

# *Assortative Mating Biases Marker-based Heritability Estimators*

## Supplementary Information

### Contents

<b>Supplementary Notes</b>	<b>2</b>
<b>1 Assortative mating and the joint distribution of causal variants; previous results</b>	<b>2</b>
1.1 Quantitative genetic model and notation . . . . .	3
1.2 Causal variant dynamics across generations . . . . .	5
<b>2 Assortative mating and the joint distribution of causal variants; novel results</b>	<b>10</b>
2.1 A thermodynamic approach . . . . .	10
2.1.1 Energy functions and maximum entropy . . . . .	10
2.1.2 The joint distribution of genotypes under simplifying assumptions . . . . .	11
2.1.3 The joint distribution of phenotypes under simplifying assumptions . . . . .	12
2.1.4 Stationarity . . . . .	13
2.1.5 Extension to the general case . . . . .	14
2.2 Moment derivations . . . . .	16
2.2.1 Deriving $\lambda$ from the first moment . . . . .	16
2.2.2 Second moments . . . . .	18
2.2.3 Third moments . . . . .	19
2.2.4 Fourth moments . . . . .	22
2.3 The limiting spectral distribution of covariance matrices for a class of random matrices with dependent entries . . . . .	26
2.3.1 Introduction and primary result . . . . .	26
2.3.2 Notation and overview . . . . .	27
2.3.3 Tools from probability theory and linear algebra . . . . .	28

2.3.4	Concentration . . . . .	31
2.3.5	Hanson–Wright inequality . . . . .	33
2.3.6	Proof of Theorem 1 . . . . .	36
2.4	The limiting spectral distribution of the genetic relatedness matrix . . . . .	39
2.4.1	From diploid genotypes to haploid genotypes . . . . .	39
<b>3</b>	<b>Impact of assortative mating on marker-based heritability estimators</b>	<b>40</b>
3.1	Haseman-Elston regression . . . . .	40
3.2	Residual maximum likelihood . . . . .	43
3.2.1	Notation and problem statement . . . . .	43
3.2.2	The limiting spectral distribution of $\zeta\zeta^T = A^T \tilde{Z}\tilde{Z}^T A$ . . . . .	44
3.2.3	Spectral functions of $\zeta\zeta^T$ . . . . .	46
3.2.4	Conditional expectation of the stationarity function . . . . .	47
3.2.5	Asymptotic variance of the stationarity function . . . . .	48
	<b>Supplementary Figures</b>	<b>51</b>
	Supplementary Figure 1 - High-dimensional behavior of the REML estimator . . . . .	51
	<b>Supplementary References</b>	<b>52</b>

# Supplementary Notes

## 1 Assortative mating and the joint distribution of causal variants; previous results

Here we briefly introduce the model of assortative mating (AM) on a quantitative trait first introduced in by Fisher in the early twentieth century<sup>1</sup> and later developed with greater rigor by a number of quantitative geneticists<sup>2,3,4,5,6</sup>. Results, all of which are presented directly in Nagylaki<sup>4</sup> or, in light of a few additional assumptions, are immediate corollaries thereof, are stated without proof.

## 1.1 Quantitative genetic model and notation

Consider a phenotype  $Y$  with zero expectation influenced by  $m$  standardized diploid causal SNP loci  $\{\mathcal{Z}_k\}_{k=1}^m$ ,

$$Y = \sum_{k=1}^m \mathcal{Z}_k \tilde{u}_k + E, \quad (1)$$

$$E \sim \mathcal{N}(0, \sigma_e^2), \quad (2)$$

where  $E$  represents the non-heritable component of  $Y$  and is assumed to be independent of the genotypes. “Standardization” here is with respect to the variance of a diploid biallelic genotype under Hardy-Weinberg equilibrium (HWE). Thus, standardized genotypes at polymorphic loci not in HWE will not have unit variance.

