (-1)^n \text{Det}(\varepsilon \mathbf{T} - \mathbf{A} \mathbf{A}^\top),$$

where in the latter equality we used the Schur complement of the block $-\mathbf{I}_n$. The last expression is nonzero by Lemma B.5.

The determinant of \mathbf{H}^c is the product of its eigenvalues, which in turn equals $(-1)^{n-r} p(0)$ divided by the leading coefficient of $p(\lambda)$. The latter has already been computed. The former, using the Schur complement of the invertible \mathbf{H} , is equal to $[\varepsilon^{m-r}] \text{Det}(\mathbf{H}) \text{Det}(\varepsilon \mathbf{T} + \mathbf{A} \mathbf{H}^{-1} \mathbf{A}^\top)$. This concludes the proof. \square

Finally, we combine these lemmas to obtain the following.

LEMMA B.7. Let Υ be a function of $\mathbf{z} \in \mathbb{R}^n$ subject to the linear constraints $\mathbf{A} \mathbf{z} = \mathbf{b}$, where $\mathbf{A} \in \mathbb{R}^{m \times n}$ and \mathbf{A} has rank r . Let $(\mathbf{A}_f, \mathbf{A}_{f_s})$ specify a full dimensional representation of Υ and let \mathbf{H}^f be the corresponding Hessian of Υ evaluated at a point \mathbf{z}_0 .

Suppose \mathbf{T} is a positive semidefinite diagonal matrix with dimensions $m \times m$ such that $[\mathbf{T} \mathbf{A}]$ has full row rank. Let \mathbf{H} be the unconstrained Hessian of Υ evaluated at \mathbf{z}_0 . If \mathbf{H} is invertible, then

$$\text{Det}(-\mathbf{H}^f) = \frac{L(\mathbf{A}_f, \mathbf{A}, \mathbf{T})}{\text{Det}(\mathbf{A}_{f_s})^2} \text{Det}(-\mathbf{H}) [\varepsilon^{m-r}] \text{Det}(\varepsilon \mathbf{T} - \mathbf{A} \mathbf{H}^{-1} \mathbf{A}^\top), \quad (176)$$

where $L(\mathbf{A}_f, \mathbf{A}, \mathbf{T}) = (-1)^r \text{Det}(\mathbf{A}_f \mathbf{A}_f^\top) / [\varepsilon^{m-r}] \text{Det}(\varepsilon \mathbf{T} - \mathbf{A} \mathbf{A}^\top)$.

PROOF OF LEMMA B.7. Just combine Lemmas B.4 and B.6. The minor sign change $-\mathbf{H}^f$ in the statement can easily be accounted by applying the lemmas to the function $-\Upsilon$. \square

The right-hand side of (176) has two qualitatively different factors: the factor $L(\mathbf{A}_f, \mathbf{A}, \mathbf{T})/\text{Det}(\mathbf{A}_{fs})^2$ depends on the specific full dimensional representation, while the remaining factor is tied to the Hessian of Υ . The technical convenience of Lemma B.7 is dual: first, it gives an explicit formula for $\text{Det}(-\mathbf{H}^f)$ without doing substitutions which would hinder the combinatorial view of the constraints \mathbf{A} ; second, it isolates the deletions of rows of \mathbf{A} in the factor $L(\mathbf{A}_f, \mathbf{A}, \mathbf{T})$ and leaves untouched the more complicated matrix $\mathbf{A}\mathbf{H}^{-1}\mathbf{A}^\top$.

B.1.2. The Computations. In this section, we utilize Lemma B.7 to compute the determinants in Lemma B.3.

Notation. For a vector $\mathbf{z} \in \mathbb{R}^n$ we denote by \mathbf{z}^D the $n \times n$ diagonal matrix $\text{diag}\{z_1, \dots, z_n\}$. For vectors $\mathbf{z}_i \in \mathbb{R}^{m_i}$, $i = 1, \dots, t$ we denote by $[\mathbf{z}_1, \dots, \mathbf{z}_t]^\top$ the $\mathbb{R}^{\sum_i m_i}$ vector which is the concatenation of the vectors $\mathbf{z}_1, \dots, \mathbf{z}_t$. For matrices \mathbf{A} and \mathbf{B} , $\mathbf{A} \otimes \mathbf{B}$ will denote the Kronecker product of \mathbf{A} , \mathbf{B} , while $\mathbf{A} \oplus \mathbf{B}$ is the direct sum of \mathbf{A} , \mathbf{B} , that is, the block diagonal matrix $\text{diag}\{\mathbf{A}, \mathbf{B}\}$. The expression $\oplus_2 \mathbf{A}$ is a shorthand for $\mathbf{A} \oplus \mathbf{A}$. Further, \mathbf{I}_n denotes the identity matrix of dimensions $n \times n$. Finally, $\mathbf{1}_n$, $\mathbf{0}_n$ denote the all-one and all-zero n -dimensional vector.

To start, the equality constraints in (164) may be written in the form

$$\mathbf{A}_1 [\alpha, \beta, \mathbf{x}]^\top = \mathbf{0}.$$

Note, we do need to pay attention at this point to the constant terms in (164), since we only care for the tangent space that the constraints induce (this is why the right-hand side in the previous equation is $\mathbf{0}$). The matrix \mathbf{A}_1 has dimensions $(2q + 2) \times (|P_1| + 2q)$ (cf., (163) for the definition of P_1). Note that we exclude from consideration variables x_{ij} which are hard-coded to zero. This is done to ensure that the unconstrained Hessians are invertible, so that Lemma B.7 applies directly. It will be useful to decompose the matrix \mathbf{A}_1 as

$$\mathbf{A}_1 = \begin{bmatrix} \mathbf{A}_{1,\alpha\beta} & \mathbf{0} \\ -\mathbf{I}_{2q} & \mathbf{A}_{1,\mathbf{x}} \end{bmatrix}, \quad (177)$$

where $\mathbf{A}_{1,\alpha\beta}$, $\mathbf{A}_{1,\mathbf{x}}$ have dimensions $2 \times 2q$ and $2q \times |P_1|$, respectively.

The easiest way to handle the matrix $\mathbf{A}_{1,\mathbf{x}}$ is as the incidence matrix of a bipartite graph $G_{\mathbf{x}}$. First, we introduce some notation: for an undirected graph G , we denote by \mathbf{A}_G the 0,1 incidence matrix of G , by \mathbf{R}_G the adjacency matrix of G , by \mathbf{D}_G the diagonal matrix whose diagonal entries are equal to the degrees of the vertices in G and by Λ_G the matrix $\mathbf{D}_G + \mathbf{R}_G$. We will also be interested in the case where the graph G is weighted, in which case we assume that the weights on the edges are given by the diagonal entries of a square diagonal matrix \mathbf{W}_G . We denote by \mathbf{R}_G^w , \mathbf{D}_G^w , Λ_G^w the weighted versions of the matrices \mathbf{R}_G , \mathbf{D}_G , Λ_G . It is well known that

$$\mathbf{A}_G \mathbf{A}_G^\top = \Lambda_G, \quad \mathbf{A}_G \mathbf{W}_G \mathbf{A}_G^\top = \Lambda_G^w. \quad (178)$$

The bipartite graph $G_{\mathbf{x}}$ has vertex bipartition $([q], [q])$ and an edge (i, j) is present iff $(i, j) \in P_1$, that is, $B_{ij} > 0$. Since \mathbf{B} is irreducible, $G_{\mathbf{x}}$ is undirected and connected. An edge (i, j) in $G_{\mathbf{x}}$ has weight x_{ij} . In the language of (178), $\mathbf{W}_{G_{\mathbf{x}}} = \mathbf{x}^D$ (the choice of $\mathbf{W}_{G_{\mathbf{x}}}$ will become apparent when we consider the unconstrained Hessian). Applying (178) to the graph $G_{\mathbf{x}}$ is useful to do explicitly in order to decompose the resulting matrices. In particular, since these graphs are undirected and bipartite, we have

$$\Lambda^w(G_{\mathbf{x}}) = \begin{bmatrix} \alpha^D & \mathbf{S}_{\mathbf{x}} \\ \mathbf{S}_{\mathbf{x}}^\top & \beta^D \end{bmatrix}, \quad (179)$$

where $\mathbf{S}_\mathbf{x}$ is the $q \times q$ matrix whose (i, j) entry is x_{ij} . Note that the total weight of the edges incident to vertices in $G_\mathbf{x}$ (in other words, the diagonal entries of the matrix \mathbf{D}^w) was substituted using (164).

We next state the unconstrained Hessians that will be of interest to us. From Lemma B.3, these are: (i) $\mathbf{H}_{1,\mathbf{x}}$, the Hessian of Υ_1/Δ with respect to \mathbf{x} when α, β are fixed, (ii) \mathbf{H}_1 , the Hessian of Υ_1/Δ with respect to $\alpha, \beta, \mathbf{x}$. These two matrices are all diagonal and by inspection one can check that

$$\begin{aligned} (\mathbf{H}_{1,\mathbf{x}})^{-1} &= -\mathbf{x}^D, \quad (\mathbf{H}_1)^{-1} = \frac{\Delta}{\Delta-1} \alpha^D \oplus \frac{\Delta}{\Delta-1} \beta^D \oplus (\mathbf{H}_{1,\mathbf{x}})^{-1}, \\ \text{Det}(-\mathbf{H}_{1,\mathbf{x}})^{-1} &= \prod_{(i,j) \in P_1} x_{ij}, \quad \text{Det}(-\mathbf{H}_1)^{-1} = \text{Det}(-\mathbf{H}_{1,\mathbf{x}})^{-1} \left(\frac{\Delta}{\Delta-1} \right)^{2q} \prod_{i \in [q]} \alpha_i \prod_{j \in [q]} \beta_j. \end{aligned} \quad (180)$$

We are now ready to evaluate these matrices at a global maximum $(\alpha^*, \beta^*, \mathbf{x}^*)$ of Υ_1 . Henceforth, we will not explicitly use asterisks in the notation with the understanding that the values of all the variables are fixed to their optimal values.

We will apply Lemma B.7 to the matrices $\mathbf{H}_{1,\mathbf{x}}^f, \mathbf{H}_1^f$ using the matrices

$$\mathbf{T}_{1,\mathbf{x}} = \alpha^D \oplus \beta^D, \quad \mathbf{T}_1 = \mathbf{I}_2 \oplus \mathbf{0}_{2q \times 2q}, \quad (181)$$

respectively (in (181), $\mathbf{0}_{2q \times 2q}$ denotes the $2q \times 2q$ matrix with all zeros). We first compute the determinants of $\mathbf{M}_{1,\mathbf{x}} := \varepsilon \mathbf{T}_{1,\mathbf{x}} - \mathbf{A}_{1,\mathbf{x}} (\mathbf{H}_{1,\mathbf{x}})^{-1} \mathbf{A}_{1,\mathbf{x}}^\top$, $\mathbf{M}_1 := \varepsilon \mathbf{T}_1 - \mathbf{A}_1 (\mathbf{H}_1)^{-1} \mathbf{A}_1^\top$, which contribute the most interesting factors in Lemma B.7.

We begin with the simplest of these matrices, $\mathbf{M}_{1,\mathbf{x}}$. Note that $\mathbf{A}_{1,\mathbf{x}}$ has rank $2q - 1$, so by Lemma B.7 we want to compute $[\varepsilon] \text{Det}(\mathbf{M}_{1,\mathbf{x}})$. Using (178), (180), and (181), we obtain that $\mathbf{M}_{1,\mathbf{x}}$ has the following form

$$\begin{aligned} \mathbf{M}_{1,\mathbf{x}} &= \begin{bmatrix} \alpha^D(\varepsilon \mathbf{I}_q + \mathbf{I}_q) & \mathbf{S}_\mathbf{x} \\ \mathbf{S}_\mathbf{x}^\top & \beta^D(\varepsilon \mathbf{I}_q + \mathbf{I}_q) \end{bmatrix}, \\ \text{so } \text{Det}(\mathbf{M}_{1,\mathbf{x}}) &= \left(\prod_{i \in [q]} \alpha_i \prod_{j \in [q]} \beta_j \right) \text{Det}(\varepsilon \mathbf{I}_{2q} + \mathbf{I}_{2q} + \mathbf{J}), \end{aligned} \quad (182)$$

where \mathbf{J} is the matrix in Lemma A.4. Note that in Eq. (182), to get the second equality, we did the following operations on $\mathbf{M}_{1,\mathbf{x}}$: for $i = 1, \dots, q$, we divided the i th row of by $\sqrt{\alpha_i}$, the i th column by $\sqrt{\alpha_i}$, the $(i+q)$ -th row by $\sqrt{\beta_i}$, the $(i+q)$ -th column by $\sqrt{\beta_i}$. The eigenvalues of the matrix $\varepsilon \mathbf{I}_{2q} + \mathbf{I}_{2q} + \mathbf{J}$ are shifts of the eigenvalues of \mathbf{J} and are given by

$$\varepsilon, \varepsilon + 2, \varepsilon + 1 \pm \lambda_1, \dots, \varepsilon + 1 \pm \lambda_{q-1},$$

cf., Lemma A.4 for the definition of the λ_i and their properties. We thus obtain

$$[\varepsilon] \text{Det}(\mathbf{M}_{1,\mathbf{x}}) = 2 \prod_{i \in [q]} \alpha_i \prod_{j \in [q]} \beta_j \prod_{i \in [q-1]} (1 - \lambda_i^2). \quad (183)$$

The determinant of the matrix \mathbf{M}_1 is more complicated to compute due to its more intricate block structure, which requires using Schur's complement formula to handle. As in the previous argument, we first write out its block structure and then appropriately normalize the resulting matrix. Here, the normalization is slightly more intricate. The analog of (182) is

$$\text{Det}(\mathbf{M}_1) = \text{Det}(\mathbf{H}_1^f) \prod_{i \in [q]} \alpha_i \prod_{j \in [q]} \beta_j, \quad (184)$$

where

$$\mathbf{H}'_1 := \frac{\Delta}{\Delta-1} \begin{bmatrix} (\varepsilon \frac{\Delta-1}{\Delta} - 1) \mathbf{I}_2 & \mathbf{V} \\ \mathbf{V}^\top & -\frac{\Delta-1}{\Delta} \mathbf{W} \end{bmatrix}, \quad \mathbf{W} := \frac{1}{\Delta-1} \mathbf{I}_{2q} - \mathbf{J}, \quad \mathbf{V}^\top := \begin{bmatrix} \sqrt{\alpha} & \mathbf{0}_q \\ \mathbf{0}_q & \sqrt{\beta} \end{bmatrix}. \quad (185)$$

(\mathbf{J} is the matrix in Lemma A.4; note also that \mathbf{V} has dimension $2 \times 2q$ and its two rows are given by $\sqrt{\alpha_1}, \dots, \sqrt{\alpha_q}, 0, \dots, 0$ and $0, \dots, 0, \sqrt{\beta_1}, \dots, \sqrt{\beta_q}$ where each row has q zeros.) Equation (184) is obtained by performing the following operations on \mathbf{M}_1 : for $i, j = 1, \dots, q$, divide the $2+i$ row by $\sqrt{\alpha_i}$ and the $2+q+j$ row by $\sqrt{\beta_j}$; and the same operations on columns.

In light of (184), it suffices to compute $\text{Det}(\mathbf{H}'_1)$. To do this, we proceed by taking the Schur complement of the matrix \mathbf{W} . The spectrum of \mathbf{W} is

$$t \pm 1, t \pm \lambda_1, \dots, t \pm \lambda_{q-1},$$

where $t = 1/(\Delta-1)$. It follows that

$$\text{Det}(\mathbf{W}) = -\frac{\Delta(\Delta-2)}{(\Delta-1)^{2q}} \prod_{i \in [q-1]} (1 - (\Delta-1)^2 \lambda_i^2), \quad (186)$$

where $\lambda_i, i \in [q-1]$ are as in Lemma A.4. Note that \mathbf{W} is invertible, since the λ_i 's are nonnegative and $\max \lambda_i < \frac{1}{\Delta-1}$. By taking the Schur complement of the matrix \mathbf{W} in \mathbf{H}'_1 , we obtain

$$\text{Det}(\mathbf{H}'_1) = \left(\frac{\Delta}{\Delta-1} \right)^2 \text{Det}(\mathbf{W}) \text{Det} \left(\varepsilon \frac{\Delta-1}{\Delta} \mathbf{I}_2 + \mathbf{Z} \right), \quad \text{where } \mathbf{Z} = -\mathbf{I}_2 + \frac{\Delta}{\Delta-1} \mathbf{V} \mathbf{W}^{-1} \mathbf{V}^\top. \quad (187)$$

We are left with the evaluation of $\text{Det}(\varepsilon \frac{\Delta-1}{\Delta} \mathbf{I}_2 + \mathbf{Z})$. The complication here is the nontrivial inverse of \mathbf{W} appearing in the formulation of \mathbf{Z} . The key idea to circumvent the computation of \mathbf{W}^{-1} is the following equality

$$\mathbf{V} \mathbf{W} = \left(\frac{1}{\Delta-1} \mathbf{I}_2 - \mathbf{J}' \right) \mathbf{V}, \quad \text{where } \mathbf{J}' = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

The equality can be checked using the relations $\sum_j x_{ij} = \alpha_i$ and $\sum_i x_{ij} = \beta_j$. Using that $\mathbf{V} \mathbf{V}^\top = \mathbf{I}_2$, we obtain

$$\mathbf{Z} = -\mathbf{I}_2 + \frac{\Delta}{\Delta-1} \left(\frac{1}{\Delta-1} \mathbf{I}_2 - \mathbf{J}' \right)^{-1} \mathbf{V} \mathbf{V}^\top = -\frac{\Delta-1}{\Delta-2} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \quad (188)$$

We thus obtain

$$[\varepsilon] \text{Det} \left(\varepsilon \frac{\Delta-1}{\Delta} \mathbf{I}_2 + \mathbf{Z} \right) = -\frac{2(\Delta-1)^2}{\Delta(\Delta-2)}. \quad (189)$$

Plugging (186) and (189) in (187), we obtain

$$[\varepsilon] \text{Det}(\mathbf{H}'_1) = \frac{2\Delta^2}{(\Delta-1)^{2q}} \prod_{i \in [q-1]} (1 - (\Delta-1)^2 \lambda_i^2).$$

Using this and (184), we obtain

$$[\varepsilon] \text{Det}(\varepsilon \mathbf{T}_1 - \mathbf{A}_1(\mathbf{H}_1)^{-1} \mathbf{A}_1^\top) = \frac{2\Delta^2}{(\Delta-1)^{2q}} \prod_{i \in [q]} \alpha_i \prod_{j \in [q]} \beta_j \prod_{i \in [q-1]} (1 - (\Delta-1)^2 \lambda_i^2). \quad (190)$$

Equations (180), (183), and (190) deal with the factors in Lemma B.7 which are tied to the Hessians of the functions. While these contribute the most interesting factors, some care is needed to deal with the remaining factors. This is accomplished in the following lemma, which is given in the end of this section.

LEMMA B.8. *Let $((\mathbf{A}_{1,\mathbf{x}})_f, (\mathbf{A}_{1,\mathbf{x}})_{f_s}), ((\mathbf{A}_1)_f, (\mathbf{A}_1)_{f_s})$ specify arbitrary full dimensional representations of the spaces $\mathbf{A}_{1,\mathbf{x}}\mathbf{x} = \mathbf{0}$, $\mathbf{A}_1[\boldsymbol{\alpha}, \boldsymbol{\beta}, \mathbf{x}]^\top = \mathbf{0}$, respectively. Then,*

$$\text{Det}((\mathbf{A}_{1,\mathbf{x}})_{f_s})^2 = \text{Det}((\mathbf{A}_1)_{f_s})^2 = 1, \quad (191)$$

$$L((\mathbf{A}_{1,\mathbf{x}})_f, \mathbf{A}_{1,\mathbf{x}}, \mathbf{T}_{1,\mathbf{x}}) = 1/2, \quad L((\mathbf{A}_1)_f, \mathbf{A}_1, \mathbf{T}_1) = 1/2, \quad (192)$$

where $\mathbf{T}_{1,\mathbf{x}}, \mathbf{T}_1$ are given by (181) and the quantities in (192) are defined in Lemma B.7.