As in<sup>4</sup>, we assume that the vector of allele substitution effects  $\tilde{u}$  is such that  $\sigma_{g,0}^2 = \tilde{u}^T \tilde{u}$ ,  $\sum_k \tilde{u}_k = 0$ . We will make the additional assumption that the allelic substitution effects are uniformly bounded in that there exists  $0 < c < \infty$  such that, for all  $m$ ,  $\max_{k=1, \dots, m} |\tilde{u}_k| \leq c \cdot m^{-1/2}$ . In other words, for increasingly polygenic traits, effect sizes remain commensurately small. At each locus, define the *diploid genic value* as  $Z_k = \mathcal{Z}_k \tilde{u}_k$ ,  $k = 1, \dots, m$ . Each  $Z_k$  can be represented as the weighted sum of two *haploid genic values*  $Z_k = 2^{-1/2}(H_k^1 + H_k^2) = 2^{-1/2}(\mathcal{H}_k^1 + \mathcal{H}_k^2)\tilde{u}_k$ ,  $k = 1, \dots, m$ . For the sake of cleaner notation, we will sometimes singly index haploid genic values  $\{H_k^j\}_{1 \leq k \leq m}^{j \in \{1,2\}}$  as  $\{G_k\}_{1 \leq k \leq 2m}$  such that  $G_k = H_{\lfloor k/2 \rfloor}^{1+\bar{k}_2}$  (where  $\bar{k}_2$  denotes  $k \bmod 2$ ), and proceed analogously for genotypes  $\mathcal{G}_k$  such that  $\mathcal{G}_k u_k = \mathcal{H}_{\lfloor k/2 \rfloor}^{1+\bar{k}_2} \tilde{u}_{\lfloor k/2 \rfloor} := \mathcal{H}_{\lfloor k/2 \rfloor}^{1+\bar{k}_2} u_k$ .

Each causal haploid genotype  $\mathcal{H}_k$  has been standardized such that  $\mathbb{E}[\mathcal{H}_k] = 0$ ,  $\text{Var}(\mathcal{H}_k) = 1$ . The same is true of each  $\mathcal{Z}_k$  when  $\mathcal{H}_k^1$  and  $\mathcal{H}_k^2$  are independent as under random mating (i.e., under Hardy-Weinberg equilibrium). Thus, the initial genetic variance of  $Y$  is simply

$$\text{Var}\left(\sum_{k=1}^m \mathcal{Z}_k \tilde{u}_k\right) = \tilde{u}^T \tilde{u} = \sigma_{g,0}^2. \quad (3)$$

We can thus write the phenotype in terms of the genic values or genotypes in multiple ways:

$$\begin{aligned} Y &= \sum_{k=1}^m Z_k + E &&= \sum_{k=1}^m \mathcal{Z}_k \tilde{u}_k + E \\ &= \frac{1}{\sqrt{2}} \sum_{k=1}^{2m} G_k + E &&= \frac{1}{\sqrt{2}} \sum_{k=1}^{2m} \mathcal{G}_k u_k + E \\ &= \frac{1}{\sqrt{2}} \sum_{k=1}^m (H_k^1 + H_k^2) + E &&= \frac{1}{\sqrt{2}} \sum_{k=1}^m (\mathcal{H}_k^1 + \mathcal{H}_k^2) \tilde{u}_k + E. \end{aligned} \quad (4)$$

The haploid genic values follow a multivariate distribution such that, for all  $k$ ,

$$\mathbb{E}[Z_k] = \mathbb{E}[G_k] = \mathbb{E}[H_k^1] = \mathbb{E}[H_k^2] = 0, \quad (5)$$

$$\mathbb{E}[\mathcal{Z}_k] = \mathbb{E}[\mathcal{G}_k] = \mathbb{E}[\mathcal{H}_k^1] = \mathbb{E}[\mathcal{H}_k^2] = 0, \quad (6)$$

$$\text{Var}(Z_k)_{t=0} = \text{Var}(G_{2k}) = \text{Var}(H_k^1) = \text{Var}(H_k^2) = \tilde{u}_k^2, \quad (7)$$

$$\text{Var}(\mathcal{Z}_k)_{t=0} = \text{Var}(\mathcal{G}_k) = \text{Var}(\mathcal{H}_k^1) = \text{Var}(\mathcal{H}_k^2) = 1, \quad (8)$$