We are now ready to finish the proof of Lemma A.9.

PROOF OF LEMMA A.9. Apply Lemma B.7 two times to unravel the determinants appearing in Lemma B.3. Each of the resulting quantities has been computed and appears in one of (180), (183), or (190) or Lemma B.8. Straightforward substitutions yield

$$\begin{aligned} \text{Det}(-\mathbf{H}_{1,\mathbf{x}}^f) &= \left(\prod_{(i,j) \in P_1} x_{ij} \right)^{-1} \prod_{i \in [q]} \alpha_i \prod_{j \in [q]} \beta_j \prod_{i \in [q-1]} (1 - \lambda_i)^2, \\ \text{Det}(-\mathbf{H}_1^f) &= \frac{1}{\Delta^{2(q-1)}} \left(\prod_{(i,j) \in P_1} x_{ij} \right)^{-1} \prod_{i \in [q-1]} (1 - (\Delta - 1)^2 \lambda_i^2). \end{aligned}$$

Thus, Lemma B.3 gives

$$\lim_{n \rightarrow \infty} \frac{\mathbf{E}_G[\mathbf{Z}_G^{\mathbf{p}}]}{e^{n\Upsilon_1(\boldsymbol{\alpha}^*, \boldsymbol{\beta}^*, \mathbf{x}^*)}} = \prod_{i=1}^{q-1} (1 - (\Delta - 1)^2 \lambda_i^2)^{-1/2} \prod_{i=1}^{q-1} (1 - \lambda_i^2)^{-(\Delta-1)/2}. \quad (193)$$

Note that in the last expression, only the eigenvalues of the matrix \mathbf{J} (different from 1) in Lemma A.4 appear. For the asymptotics of $\mathbf{E}_G[(\mathbf{Z}_G^{\mathbf{p}})^2]$, by Lemma B.2, it suffices to consider the spin system with interaction matrix $\mathbf{B} \otimes \mathbf{B}$ (and dominant phase $\boldsymbol{\gamma} = \boldsymbol{\alpha}^* \otimes \boldsymbol{\alpha}^*, \boldsymbol{\delta}^* = \boldsymbol{\beta}^* \otimes \boldsymbol{\beta}^*, \mathbf{y} = \mathbf{x}^* \otimes \mathbf{x}^*$). The eigenvalues (different from 1) of the matrix $\mathbf{J} \otimes \mathbf{J}$ are λ_i for $i \in [q-1]$ and $\lambda_i \lambda_j$ for $i, j \in [q-1]$. Thus, we obtain

$$\lim_{n \rightarrow \infty} \frac{\mathbf{E}_G[(\mathbf{Z}_G^{\mathbf{p}})^2]}{e^{n\Upsilon_2(\boldsymbol{\gamma}^*, \boldsymbol{\delta}^*, \mathbf{y}^*)}} = C \cdot \prod_{i=1}^{q-1} (1 - (\Delta - 1)^2 \lambda_i^2)^{-1/2} \prod_{i=1}^{q-1} (1 - \lambda_i^2)^{-(\Delta-1)/2}, \quad (194)$$

where C is the constant in the statement of the lemma. Since $\boldsymbol{\alpha}^*, \boldsymbol{\beta}^*$ is a dominant phase, Theorem 1.4 implies $\Psi_2(\boldsymbol{\alpha}^*, \boldsymbol{\beta}^*) = 2\Psi_1(\boldsymbol{\alpha}^*, \boldsymbol{\beta}^*)$ (recall that Ψ_1, Ψ_2 are given by (8) and (9), respectively), so we obtain that

$$\Upsilon_2(\boldsymbol{\gamma}^*, \boldsymbol{\delta}^*, \mathbf{y}^*) = \Psi_2(\boldsymbol{\alpha}^*, \boldsymbol{\beta}^*) = 2\Psi_1(\boldsymbol{\alpha}^*, \boldsymbol{\beta}^*) = 2\Upsilon_1(\boldsymbol{\alpha}^*, \boldsymbol{\beta}^*, \mathbf{x}^*).$$

Thus, combining (193) and (194) with Lemma B.1 yields the result. \square

Finally, we give the proof of Lemma B.8.

PROOF OF LEMMA B.8. We first prove (191). Since $\mathbf{A}_{1,\mathbf{x}}$ is the incidence matrix of the bipartite graph $G_{\mathbf{x}}$, it is a totally unimodular matrix. By the way full dimensional representations are chosen, the matrix $(\mathbf{A}_{1,\mathbf{x}})_{f_s}$ is invertible and hence its determinant squared equals 1. For $(\mathbf{A}_1)_{f_s}$, observe that $(\mathbf{A}_1)_{f_s}$ has the following block decomposition

(for some appropriate matrix \mathbf{M} which we do not need to specify explicitly):

$$(\mathbf{A}_1)_{f_s} = \begin{bmatrix} (\mathbf{A}_{1,\alpha\beta})_{f_s} & \mathbf{0} \\ \mathbf{M} & (\mathbf{A}_{1,\mathbf{x}})_{f_s} \end{bmatrix}, \text{ so that } \text{Det}((\mathbf{A}_1)_{f_s}) = \text{Det}((\mathbf{A}_{1,\alpha\beta})_{f_s}) \text{Det}((\mathbf{A}_{1,\mathbf{x}})_{f_s}).$$

Since $\mathbf{A}_{1,\alpha\beta}$, $\mathbf{A}_{1,\mathbf{x}}$ are totally unimodular, any invertible submatrix of them has determinant ± 1 . This concludes the proof of (191).

We next turn to (192). We begin with $L((\mathbf{A}_{1,\mathbf{x}})_f, \mathbf{A}_{1,\mathbf{x}}, \mathbf{T}_{1,\mathbf{x}})$. The argument is closely related to the proof of Kirchoff's Matrix-Tree Theorem, but is written in a way that it easily extends to the more complicated $L((\mathbf{A}_1)_f, \mathbf{A}_1, \mathbf{T}_1)$.

Denote by μ_1, \dots, μ_{2q-1} the nonzero eigenvalues of $\mathbf{A}_{1,\mathbf{x}}\mathbf{A}_{1,\mathbf{x}}^\top$; there are exactly $2q-1$ of those since $G_{\mathbf{x}}$ is a connected bipartite graph. Moreover, $\mathbf{v}_0^\top = \frac{1}{\sqrt{2q}}[-\mathbf{1}_q \ \mathbf{1}_q]$ is the unit eigenvector of $\mathbf{A}_{1,\mathbf{x}}\mathbf{A}_{1,\mathbf{x}}^\top$ with eigenvalue 0. We claim that

$$[\varepsilon]\text{Det}(\varepsilon\mathbf{T}_{1,\mathbf{x}} - \mathbf{A}_{1,\mathbf{x}}\mathbf{A}_{1,\mathbf{x}}^\top) = -\frac{\prod_{i \in [2q-1]} \mu_i}{q}, \quad \text{Det}((\mathbf{A}_{1,\mathbf{x}})_f (\mathbf{A}_{1,\mathbf{x}})_f^\top) = \frac{\prod_{i \in [2q-1]} \mu_i}{2q}, \quad (195)$$

which yields that $L((\mathbf{A}_{1,\mathbf{x}})_f, \mathbf{A}_{1,\mathbf{x}}, \mathbf{T}_{1,\mathbf{x}}) = 1/2$, as wanted. The first equality is a direct application of Lemma B.5, after observing that $\mathbf{v}_0^\top \mathbf{T}_{1,\mathbf{x}} \mathbf{v}_0 = 1/q$. The second can be proved as follows. The matrix $(\mathbf{A}_{1,\mathbf{x}})_f (\mathbf{A}_{1,\mathbf{x}})_f^\top$ is a principal minor of $\mathbf{A}_{1,\mathbf{x}}\mathbf{A}_{1,\mathbf{x}}^\top$, the specific principal minor is clearly determined by which row of \mathbf{A}_1 we chose to delete to obtain $(\mathbf{A}_1)_f$. Since $\mathbf{A}_{1,\mathbf{x}}\mathbf{A}_{1,\mathbf{x}}^\top$ has exactly one zero eigenvalue, we have

$$\prod_{i \in [2q-1]} \mu_i = \sum_{W \in \binom{[2q]}{2q-1}} \text{Det}((\mathbf{A}_{1,\mathbf{x}})_W (\mathbf{A}_{1,\mathbf{x}})_W^\top), \quad (196)$$

where $(\mathbf{A}_{1,\mathbf{x}})_W$ is the submatrix of $\mathbf{A}_{1,\mathbf{x}}$ induced by the rows indexed with W . It is easily checked that for any $W, W' \in \binom{[2q]}{2q-1}$, there exists a unitary matrix \mathbf{P} such that $(\mathbf{A}_{1,\mathbf{x}})_W = \mathbf{P}(\mathbf{A}_{1,\mathbf{x}})_{W'}$, so that all summands in (196) are equal. Indeed, since $\mathbf{A}_{1,\mathbf{x}}$ corresponds to the incidence matrix of a bipartite graph, the sum of the first q rows (as vectors) equals the sum of the last q rows. It follows that any row of $\mathbf{A}_{1,\mathbf{x}}$ can be expressed as a $\{1, -1\}$ linear combination of the remaining rows, which easily yields the existence of \mathbf{P} with the desired properties. Hence, for any $(\mathbf{A}_{1,\mathbf{x}})_f$ as in the statement of the lemma, the second equality in (195) holds as well.

We finally give a proof sketch for $L((\mathbf{A}_1)_f, \mathbf{A}_1, \mathbf{T}_1) = 1/2$. The matrix $\mathbf{A}_1\mathbf{A}_1^\top$ has zero as an eigenvalue by multiplicity one. Denote by $\sigma_1, \dots, \sigma_{2q+1}$ the nonzero eigenvalues of $\mathbf{A}_1\mathbf{A}_1^\top$. By looking at the space $\mathbf{z}\mathbf{A}_1 = \mathbf{0}$, it is easy to see that $\mathbf{v}_1 = \frac{1}{\sqrt{2(q+1)}}[-1, 1, -\mathbf{1}_q, \mathbf{1}_q]^\top$ is the unit length eigenvector for the eigenvalue 0. Moreover, the analog of (196) is

$$\prod_{i \in [2q+1]} \sigma_i = \sum_{W \in \binom{[2q+2]}{2q+1}} \text{Det}((\mathbf{A}_1)_W (\mathbf{A}_1)_W^\top). \quad (197)$$

Hence, the equality $L((\mathbf{A}_1)_f, \mathbf{A}_1, \mathbf{T}_1) = 1/2$ is obtained by the following analog of (195)

$$[\varepsilon]\text{Det}(\varepsilon\mathbf{T}_1 - \mathbf{A}_1\mathbf{A}_1^\top) = -\frac{\prod_{i \in [2q+1]} \sigma_i}{q+1}, \quad \text{Det}((\mathbf{A}_1)_f (\mathbf{A}_1)_f^\top) = \frac{\prod_{i \in [2q+2]} \sigma_i}{2(q+1)}. \quad \square$$

C. UNIQUENESS OF SEMITRANSLATION INVARIANT MEASURES (ANTIFERROMAGNETIC POTTS)

In this section, we prove Lemma 7.3. As noted earlier, the proof extends the respective argument in Brightwell and Winkler [2002] for colorings in the antiferromagnetic Potts

model setting. The technical details, due to the presence of the extra parameter B , are relatively more intricate.

PROOF OF LEMMA 7.3. Without loss of generality, we may assume that the scaling factors in (107) are equal to 1. We may also assume that $R_1 \geq \dots \geq R_q$. Since $B < 1$, (107) implies $C_1 \leq \dots \leq C_q$. Define

$$\alpha = \frac{R_1}{R_q}, \quad \beta = \frac{R_1 + \dots + R_{q-1}}{(q-1)R_q}, \quad S = R_1 + \dots + R_{q-1}.$$

We clearly have $\alpha \geq \beta \geq 1$, and we may assume for the sake of contradiction that $\beta > 1$. Note that

$$\begin{aligned} \alpha^{1/d} &= \left(\frac{R_1}{R_q}\right)^{1/d} = 1 + \frac{(1-B)(C_q - C_1)}{C_1 + \dots + C_{q-1} + BC_q} \\ C_q &= (R_1 + \dots + R_{q-1} + BR_q)^d = [(q-1)\beta + B]^d R_q^d \\ C_1 &= (BR_1 + R_2 + \dots + R_q)^d = [(q-1)\beta + 1 - (1-B)\alpha]^d R_q^d. \end{aligned}$$

Moreover, by Hölder's inequality, we have

$$\begin{aligned} C_1 + \dots + C_{q-1} + BC_q &= \sum_{i=1}^{q-1} [S + R_q - (1-B)R_i]^d + B(S + BR_q)^d \\ &\geq (q-1) \left[\frac{q-2+B}{q-1} S + R_q \right]^d + B(S + BR_q)^d \\ &= (q-1)[(q-2+B)\beta + 1]^d R_q^d + B[(q-1)\beta + B]^d R_q^d. \end{aligned}$$

Thus, we obtain that every solution must satisfy

$$\begin{aligned} \alpha^{1/d} &\leq 1 + \frac{(1-B)\{[(q-1)\beta + B]^d - [1 - (1-B)\alpha + (q-1)\beta]^d\}}{(q-1)[(q-2+B)\beta + 1]^d + B[(q-1)\beta + B]^d} \iff \\ 0 &\leq 1 - \alpha^{1/d} + \frac{(1-B) \left[1 - \left(1 - \frac{(1-B)(\alpha-1)}{(q-1)\beta+B} \right)^d \right]}{(q-1) \left[1 - \frac{(1-B)(\beta-1)}{(q-1)\beta+B} \right]^d + B} =: f(\alpha, \beta, B). \end{aligned}$$

To obtain a contradiction, our goal is to prove that for q and B as in the statement of the lemma, when $(q-1)\beta > \alpha \geq \beta > 1$, it holds that $f(\alpha, \beta, B) < 0$.

It is easy to see that f is decreasing in B . This immediately yields the lemma for $q \geq \Delta$: it holds that $f(\alpha, \beta, B) \leq f(\alpha, \beta, 0) < 0$, since the last inequality was proved by Brightwell and Winkler [2002]. For $q \leq d$ and $B \geq \frac{d+1-q}{d+1} := B_c$, this yields

$$f(\alpha, \beta, B) \leq f(\alpha, \beta, B_c) =: g(\alpha, \beta).$$

We first prove that $g(\alpha, \beta) \leq g(\beta, \beta)$. For $q = 2$, there is nothing to prove. Hence, we may assume that $d \geq q \geq 3$. Clearly, it suffices to prove that g is decreasing in α . This requires a fair bit of work, so we state it as a lemma to prove later.

LEMMA C.1. For $d \geq q \geq 3$ and $B_c = \frac{d+1-q}{d+1}$, the function $g(\alpha, \beta)$ is decreasing in α for $\alpha \geq \beta > 1$.

We finish the proof by showing that for $\beta \geq 1$, it holds that $g(\beta, \beta) \leq 0$ with equality iff $\beta = 1$. After massaging the inequality, this reduces to

$$1 \leq \left[1 - \frac{(1 - B_c)(\beta - 1)}{(q - 1)\beta + B_c} \right]^d [(q - 1)(\beta^{1/d} - 1) + 1 - B_c] + B_c \beta^{1/d} =: h(\beta).$$

Note that the inequality holds at equality for $\beta = 1$, so it suffices to prove $h'(\beta) > 0$ for $\beta > 1$, which is the assertion of the next lemma.

LEMMA C.2. *For $d \geq q \geq 3$ and $B_c = \frac{d+1-q}{d+1}$, the function $h(\beta)$ is increasing for $\beta \geq 1$.*

Modulo the proofs of Lemmas C.1 and C.2, which are given here, the proof is complete. \square

PROOF OF LEMMA C.1. We compute

$$\frac{\partial g}{\partial \alpha} = -\frac{1}{d} \alpha^{-(d-1)/d} + \frac{(1 - B_c)^2}{(q - 1)\beta + B_c} \cdot \frac{d \left[1 - \frac{(1 - B_c)(\alpha - 1)}{(q - 1)\beta + B_c} \right]^{d-1}}{(q - 1) \left[1 - \frac{(1 - B_c)(\beta - 1)}{(q - 1)\beta + B_c} \right]^d + B_c}.$$

Let $F(x) = x \left[1 - \frac{(1 - B_c)(x - 1)}{(q - 1)\beta + B_c} \right]^d$ for $x \in [\beta, (q - 1)\beta]$. We then have that $\frac{\partial g}{\partial \alpha} < 0$ is equivalent to

$$d^2(1 - B_c)^2 F(\alpha)^{(d-1)/d} \leq [(q - 1)\beta + B_c] \left[(q - 1) \left(1 - \frac{(1 - B_c)(\beta - 1)}{(q - 1)\beta + B_c} \right)^d + B_c \right]. \quad (198)$$

We first prove that $F(x)$ is decreasing in $[\beta, (q - 1)\beta]$. We calculate

$$F'(x) = \left[1 - \frac{(1 - B_c)(x - 1)}{(q - 1)\beta + B_c} \right]^{d-1} \frac{(q - 1)\beta + 1 - (d + 1)(1 - B_c)x}{(q - 1)\beta + B_c}.$$

For $x \in [\beta, (q - 1)\beta]$, we have $(d + 1)(1 - B_c)x = qx = (q - 1)x + x > (q - 1)\beta + 1$, where in the last inequality we used that $\beta > 1$. It follows that $F(x)$ is indeed decreasing and thus $F(\alpha) \leq F(\beta)$.

To prove (198), it thus suffices to argue that for $\beta > 1$ it holds

$$d^2(1 - B_c)^2 F(\beta)^{(d-1)/d} \leq [(q - 1)\beta + B_c] \left[(q - 1) \left(1 - \frac{(1 - B_c)(\beta - 1)}{(q - 1)\beta + B_c} \right)^d + B_c \right]. \quad (199)$$

Note that $q - 1 + B_c = d(1 - B_c)$ so that the inequality is tight for $\beta = 1$. By the weighted AM-GM inequality on A^d and 1 with weights $(q - 1)$ and B_c respectively, we obtain

$$(q - 1)A^d + B_c \geq (q - 1 + B_c)A^{d(q-1)/(q-1+B_c)} = d(1 - B_c)A^{(q-1)(d+1)/q}.$$

We use this for $A = 1 - \frac{(1 - B_c)(\beta - 1)}{(q - 1)\beta + B_c}$ so that, after simplifications, it suffices to show that

$$d(1 - B_c)\beta^{(d-1)/d} \leq [(q - 1)\beta + B_c] \left[1 - \frac{(1 - B_c)(\beta - 1)}{(q - 1)\beta + B_c} \right]^{-(d+1-2q)/q}.$$

This can further be massaged into

$$G(\beta) := \beta^{(d-1)/d} [(q - 1)\beta + B_c]^{-(d+1-q)/q} [(q - 2 + B_c)\beta + 1]^{(d+1-2q)/q} \leq \frac{1}{d(1 - B_c)}.$$