and

$$\mathbb{E}_k[\text{Var}(Z_k)_{t=0}] = \mathbb{E}_k[\text{Var}(G_k)] = \dots = \sigma_{g,0}^2/m, \quad (9)$$

$$\mathbb{E}_k[\text{Var}(\mathcal{Z}_k)_{t=0}] = \mathbb{E}_k[\text{Var}(\mathcal{G}_k)] = \dots = 1, \quad (10)$$

where  $\mathbb{E}_k[\cdot]$  denotes the average over loci (i.e., expectation with respect to the probability measure assigning equal probability to each causal locus). The qualification that the above holds for diploid quantities only at generation zero reflects the fact that the univariate marginal distributions of haploid quantities are time-invariant whereas diploid quantities are subject to departures from Hardy-Weinberg equilibrium.

When all of haploid genic values  $\{G_k\}_{k=1}^{2m}$  are independent, as is expected for unlinked loci under random mating, the variance of the phenotype is simply

$$\text{Var}(Y) = \mathbb{E} \left[ \frac{1}{2} \sum_{k=1}^{2m} G_k^2 \right] + \sigma_e^2 \quad (11)$$

$$= \sigma_{g,0}^2 + \sigma_e^2. \quad (12)$$

On the other hand, if the genic values  $\{G_k\}_{k=1}^{2m}$  have non-zero second moments, the total variance is

$$\begin{aligned} \text{Var}(Y) &= \frac{1}{2} \sum_{k,l=1}^{2m} \mathbb{E}[G_k G_l] + \sigma_e^2 \\ &= \frac{1}{2} \sum_{k,l=1}^{2m} u_k u_l \text{Corr}(G_k G_l) + \sigma_e^2. \end{aligned} \quad (13)$$

For the special case where the allelic substitution effects  $u_i$  and the correlation among haploid genic values are the same for all  $i, j, k, l$ ,  $i \neq j \vee k \neq l$  (i.e., for exchangeable loci),  $\text{Corr}(H_k^i, H_l^j) \equiv \mu$ , we have

$$\text{Var}(Y) = \frac{\mu}{2} \sum_{k,l=1}^{2m} u_k u_l + (1 - \mu) \sum_{k=1}^m \tilde{u}_k^2 + \sigma_e^2 \quad (14)$$

$$= 2\mu \left( \sum_{k=1}^m \tilde{u}_k \right)^2 + (1 - \mu) \sigma_{g,0}^2 + \sigma_e^2 \quad (15)$$

$$= \sigma_{g,0}^2 (1 + (2m - 1)\mu) + \sigma_e^2 \quad (16)$$

## 1.2 Causal variant dynamics across generations

Let  $Y^*$ ,  $Y^{**}$ , and  $\tilde{Y}$  and denote the respective phenotypes of parent-parent-offspring trio and similarly denote related quantities. Index generations with  $t \in \mathbb{Z}^+$  and define the following parameters:

- $\ell_{kl,t} = \text{Corr}(H_{k,t}^1, H_{l,t}^2) = \text{Corr}(H_{k,t}^2, H_{l,t}^1)$ ,  $k, l \in \{1, \dots, m\}$ , the correlation between haploid genic values on uniting gametes at generation  $t$  (within either parent),
- $\kappa_{kl,t} = \text{Corr}(H_{k,t}^1, H_{l,t}^1) = \text{Corr}(H_{k,t}^2, H_{l,t}^2)$ ,  $k, l \in \{1, \dots, m\}$ , the correlation between haploid genic values within either parent on the same gamete at generation  $t$  (note that for  $k = l$ ,  $\kappa_{kl,t} = 1$ ),
- $\mu_{kl,t} = \text{Corr}(H_{k,t}^{i*}, H_{l,t}^{j**}) = \text{Corr}(H_{k,t}^{i*}, H_{l,t}^{i**})$ ,  $k, l \in \{1, \dots, m\}$ ,  $i, j \in \{1, 2\}$ , the cross mate correlation between haploid genic values at generation  $t$ ,
- $\Upsilon_t = \{\kappa_{kl,t} + \ell_{kl,t}\}_{kl}$ , the  $m \times m$  covariance matrix for diploid genotypes  $\{\mathcal{Z}_k\}_{k=1}^m$  at generation  $t$ ,
- $u_k = \sqrt{\text{Var}(G_k)}$ , the time-invariant standard deviation of the  $k^{\text{th}}$  haploid locus; because each  $G_k$  has unit variance, we have  $\mathbb{E}_k[u_k^2] = \sigma_{g,0}^2/m$ , and  $u^T u = 2\sigma_{g,0}^2$ ; we also denote the  $k^{\text{th}}$  diploid locus standard deviation by  $\tilde{u}_k = u_{\lceil k/2 \rceil}$  with  $\tilde{u}^T \tilde{u} = \sigma_{g,0}^2$ ,
- $\sigma_{g,t}^2 = 2^{-1} \text{Var}(\sum_{k=1}^{2m} G_k) = \tilde{u}^T \Upsilon_t \tilde{u}$ , the genetic variance at generation  $t$  given genotypes,
- $\sigma_e^2 = \text{Var}(E)$ , the time-invariant residual variance,
- $\sigma_{y,t}^2 = \sigma_{g,t}^2 + \sigma_e^2$ , the phenotypic variance at generation  $t$ ,
- $h_t^2 = \sigma_{g,t}^2 / \sigma_{y,t}^2$ , the heritability at generation  $t$ ,
- $r = \text{Corr}(Y^*, Y^{**}) \in (0, 1)$ , the time-invariant phenotypic correlation between mates,
- $\text{cov}_{g,t} = 2^{-1} \text{Cov}(\sum_{k=1}^{2m} G_k^*, \sum_{k=1}^{2m} G_k^{**})_t = 2^{-1} \sum_{k,l=1}^{2m} u_k u_l \mu_{\lceil k/2 \rceil \lceil l/2 \rceil, t}$ , the genetic covariance between mates at generation  $t$ ,
- $r_{g,t} = \text{cov}_{g,t} / \sigma_{g,t}^2$ , the genetic correlation between mates at generation  $t$ . Note that we also have  $r_{g,t} = r \cdot h_t^2$ , because  $\text{corr}(\sum G_k^*, \sum G_k^{**}) = \text{corr}(\sum G_k^*, Y^*) \text{corr}(Y^*, Y^{**}) \text{corr}(Y^{**}, \sum G_k^{**})$ .

At generation zero, we assume that causal variants are unlinked, i.e., that  $\kappa_{k \neq l, 0} \equiv \ell_{kl, 0} \equiv 0$ , and at all generations we assume that recombination is equally likely to occur or not to occur between causal loci. Again following<sup>4</sup>, in an infinitely large population, a single generation of assortative mating yields the following recurrence for correlations among haploid causal variants:

$$\ell_{kl,t} = \mu_{kl,t-1}, \tag{17}$$

$$\kappa_{kl,t} = \frac{1}{2}(\kappa_{kl,t-1} + \ell_{kl,t-1}) \llbracket k \neq l \rrbracket + \llbracket k = l \rrbracket, \quad (18)$$

such that  $\sigma_{g,t}^2 = \sum_{k,l=1}^m (\kappa_{kl,t} + \ell_{kl,t}) \tilde{u}_k \tilde{u}_l = \tilde{u}^T \Upsilon_t \tilde{u}$ . The above system approaches a stable equilibrium such that

$$\kappa_{kl,\infty} = \ell_{kl,\infty} = \mu_{kl,\infty}, \quad \text{for } k \neq l, \quad (19)$$

$$\ell_{kk,\infty} = \mu_{kk,\infty}. \quad (20)$$

At this fixed point, we can write the equilibrium genetic variance in terms of the generation zero genetic variance and the correlations among haploid effects, recalling that  $cov_{g,t} = \frac{1}{2} \sum_{k,l=1}^{2m} u_k u_l \mu_{\lceil k/2 \rceil \lceil l/2 \rceil, t}$ :