Once again, note that the inequality holds at equality for $\beta = 1$, so it suffices to prove that $G'(\beta) < 0$ for $\beta > 1$. This has nothing special, apart from tedious, but otherwise

straightforward, calculations. We include the details briefly. Differentiating $\ln G(\beta)$, we obtain

$$\frac{G'(\beta)}{G(\beta)} = \frac{(d-1)}{d\beta} - \frac{(d+1-q)(q-1)}{q[(q-1)\beta + B_c]} + \frac{(d+1-2q)(q-2+B_c)}{q[(q-2+B_c)\beta + 1]}.$$

By clearing denominators, it suffices to check that the following second-order polynomial $p(\beta)$ is negative whenever $\beta > 1$:

$$\begin{aligned} p(\beta) := & (d-1)q((q-1)\beta + B_c)((q-2+B_c)\beta + 1) \\ & - d(d+1-q)(q-1)\beta((q-2+B_c)\beta + 1) \\ & + d(d+1-2q)(q-2+B_c)\beta((q-1)\beta + B_c). \end{aligned}$$

Using again that $q-1+B_c = d(1-B_c)$, we obtain

$$\begin{aligned} p(1) = & (d-1)q(q-1+B_c)^2 - d(d+1-q)(q-1)(q-1+B_c) \\ & + d(d+1-2q)(q-2+B_c)(q-1+B_c) \\ = & d(q-1+B_c)[(d-1)q(1-B_c) - (d+1-q)(q-1) + (d+1-2q)(d(1-B_c)-1)] \\ = & d(q-1+B_c)[(d+1)(d-q)(1-B_c) - (d-q)q] \\ = & 0. \end{aligned}$$

The factorization of $p(\beta)$ (using the value of B_c) is given by

$$p(\beta) = -\frac{q(\beta-1)[\beta(d(q-1)^2 - (q-1)) + d(d-q) + q-1]}{d+1},$$

which is obviously negative for $\beta > 1$, whenever $d \geq q \geq 3$. \square

PROOF OF LEMMA C.2. We compute

$$\begin{aligned} h'(\beta) = & \frac{1}{d}\beta^{-(d-1)/d} \left[(q-1) \left(1 - \frac{(1-B_c)(\beta-1)}{(q-1)\beta + B_c} \right)^d + B_c \right] \\ & - d \left[1 - \frac{(1-B_c)(\beta-1)}{(q-1)\beta + B_c} \right]^{d-1} \frac{(1-B_c)(q-1+B_c)[(q-1)\beta^{1/d} - (q-2+B_c)]}{[(q-1)\beta + B_c]^2}. \end{aligned}$$

Thus, to prove $h'(\beta) > 0$, it suffices to check (using $q-1+B_c = d(1-B_c)$ and the function F defined in Lemma C.1).

$$\begin{aligned} & d^3(1-B_c)^2 F(\beta)^{(d-1)/d} \\ & \leq \frac{[(q-1)\beta + B_c]^2}{(q-1)\beta^{1/d} - (q-2+B_c)} \left[(q-1) \left(1 - \frac{(1-B_c)(\beta-1)}{(q-1)\beta + B_c} \right)^d + B_c \right]. \end{aligned}$$

This is similar to (199) and in fact follows from (199), once we prove that

$$\frac{(q-1)\beta^{1/d} - (q-2+B_c)}{(q-1)\beta + B_c} \leq \frac{1}{d}.$$

To see the last inequality, observe that $\beta+d-1 \geq d\beta^{1/d}$ as a consequence of the AM-GM inequality. Hence,

$$\frac{(q-1)\beta^{1/d} - (q-2+B_c)}{(q-1)\beta + B_c} \leq \frac{(q-1)\beta + (d-1)(q-1) - d(q-2+B_c)}{d[(q-1)\beta + B_c]} = \frac{1}{d},$$

completing the proof. \square

D. REMAINING PROOFS

D.1. Proof of Lemma B.1

In this section, we give the proof of Lemma B.1 which was used to derive the asymptotics of the ratio of the second moment to the square of the first moment for the distribution \mathcal{G}_n^r by the asymptotics of the corresponding ratio for the distribution \mathcal{G}_n .

We first recall some relevant definitions. Recall that a dominant phase $\mathbf{p} = (\alpha, \beta)$ corresponds to a fixpoint $(R_1, \dots, R_q, C_1, \dots, C_q)$ of the tree recursions (22). In particular, from Theorem 4.1, we have that

$$\alpha_i = \frac{R_i^{\Delta/(\Delta-1)}}{\sum_i R_i^{\Delta/(\Delta-1)}} \quad \text{and} \quad \beta_j = \frac{C_j^{\Delta/(\Delta-1)}}{\sum_j C_j^{\Delta/(\Delta-1)}}, \quad (23)$$

where the R_i 's and C_j 's satisfy

$$R_i = u \left(\sum_{j=1}^q B_{ij} C_j \right)^{\Delta-1} \quad \text{and} \quad C_j = v \left(\sum_{i=1}^q B_{ij} R_i \right)^{\Delta-1} \quad (22)$$

for some scaling factors $u, v > 0$. As we have mentioned before (and can readily be checked), the precise values of the scaling factors u, v in (22) do not matter for the correspondence (23) since scaling all the R_i 's or all the C_j 's by a positive factor leaves invariant the values of α_i 's and β_j 's in (23). For the purposes of this section however it will be easier to have a handle on the scaling factors, which is the reason u, v are explicitly included in (22).

For the purposes of this section, we will assume that the R_i 's and C_j 's in (22) further satisfy

$$\sum_i R_i = \sum_j C_j = 1, \quad (200)$$

this can be achieved by scaling the R_i 's and C_j 's appropriately. In particular, if we set $\hat{R}_i = R_i / (\sum_i R_i)$ and $\hat{C}_j = C_j / (\sum_j C_j)$, we obtain that the \hat{R}_i 's and \hat{C}_j 's also satisfy the tree recursions (22) (possibly with different values for the scaling factors u, v).

Recall from Section 6.4 that for a configuration $\eta : W \rightarrow [q]$, the product measure $\nu_{\mathbf{p}}^{\otimes}(\eta)$ was defined as

$$\nu_{\mathbf{p}}^{\otimes}(\eta) = \prod_{i \in [q]} (R_i)^{|\eta^{-1}(i) \cap W^+|} \prod_{j \in [q]} (C_j)^{|\eta^{-1}(j) \cap W^-|}. \quad (96)$$

The main lemma to prove Lemma B.1 will be the following.

LEMMA D.1. *Let $\mathbf{p} = (\alpha, \beta) \in \mathcal{Q}$ be a Hessian dominant phase. For every fixed $r > 0$, for every $\eta : W \rightarrow [q]$, it holds that*

$$\lim_{n \rightarrow \infty} \frac{\mathbf{E}_{\mathcal{G}_n^r} [Z_G^{\mathbf{p}}(\eta)]}{\mathbf{E}_{\mathcal{G}_n} [Z_G^{\mathbf{p}}]} = C^r \nu_{\mathbf{p}}^{\otimes}(\eta), \quad (201)$$

where the constant $C = C(\mathbf{p})$ is given in terms of the fixpoint of (22) corresponding to the phase \mathbf{p} and equals

$$C = \exp(\Phi(\mathbf{r}, \mathbf{c})) \left(\frac{\sum_{i,j} B_{ij} R_i C_j}{\sum_{i,j} R_i C_j} \right)^{\Delta-2}, \quad (202)$$

where in the expression for C , Φ denotes the function of $\mathbf{r} = (R_1, \dots, R_q)$ and $\mathbf{c} = (C_1, \dots, C_q)$ defined in Section 3.2.

We remark here that while (200) is obviously important for the definition of the measure $\nu_{\mathbf{p}}^{\otimes(\cdot)}$ (to have a valid probability distribution), it is not important for the expression (202). More precisely, (analogously to (23)) the particular scaling of the R_i 's and C_j 's in the expression (202) does not matter, as it can readily be checked. From (202), it is also simple to see that the value of $C(\mathbf{p})$ does not depend on \mathbf{p} or η when the phases \mathbf{p} are assumed to be permutation symmetric (this is relevant for Lemma 6.11).

LEMMA D.2. *If $\mathbf{p} = (\alpha, \beta)$ is a Hessian dominant phase, it holds that*

$$\lim_{n \rightarrow \infty} \frac{\mathbf{E}_{\mathcal{G}_n}[(Z_G^{\mathbf{p}}(\eta))^2]}{\mathbf{E}_{\mathcal{G}_n}[(Z_G^{\mathbf{p}})^2]} = C^{2r} (\nu_{\mathbf{p}}^{\otimes(\eta)})^2,$$

where C is the same constant as in (202).

Using Lemmas D.1 and D.2, the proof of Lemma B.1 is immediate.

PROOF OF LEMMA B.1. By Lemma D.1, we have

$$\mathbf{E}_{\mathcal{G}_n}[Z_G^{\mathbf{p}}(\eta)] = (1 + o(1))C^r \nu_{\mathbf{p}}^{\otimes(\eta)} \mathbf{E}_{\mathcal{G}_n}[Z_G^{\mathbf{p}}], \quad (102)$$

where $C = C(\mathbf{p})$ is the constant in (202) and $\nu_{\mathbf{p}}^{\otimes(\eta)}$ is defined in (96). By Lemma D.2, we also obtain

$$\mathbf{E}_{\mathcal{G}_n}[(Z_G^{\mathbf{p}}(\eta))^2] = (1 + o(1))C^{2r} (\nu_{\mathbf{p}}^{\otimes(\eta)})^2 \mathbf{E}_{\mathcal{G}_n}[(Z_G^{\mathbf{p}})^2], \quad (203)$$

where $C = C(\mathbf{p})$ is again the constant in (202). Combining (102) and (203) proves the lemma. \square

We conclude by giving the proofs of Lemmas D.1 and D.2.