$$\sigma_{g,\infty}^2 = \sum_{k,l=1}^m (\kappa_{kl,\infty} + \ell_{kl,\infty}) \tilde{u}_k \tilde{u}_l \quad (21)$$

$$= \sigma_{g,0}^2 + \frac{1}{2} \sum_{k,l=1}^{2m} u_k u_l \mu_{\lceil k/2 \rceil \lceil l/2 \rceil, \infty} - \frac{1}{2} \sum_{k=1}^{2m} u_k^2 \mu_{\lceil k/2 \rceil \lceil k/2 \rceil, \infty} \quad (22)$$

$$= \sigma_{g,0}^2 + (1 - m_e^{-1}) cov_{g,\infty} \quad (23)$$

where the effective number of loci is defined as

$$m_e := \left( \sum_{k,l}^{2m} \mu_{\lceil k/2 \rceil \lceil l/2 \rceil, \infty} u_k u_l \right) / \left( \sum_{k=1}^{2m} \mu_{\lceil k/2 \rceil \lceil k/2 \rceil, \infty} u_k^2 \right) \quad (24)$$

Defining  $Q = 1 - 1/m_e$  we have that  $\sigma_{g,\infty}^2 = \sigma_{g,0}^2 + Q \cdot cov_{g,\infty}$ , noting that that  $Q \rightarrow 1$  as  $m_e \rightarrow \infty$ . We assume that  $m_e$  is large and replace  $Q$  with the approximation  $Q \approx 1$ .

Using the relation  $\mu_{\lceil k/2 \rceil \lceil l/2 \rceil, \infty} = r \nu_{k,\infty} \nu_{l,\infty}$  derived below, this assumption reduces to the assumption

$$\sum_{k,l}^{2m} \nu_{k,\infty} \nu_{l,\infty} u_k u_l \gg \sum_{k=1}^{2m} (\nu_{k,\infty}^2 u_k^2), \quad (25)$$

or, equivalently, that

$$\sum_{k,l}^{2m} \nu_{k,\infty} \nu_{l,\infty} u_k u_l = \nu_{\infty}^T u u^T \nu_{\infty} = \left( \sum_{k=1}^{2m} \nu_{k,\infty} u_k \right)^2 \gg \sum_{k=1}^{2m} (\nu_{k,\infty}^2 u_k^2). \quad (26)$$

Substituting via  $r_{g,t} = cov_{g,t} / \sigma_{g,t}^2$ , the equilibrium genetic variance for polygenic traits is approximated as

$\sigma_{g,\infty}^2 \approx \sigma_{g,0}^2(1 - r_{g,\infty})^{-1}$ . Substituting  $r_{g,\infty} = rh_\infty^2$ , the equilibrium heritability is then computed

$$h_\infty^2 \approx \frac{\sigma_{g,0}^2(1 - rh_\infty^2)^{-1}}{\sigma_{g,0}^2(1 - rh_\infty^2)^{-1} + \sigma_e^2}. \quad (27)$$

In terms of the time-invariant phenotypic correlations between mates  $r = r_{g,\infty}/h_\infty^2$ , the equilibrium heritability can be written

$$h_\infty^2 \approx \frac{\sigma_{g,0}^2(1 - rh_\infty^2)^{-1}}{\sigma_{g,0}^2(1 - rh_\infty^2)^{-1} + \sigma_e^2} \quad (28)$$

$$\implies h_\infty^2 \approx (2r\sigma_e^2)^{-1} \left( \sigma_e^2 + \sigma_{g,0}^2 \pm \sqrt{(2 - 4r)\sigma_e^2\sigma_{g,0}^2 + \sigma_e^4 + \sigma_{g,0}^4} \right) \quad (29)$$

$$= (2r\sigma_e^2)^{-1} \left( \sigma_e^2 + \sigma_{g,0}^2 \pm \sqrt{\sigma_{y,0}^4 - 4r\sigma_e^2\sigma_{g,0}^2} \right). \quad (30)$$