PROOF OF LEMMA D.1. Fix $\eta : W \rightarrow [q]$. Recall that $W = W^+ \cup W^-$, where W^+ and W^- are the vertices of degree $\Delta - 1$ in the $+$ and $-$ parts of a graph $G \sim \mathcal{G}_n^r$ and it holds that $|W^+| = |W^-| = r$. For $i \in [q]$, let $\eta_i^{\pm} = |\eta^{-1}(i) \cap W^{\pm}|$.

Recall that, with Σ_1 as in (162) and $P_1 = \{(i, j) \mid B_{ij} > 0\}$ (see (163)),

$$\begin{aligned} \mathbf{E}_{\mathcal{G}_n}[Z_G^{\mathbf{p}}] &= \sum_{\alpha', \beta' \in \Sigma_1} \binom{n}{\alpha'_1 n, \dots, \alpha'_q n} \binom{n}{\beta'_1 n, \dots, \beta'_q n} \\ &\quad \times \left(\sum_{\mathbf{x}'} \frac{\prod_i \binom{\alpha'_i n}{x'_{i1}, \dots, x'_{iq}} \prod_j \binom{\beta'_j n}{x'_{j1}, \dots, x'_{jq}} \prod_{(i,j) \in P_1} B_{ij}^{nx'_{ij}}}{\binom{n}{x'_{11}, \dots, x'_{qq}}} \right)^{\Delta}, \end{aligned} \quad (4)$$

where the sum ranges over nonnegative $\alpha' = (\alpha'_1, \dots, \alpha'_q)$, $\beta' = (\beta'_1, \dots, \beta'_q)$ and $\mathbf{x}' = (x'_{11}, \dots, x'_{qq})$ satisfying $x'_{ij} = 0$ whenever $(i, j) \in P_1$ and the following constraints:

$$\begin{aligned} \sum_i \alpha'_i &= 1, & \sum_j \beta'_j &= 1, \\ \sum_j x'_{ij} &= \alpha'_i \quad (\forall i \in [q]), & \sum_i x'_{ij} &= \beta'_j \quad (\forall j \in [q]). \end{aligned} \quad (5)$$

Analogously, we have the following expression for $\mathbf{E}_{\mathcal{G}_n^r}[\mathbf{Z}_G^{\mathbf{p}}(\eta)]$:

$$\begin{aligned} \mathbf{E}_{\mathcal{G}_n^r}[\mathbf{Z}_G^{\mathbf{p}}(\eta)] &= \sum_{\alpha', \beta' \in \Sigma_1} \binom{n}{\alpha'_1 n, \dots, \alpha'_q n} \binom{n}{\beta'_1 n, \dots, \beta'_q n} \\ &\times \left(\sum_{\mathbf{x}''} \frac{\prod_i \binom{\alpha'_i n + \eta_i^+}{x''_{i1} n, \dots, x''_{iq} n} \binom{\beta'_j n + \eta_j^-}{x''_{j1} n, \dots, x''_{jq} n} \prod_{(i,j) \in P_1} B_{ij}^{x''_{ij}}}{\binom{n+r}{x''_{11} n, \dots, x''_{qq} n}} \right)^{\Delta-1} \\ &\times \left(\sum_{\mathbf{x}'} \frac{\prod_i \binom{\alpha'_i n}{x'_{i1} n, \dots, x'_{iq} n} \prod_j \binom{\beta'_j n}{x'_{j1} n, \dots, x'_{jq} n} \prod_{(i,j) \in P_1} B_{ij}^{x'_{ij}}}{\binom{n}{x'_{11} n, \dots, x'_{qq} n}} \right), \end{aligned} \quad (204)$$

where the summation ranges over non-negative $\alpha', \beta', \mathbf{x}', \mathbf{x}''$ satisfying $x'_{ij} = x''_{ij} = 0$ whenever $(i, j) \in P_1$ and, in addition to (5), the following equalities:

$$\sum_j x''_{ij} = \alpha'_i + \eta_i^+ / n \quad (\forall i \in [q]), \quad \sum_i x''_{ij} = \beta'_j + \eta_j^- / n \quad (\forall j \in [q]). \quad (205)$$

Let $\mathbf{x} = \arg \max_{\mathbf{x}'} \Upsilon_1(\alpha, \beta, \mathbf{x}')$ where Υ_1 is given in (8), that is, \mathbf{x} captures the terms in (4) whose contribution determines the order of $\frac{1}{n} \log \mathbf{E}_{\mathcal{G}}[\mathbf{Z}_G^{\alpha', \beta'}]$ for $(\alpha', \beta') = (\alpha, \beta)$. Recall from Lemma 4.3 that the vector \mathbf{x} is unique and its value is given in Lemma 4.3. Since $\mathbf{p} = (\alpha, \beta)$ is the unique dominant phase contained in Σ_1 , for every constant $\delta > 0$ terms in (4) with $\|(\alpha', \beta', \mathbf{x}') - (\alpha, \beta, \mathbf{x})\|_2 \geq \delta$ have exponentially small contribution and we will thus ignore them henceforth. Similarly, for every constant $\delta > 0$, terms in (204) with $\|(\alpha', \beta', \mathbf{x}', \mathbf{x}'') - (\alpha, \beta, \mathbf{x}, \mathbf{x})\|_2 \geq \delta$ have exponentially small contribution and we will ignore such terms as well. In particular, we consider δ sufficiently small so that $\|(\alpha', \beta', \mathbf{x}', \mathbf{x}'') - (\alpha, \beta, \mathbf{x}, \mathbf{x})\|_2 < \delta$ implies $\alpha'_i > 0, \beta'_j > 0$ for $i, j \in [q]$ and $x'_{ij} > 0, x''_{ij} > 0$ for $(i, j) \in P_1$.

Let $\boldsymbol{\rho} = (\rho_{11}, \dots, \rho_{qq})$ be a (fixed) q^2 -dimensional vector with integer entries satisfying

$$\sum_j \rho_{ij} = \eta_i^+ \quad (\forall i \in [q]), \quad \sum_i \rho_{ij} = \eta_j^- \quad (\forall j \in [q]), \quad (206)$$

and $\rho_{ij} = 0$ whenever $(i, j) \in P_1$. Note that such a vector exists because of ergodicity of \mathbf{B} . We decompose vectors \mathbf{x}'' in (205) as $\mathbf{x}'' = \mathbf{x}' + \boldsymbol{\rho}/n$ where \mathbf{x}' satisfies (5). Standard approximations of binomial coefficients (see, e.g., Galanis et al. [2012, Lemma 28]) yield for $i, j \in [q]$:

$$\begin{aligned} \frac{\binom{\alpha_i n + \eta_i^+}{x_{i1} n + \rho_{i1}, \dots, x_{iq} n + \rho_{iq}}}{\binom{\alpha_i n}{x_{i1} n, \dots, x_{iq} n}} &\sim \frac{\alpha_i^{\eta_i^+}}{\prod_{j: (i,j) \in P_1} x_{ij}^{\rho_{ij}}}, & \frac{\binom{\beta_j n + \eta_j^-}{x_{j1} n + \rho_{j1}, \dots, x_{jq} n + \rho_{jq}}}{\binom{\beta_j n}{x_{j1} n, \dots, x_{jq} n}} &\sim \frac{\beta_j^{\eta_j^-}}{\prod_{i: (i,j) \in P_1} x_{ij}^{\rho_{ij}}}, \\ \frac{\binom{n+r}{x_{11} n + \rho_{11}, \dots, x_{qq} n + \rho_{qq}}}{\binom{n}{x_{11} n, \dots, x_{qq} n}} &\sim \frac{1}{\prod_{(i,j) \in P_1} x_{ij}^{\rho_{ij}}}. \end{aligned}$$

By considering the limit $\delta \downarrow 0$, we thus obtain

$$\frac{\mathbf{E}_{\mathcal{G}_n^r}[\mathbf{Z}_G^{\mathbf{p}}(\eta)]}{\mathbf{E}_{\mathcal{G}_n}[\mathbf{Z}_G^{\mathbf{p}}]} \sim \left(\frac{\prod_{i \in [q]} \alpha_i^{\eta_i^+} \prod_{j \in [q]} \beta_j^{\eta_j^-}}{\prod_{(i,j) \in P_1} (x_{ij} / B_{ij})^{\rho_{ij}}} \right)^{\Delta-1} =: \mathbf{E}.$$

It remains to massage the expression for E . From the correspondence given in (23) and the fact that the R_i 's and C_j 's satisfy (22), we obtain that for all $i, j \in [q]$, it holds that

$$\alpha_i = \frac{R_i(\sum_j B_{ij}C_j)}{\sum_{i,j} B_{ij}R_iC_j}, \quad \beta_j = \frac{C_j(\sum_i B_{ij}R_i)}{\sum_{i,j} B_{ij}R_iC_j}, \quad x_{ij} = \frac{B_{ij}R_iC_j}{\sum_{i,j} B_{ij}R_iC_j}.$$

Plugging these into the expression for E and using (22) and (206), we obtain that

$$\begin{aligned} E &= \prod_{i \in [q]} \left(\sum_j B_{ij}C_j \right)^{(\Delta-1)\eta_i^+} \prod_{j \in [q]} \left(\sum_i B_{ij}R_i \right)^{\eta_j^-} = \frac{1}{(uv)^r} \prod_{i \in [q]} R_i^{\eta_i^+} \prod_{j \in [q]} C_j^{\eta_j^-} \\ &= \frac{1}{(uv)^r} v_{\mathbf{p}}^{\otimes}(\eta), \end{aligned}$$

where, in the last equality, we used the definition (96) of $v_{\mathbf{p}}^{\otimes}(\eta)$. This yields the statement of the lemma with $C = 1/(uv)$. To massage the latter into (202), observe that

$$\begin{aligned} \frac{1}{uv} &= \frac{(\sum_{i,j} B_{ij}R_iC_j)^{2(\Delta-1)}}{\left(\sum_i R_i^{\Delta/\Delta-1}\right)^{\Delta-1} \left(\sum_j C_j^{\Delta/\Delta-1}\right)^{\Delta-1}} \\ &= \frac{(\sum_{i,j} B_{ij}R_iC_j)^{2(\Delta-1)}}{\left(\sum_i R_i^{\Delta/\Delta-1}\right)^{\Delta-1} \left(\sum_j C_j^{\Delta/\Delta-1}\right)^{\Delta-1} (\sum_i R_i)^{\Delta-2} (\sum_j C_j)^{\Delta-2}}, \end{aligned} \quad (207)$$

where the first equality is a consequence of the tree recursions (22), while the second equality follows from our assumption that the R_i 's and C_j 's satisfy (200). The last expression is equal to the right-hand side in (202) which concludes the proof of Lemma D.1. \square

PROOF OF LEMMA D.2. Recall from the first part of Lemma B.2 that if $\mathbf{p} = (\boldsymbol{\alpha}, \boldsymbol{\beta}) \in \mathcal{Q}$ is a dominant phase for the spin system with interaction matrix \mathbf{B} , then $\mathbf{p}' = (\boldsymbol{\alpha} \otimes \boldsymbol{\alpha}, \boldsymbol{\beta} \otimes \boldsymbol{\beta})$ is a dominant phase for the spin system with interaction matrix $\mathbf{B} \otimes \mathbf{B}$. The corresponding fixpoint of the tree recursions for $\mathbf{B} \otimes \mathbf{B}$ is given by R_{ik} 's and C_{jl} 's which satisfy for all $i, j, k, l \in [q]$,

$$R_{ik} \propto R_i R_k, \quad C_{jl} \propto C_j C_l. \quad (208)$$

Note that (208) follows from Lemma 3.2 and Theorem 4.1.

Using a completely analogous argument to the one used in the proof of the second part of Lemma B.2 (only slightly more care is needed in writing the range of $\gamma, \delta, \mathbf{y}$, analogously to (205)), we obtain the following. For $G \sim \mathcal{G}_n^r$ and $\eta : W \rightarrow [q]$, let $Z_G^{\mathbf{p}'}(\eta^{\otimes 2})$ be the contribution to the partition function of G for the spin system with interaction matrix $\mathbf{B} \otimes \mathbf{B}$ from configurations σ with $Y(\sigma) = \mathbf{p}'$ and $\sigma_W = \eta^{\otimes 2}$, where the configuration $\eta^{\otimes 2}$ is given by $\eta^{\otimes 2}(v) = (\eta(v), \eta(v))$ for $v \in W$. Then, $\lim_{n \rightarrow \infty} \frac{\mathbf{E}_{\mathcal{G}_n^r}[Z_G^{\mathbf{p}'}(\eta^{\otimes 2})^2]}{\mathbf{E}_{\mathcal{G}_n^r}[Z_G^{\mathbf{p}'}(\eta^{\otimes 2})]} = 1$.

Using Lemma D.1 for the spin system with interaction matrix $\mathbf{B} \otimes \mathbf{B}$ and dominant phase $\mathbf{p}' = (\boldsymbol{\alpha} \otimes \boldsymbol{\alpha}, \boldsymbol{\beta} \otimes \boldsymbol{\beta})$, we obtain that

$$\lim_{n \rightarrow \infty} \frac{\mathbf{E}_{\mathcal{G}_n^r}[Z_G^{\mathbf{p}'}(\eta^{\otimes 2})]}{\mathbf{E}_{\mathcal{G}_n^r}[Z_G^{\mathbf{p}'}]} = C(\mathbf{p}') v_{\mathbf{p}'}^{\otimes}(\eta^{\otimes 2}) = (C(\mathbf{p}'))^2 (v_{\mathbf{p}}^{\otimes}(\eta))^2,$$

where the latter equality is a consequence of (208) and the value (202) of the constant $C(\mathbf{p})$. Finally, it remains to note that, from Lemma B.2, it holds that $\lim_{n \rightarrow \infty} \frac{\mathbf{E}_{\mathcal{G}_n}[(Z_G^{\mathbf{p}})^2]}{\mathbf{E}_{\mathcal{G}_n}[Z_G^{\mathbf{p}}]} = 1$.

Combining these asymptotic statements yields the statement of the lemma. \square

D.2. Proof of Lemma A.6

In Section A.3, we gave the proof of Lemma A.6 for the graph distribution \mathcal{G}_n and the random variables $Z_G^{\alpha, \beta}$. In this section, we extend this argument to the graph distribution \mathcal{G}_n^r and the random variables $Z_G^{\mathbf{p}}$ (see Section 6.4 for the relevant definitions) for some fixed $r > 0$. The argument is essentially the same as the one given in Section A.3, modulo some technical details in the calculations, analogous to those given in the proof of Lemma D.1. For completeness, we give these details. All probabilities and expectations in the following argument will refer to the graph distribution \mathcal{G}_n^r .

In the setting of Lemma A.6, we want to show

$$\frac{\mathbf{E}[Z_G^{\mathbf{p}}(\eta)X_\ell]}{\mathbf{E}[Z_G^{\mathbf{p}}(\eta)]} \rightarrow \mu_\ell(1 + \delta_\ell) \text{ as } n \rightarrow \infty, \quad (209)$$

where X_ℓ denotes the number of cycles of length ℓ in \mathcal{G}_n^r and μ_ℓ, δ_ℓ are given in Lemmas A.5 and A.6, respectively (recall that ℓ is fixed with respect to n and that we are only interested in even integers ℓ since a graph $G \sim \mathcal{G}_n^r$ is bipartite).

We follow the argument given in Section A.3 as closely as possible. We again decompose X_ℓ as

$$X_\ell = \frac{1}{2\ell} \sum_{\xi} \sum_{\zeta} \mathbf{1}_{\xi, \zeta}, \quad (157)$$

where recall that, roughly, ξ specifies which of the Δ matchings (used to sample $G \sim \mathcal{G}_n^r$) the edges of an ℓ -cycle belong to as well as the spin assignment of the vertices on the cycle, while ζ specifies the vertices in $U \cup W$ that an ℓ -cycle traverses, in order.

More precisely, ξ denotes a proper Δ -edge colored, rooted and oriented ℓ -cycle ($r(\Delta, \ell)$ possibilities, see Lemma A.5), whose vertices are colored with $\{Y_1, \dots, Y_q, G_1, \dots, G_q\}$ and edges are colored with $\{1, \dots, \Delta\}$, see Section A.3 for details on the vertex coloring of the vertices in the cycle. Recall that y_i is the number of vertices colored with Y_i and g_j is the number of vertices colored with G_j . Recall also that for $k \in [\Delta]$ and $i, j \in [q]$, $a_{ij}(k)$ denotes the number of edges whose color is k and their endpoints are assigned colors Y_i, G_j . With $a_{ij} := \sum_{k=1}^{\Delta} a_{ij}(k)$, we have the following:

$$\sum_j a_{ij} = 2y_i, \quad \sum_i a_{ij} = 2g_j, \quad \sum_{i,j} a_{ij} = 2\ell. \quad (159)$$

Given such a ξ , for a specification ζ denote by u_i^+ the number of vertices in the cycle with color Y_i that belong to U^+ and define analogously u_j^- for U^- . Note we have that $u_i^+ \leq y_i$ and $u_j^- \leq g_j$, so (159) implies that

$$\sum_i u_i^+ + \sum_j u_j^- \leq \ell \text{ with equality iff } u_i^+ = y_i \text{ and } u_j^- = g_j \text{ for all } i, j \in [q]. \quad (210)$$

We will show that for such ξ and ζ , it holds that

$$\frac{\mathbf{E}[Z_G^{\mathbf{p}}(\eta) \mathbf{1}_{\xi, \zeta}]}{\mathbf{E}[Z_G^{\mathbf{p}}(\eta)]} \sim \prod_i \alpha_i^{u_i^+} \prod_j \beta_j^{u_j^-} \frac{\prod_{(i,j) \in P_1} x_{ij}^{\alpha_{ij}}}{\prod_i \alpha_i^{\sum_j a_{ij}} \prod_j \beta_j^{\sum_i a_{ij}}}. \quad (211)$$

Let us assume (211) for the moment and conclude the proof of (209). Set $p_1 := \Pr[\mathbf{1}_{\xi, \zeta} = 1]$ and note that $p_1 \sim n^{-\ell}$. We have that

$$\frac{\mathbf{E}[Z_G^{\mathbf{p}}(\eta)X_\ell]}{\mathbf{E}[Z_G^{\mathbf{p}}(\eta)]} = \frac{1}{2\ell} \sum_{\xi} \sum_{\zeta} p_1 \cdot \frac{\mathbf{E}[Z_G^{\mathbf{p}}(\eta) | \mathbf{1}_{\xi, \zeta} = 1]}{\mathbf{E}[Z_G^{\mathbf{p}}(\eta)]}. \quad (212)$$

Now for a given ξ , we may ignore those ζ such that $\sum_i u_i^+ + \sum_j u_j^- < \ell$, since the number of those ζ is $o(n^\ell)$ and hence, using (211), their contribution in (212) is $o(1)$. From (210), we only need to consider ζ such that $u_i^+ = y_i$ and $u_j^- = g_j$ for all $i, j \in [q]$. (We note here that this part of the argument displays that we could have chosen X_ℓ to be the number of cycles whose vertices consist solely of vertices in $U^+ \cup U^-$ as done in Sly [2010].) The number of such ζ is asymptotically n^ℓ . Plugging $u_i^+ = y_i$ and $u_j^- = g_j$ in (211), we obtain, using also (159), that

$$\frac{\mathbf{E}[Z_G^{\mathbf{p}}(\eta) | \mathbf{1}_{\xi, \zeta}]}{\mathbf{E}[Z_G^{\mathbf{p}}(\eta)]} \sim \prod_{i,j} \left(\frac{x_{i,j}}{\sqrt{\alpha_i \beta_j}} \right)^{a_{ij}}. \quad (213)$$

Combining this, we obtain

$$\frac{\mathbf{E}_{G_n^r}[Z_G^{\mathbf{p}}(\eta)X_\ell]}{\mathbf{E}_{G_n^r}[Z_G^{\mathbf{p}}(\eta)]} \sim \frac{r(\Delta, \ell)}{2\ell} \cdot \sum_{\mathbf{a}} N_{\mathbf{a}} \prod_{i,j} \left(\frac{x_{i,j}}{\sqrt{\alpha_i \beta_j}} \right)^{a_{ij}}, \quad (214)$$

where $\mathbf{a} = \{a_{11}, \dots, a_{qq}\}$ and $N_{\mathbf{a}}$ is the number of possible ξ with a_{ij} edges having assignment (Y_i, G_j) . Note that the right-hand side in (214) is the same as the right-hand side in (160), and hence the rest of the argument in Section A.3 carries over verbatim, yielding (209).

It remains to prove (211). We have already derived an expression for $\mathbf{E}[Z_G^{\mathbf{p}}(\eta)]$ in (204), so we focus on computing $\mathbf{E}[Z_G^{\mathbf{p}}(\eta) | \mathbf{1}_{\xi, \zeta} = 1]$. Set $u^\pm := \sum_{i \in [q]} u_i^\pm$ and recall that for $i \in [q]$, $\eta_i^\pm = |\eta^{-1}(i) \cap W^\pm|$. We have

$$\begin{aligned} \mathbf{E}[Z_G^{\mathbf{p}}(\eta) | \mathbf{1}_{\xi, \zeta}] &= \sum_{\alpha', \beta'} \binom{n - u^+}{\alpha'_1 n - u_1^+, \dots, \alpha'_q n - u_q^+} \binom{n - u^-}{\beta'_1 n - u_1^-, \dots, \beta'_q n - u_q^-} \\ &\times \prod_{k=1}^{\Delta-1} \left(\frac{\prod_i \binom{\alpha'_i n + \eta_i^+ - \sum_j a_{ij}(k)}{x''_{i1} n - a_{i1}(k), \dots, x''_{iq} n - a_{iq}(k)} \prod_j \binom{\beta'_j n + \eta_j^- - \sum_i a_{ij}(k)}{x''_{1j} n - a_{1j}(k), \dots, x''_{qj} n - a_{qj}(k)} \prod_{(i,j) \in P_1} B_{ij}^{n x''_{ij}}}{\binom{n + r - \sum_{i,j} a_{ij}(k)}{x''_{11} n - a_{11}(k), \dots, x''_{qq} n - a_{qq}(k)}} \right) \\ &\times \left(\frac{\prod_i \binom{\alpha'_i n - \sum_j a_{ij}(\Delta)}{x'_{i1} n - a_{i1}(\Delta), \dots, x'_{iq} n - a_{iq}(\Delta)} \prod_j \binom{\beta'_j n - \sum_i a_{ij}(\Delta)}{x'_{1j} n - a_{1j}(\Delta), \dots, x'_{qj} n - a_{qj}(\Delta)} \prod_{(i,j) \in P_1} B_{ij}^{n x'_{ij}}}{\binom{n - \sum_{i,j} a_{ij}(\Delta)}{x'_{11} n - a_{11}(\Delta), \dots, x'_{qq} n - a_{qq}(\Delta)}} \right), \end{aligned}$$

where $\alpha', \beta', \mathbf{x}', \mathbf{x}''$ range over the same values as in (204). Using asymptotic estimates for ratios of binomial coefficients (similar to those used in the proof of Lemma D.1), we obtain

$$\frac{\mathbf{E}_{G_n^r}[Z_G^{\mathbf{p}}(\eta)X_\ell]}{\mathbf{E}_{G_n^r}[Z_G^{\mathbf{p}}(\eta)]} \sim A \prod_{k=1}^{\Delta} M_k, \quad (215)$$

where $A := \prod_i \alpha_i^{u_i^+} \prod_j \beta_j^{u_j^-}$ and $M_k := \frac{\prod_{(i,j) \in P_1} x_{ij}^{a_{ij}^{(k)}}}{\prod_i \alpha_i^{\sum_j a_{ij}^{(k)}} \prod_j \beta_j^{\sum_i a_{ij}^{(k)}}}$. Note that

$$\prod_{k=1}^{\Delta} M_k = \frac{\prod_{(i,j) \in P_1} x_{ij}^{a_{ij}}}{\prod_i \alpha_i^{\sum_j a_{ij}} \prod_j \beta_j^{\sum_i a_{ij}}},$$

which gives (211), as desired.

This completes the proof of Lemma A.6.

D.3. Remaining Proof of Lemma 7.23

In this section, we show that the function

$$h(r_2) := r_2^{d+1} - \frac{(r_2^{d+1} - dr_2 + (d-1))^{d+1}}{(r_2^{d+1} - (d+1)r_2 + d)^d}$$

is strictly increasing for $r_2 > 1$ and $d \geq 2$. This monotonicity was used in the proof of Lemma 7.23.

We have

$$h'(r_2) = (d+1) \left(r_2^d - \frac{(r_2^{d+1} - dr_2 + (d-1))^d (r_2^{2d+1} - (2d+1)r_2^{d+1} + 2dr_2^d + dr_2 - d)}{(r_2^{d+1} - (d+1)r_2 + d)^{d+1}} \right).$$

Thus, to show $h'(r_2) > 0$ for $r_2 > 1$, it suffices to show that $w(r_2) < 1$, where

$$w(r_2) := \frac{(r_2^{d+1} - dr_2 + (d-1))^d (r_2^{2d+1} - (2d+1)r_2^{d+1} + 2dr_2^d + dr_2 - d)}{r_2^d (r_2^{d+1} - (d+1)r_2 + d)^{d+1}}.$$

By differentiating $\ln w(r_2)$, we have

$$\begin{aligned} & \frac{w'(r_2)}{w(r_2)} \\ &= \frac{d p(r_2)}{r_2 (r_2^{d+1} - dr_2 + d - 1) (r_2^{d+1} - (d+1)r_2 + d) (r_2^{2d+1} - (2d+1)r_2^{d+1} + 2dr_2^d + dr_2 - d)}, \end{aligned}$$

where

$$\begin{aligned} p(r_2) := & (d-1)d^2 - (d-1)(3d^2 + 2d+1)r_2 + d(3d^2 + d-1)r_2^2 - d^2(d+1)r_2^3 \\ & + (d-2)(2d+1)r_2^{d+1} - 2d(2d-1)r_2^{d+2} + d(2d+1)r_2^{d+3} \\ & + (d+1)^2 r_2^{2d+1} - d(2d+1)r_2^{2d+2} + d(d-1)r_2^{2d+3}. \end{aligned}$$

We will show that $p(r_2) > 0$ for $r_2 > 1$, which yields that w is strictly increasing in the interval $(1, +\infty)$ and thus $w(r_2) < w(+\infty) = 1$ for all $r_2 > 1$.

We have that $p(1) = p'(1) = p''(1) = p'''(1) = 0$, so $r_2 = 1$ is a root of $p(r_2)$ by multiplicity 4. Thus, if there exists $r_2 > 1$ such that $p(r_2) \leq 0$, by four applications of the Mean Value Theorem, there must exist $r_2 > 1$ such that $p''''(r_2) \leq 0$. For $d = 2$, we have

$$p''''(r_2) = 24(-12 + 95r_2 - 150r_2^2 + 70r_2^3),$$

which is positive for all $r_2 > 1$. We may thus focus on $d \geq 3$.

Note that $p''''(r_2) = d(d+1)r_2^{d-3} p_1(r_2)$, where

$$\begin{aligned} p_1(r_2) := & (d-2)^2(d-1)(2d+1) - 2(d-1)d(d+2)(2d-1)r_2 + d(d+2)(d+3)(2d+1)r_2^2 \\ & + 2(d+1)(2d-2)(2d-1)(2d+1)r_2^d - 4d(2d-1)(2d+1)^2r_2^{d+1} \\ & + 4(d-1)d(2d+1)(2d+3)r_2^{d+2}. \end{aligned}$$

It suffices to show that $p_1(r_2) > 0$ for $r_2 > 1$. We have that $p_1(1) = 6d(d-1) > 0$, so if there exists $r_2 > 1$ such that $p_1(r_2) \leq 0$, by the Mean Value Theorem, there exists $r_2 > 1$ such that $p_1'(r_2) < 0$. By iterating this process four more times, one obtains polynomials $p_i(r_2)$ for $i = 2, 3, 4, 5$ with $p_i(1) > 0$ and the following property. For $i = 2, 3, 4$, there exists $r_2 > 1$ with $p_i(r_2) < 0$ if there exists $r_2 > 1$ such that $p_i'(r_2) < 0$ if there exists $r_2 > 1$ such that $p_{i+1}(r_i) < 0$. The polynomials p_i are given by

$$\begin{aligned} p_2(r_2) = & -(d-1)(d+2)(2d-1) + (d+2)(d+3)(2d+1)r_2 \\ & + (d+1)(2d-2)(2d-1)(2d+1)r_2^{d-1} - 2(d+1)(2d-1)(2d+1)^2r_2^d \\ & + 2(d-1)(d+2)(2d+1)(2d+3)r_2^{d+1}, \\ p_3(r_2) = & (d+2)(d+3) + (d-1)(d+1)(2d-2)(2d-1)r_2^{d-2} \\ & - 2d(d+1)(2d-1)(2d+1)r_2^{d-1} + 2(d-1)(d+1)(d+2)(2d+3)r_2^d, \\ p_4(r_2) = & (d-1)(d-2)(2d-1) - d(2d-1)(2d+1)r_2 + d(d+2)(2d+3)r_2^2, \\ p_5(r_2) = & -(2d-1)(2d+1) + 2(d+2)(2d+3)r_2. \end{aligned}$$

We have that $p_5(r_2) > p_5(1) > 0$ for all $r_2 > 1$. It follows that $p(r_2) > 0$ for all $r_2 > 1$, thus concluding the proof that h is increasing in the interval $(1, +\infty)$.