The upper root in this expression can be excluded because it is necessarily greater than or equal to 1, with equality only when  $r = 1$ . This can be shown from the following observation together with  $\sqrt{(1 - h_0^2)^{-2} - 4rh_0^2(1 - h_0^2)^{-1}} \geq |(1 - h_0^2)^{-1} - 2r|$ . Observing that  $\sigma_{g,0}^2/\sigma_e^2 = h_0^2(\sigma_e^2 + \sigma_{g,0}^2)/\sigma_e^2 = h_0^2/(1 - h_0^2)$ , we can then write the equilibrium heritability in terms of the phenotypic correlation between mates and the generation zero heritability:

$$h_\infty^2 \approx (2r)^{-1} \left( (1 - h_0^2)^{-1} - \sqrt{(1 - h_0^2)^{-2} - 4rh_0^2(1 - h_0^2)^{-1}} \right). \quad (31)$$

Similarly, the equilibrium genetic correlation between mates is

$$r_{g,\infty} \approx \frac{r\sigma_{g,0}^2(1 - r_{g,\infty})^{-1}}{\sigma_{g,0}^2(1 - r_{g,\infty})^{-1} + \sigma_e^2} \quad (32)$$

$$\implies r_{g,\infty} = (2\sigma_e^2)^{-1} \left( \sigma_e^2 + \sigma_{g,0}^2 - \sqrt{(2 - 4r)\sigma_e^2\sigma_{g,0}^2 + \sigma_e^4 + \sigma_{g,0}^4} \right) \quad (33)$$

$$= \frac{1}{2} \left( (1 - h_0^2)^{-1} - \sqrt{(1 - h_0^2)^{-2} - 4rh_0^2(1 - h_0^2)^{-1}} \right). \quad (34)$$

Finally, the equilibrium variance is computed

$$\sigma_{y,\infty}^2 \approx \sigma_e^2 + \sigma_{g,0}^2(1 - r_{g,\infty})^{-1} \quad (35)$$

$$= \left( \frac{1 - r_{g,\infty}(1 - h_0^2)}{1 - r_{g,\infty}} \right) \sigma_{y,0}^2. \quad (36)$$

Further assumptions are required to make any explicit claims about the equilibrium values of the correlations between haploid causal variants,  $\mu_{kl,\infty}$ ,  $k, l = 1, \dots, m$ . As in<sup>4</sup>, we assume the regressions of individuals'

genetic scores on their phenotypes, and on their mates' values, are linear:

$$\mathbb{E}\left[2^{-1/2}\sum_{k=1}^{2m}G_k^* \middle| Y^*\right] = h_t^2 Y^*, \quad \mathbb{E}\left[2^{-1/2}\sum_{k=1}^{2m}G_k^* \middle| Y^{**}\right] = r h_t^2 Y^{**}, \quad (37)$$

and that the same is true of their individual haploid genic values:

$$\mathbb{E}[G_k^* | Y^*] = \nu_{k,t} u_k \sigma_{y,t}^{-1} Y^*, \quad (38)$$

where  $\nu_k$  denotes the correlation between the haploid genic value and the phenotype. At equilibrium, we then have

$$\mathbb{E}[G_k^* | Y^{**}] = r \nu_{k,\infty} u_k \sigma_{y,\infty}^{-1} Y^{**}. \quad (39)$$

Thus, by conditional independence of mates' genotypes, we have

$$\mu_{kl,\infty} u_k u_l = \mathbb{E}[G_k^* G_l^{**}] \quad (40)$$

$$= \mathbb{E}[\mathbb{E}[G_k^* | Y^{**}] \mathbb{E}[G_l^{**} | Y^{**}]] \quad (41)$$

$$= r u_k u_l \nu_{k,\infty} \nu_{l,\infty} \quad (42)$$

$$\implies \mu_{kl,\infty} = r \nu_{k,\infty} \nu_{l,\infty}. \quad (43)$$

We abuse notation here such that  $\mu_{kl,\infty}$  refers to  $\mu_{\lceil k/2 \rceil \lceil l/2 \rceil, \infty}$ . Compute  $\nu_{k,\infty}$  as follows:

$$\nu_{k,\infty} \sigma_{y,\infty} u_k = \mathbb{E}[G_k Y] \quad (44)$$

$$= \sum_{l=1}^{2m} 2^{-1/2} u_k u_l \mathbb{E}[G_k G_l] \quad (45)$$

$$\implies \nu_{k,\infty} = 2^{-1/2} u_k \sigma_{y,\infty}^{-1} (1 - r \nu_{k,\infty}^2) + 2^{-1/2} r \nu_{k,\infty} \sigma_{y,\infty}^{-1} \sum_{l=1}^{2m} \nu_{l,\infty} u_l. \quad (46)$$

Now, writing

$$h_\infty^2 Y = \mathbb{E}\left[2^{-1/2}\sum_{l=1}^{2m}G_l \middle| Y\right] = 2^{-1/2} \sigma_{y,\infty}^{-1} \sum_{l=1}^{2m} \nu_{l,\infty} u_l Y, \quad (47)$$

yields  $h_\infty^2 = \sigma_{y,\infty}^{-1} 2^{-1/2} \sum_{l=1}^{2m} \nu_{l,\infty} u_l$ , which we then substitute into the previous expression to yield

$$\nu_{k,\infty} = \sigma_{y,\infty} \left(2^{1/2} u_k r\right)^{-1} \left(\sqrt{\frac{2u_k^2 r}{\sigma_{y,\infty}^2} + (1 - r_{g,\infty})^2} - (1 - r_{g,\infty})\right). \quad (48)$$



Note that, under exchangeable loci, we then have  $\nu_{k,\infty} \equiv \pm\nu_\infty$  for all  $k$ , which in turn implies that  $\mu_{kl,\infty} = \pm r\nu_\infty^2 \equiv \pm\mu_\infty$  for all  $k, l$ , with  $\text{sgn } \mu_{kl,\infty} = \text{sgn } u_k u_l$ . That is, the equilibrium correlations between distinct haploid loci are all of equal magnitude, which can be bounded as:

$$\text{cov}_{g,\infty} = \frac{1}{2} \sum_{k,l=1}^{2m} u_k u_l \mu_\infty \quad (49)$$

$$= 4m^2 (\sigma_{g,0}^2 / 2m) \mu_\infty \quad (50)$$

$$\implies \mu_\infty = \pm \frac{r_{g,\infty} \sigma_{g,\infty}^2}{2m \sigma_{g,0}^2} = \mathcal{O}(m^{-1}). \quad (51)$$

In the more general case, for a highly polygenic trait with fixed generation zero conditions, we can employ our assumption that  $\max_{k=1,\dots,m} |\tilde{u}_k| \leq c \cdot m^{-1/2}$  to obtain a similar bound. The equilibrium correlations between genic values are  $\mu_{kl,\infty} = r\nu_{k,\infty}\nu_{l,\infty}$ . Defining  $\beta_k = u_k \sigma_{y,\infty}^{-1} = \mathcal{O}(m^{-1/2})$ , we compute the equilibrium correlation between mates' haploid genic effects as

$$\mu_{kl,\infty} = r\nu_{k,\infty}\nu_{l,\infty} \quad (52)$$

$$= \sigma_{y,\infty}^2 (2r^2 u_k u_l)^{-1} \left( \sqrt{(1-r_{g,\infty})^2 + 2u_k^2 r \sigma_{y,\infty}^{-2}} - (1-r_{g,\infty}) \right) \left( \sqrt{(1-r_{g,\infty})^2 + 2u_l^2 r \sigma_{y,\infty}^{-2}} - (1-r_{g,\infty}) \right) \quad (53)$$

$$= (2r^2 \beta_k \beta_l)^{-1} \left( \sqrt{(1-r_{g,\infty})^2 + 2\beta_k^2 r} - (1-r_{g,\infty}) \right) \left( \sqrt{(1-r_{g,\infty})^2 + 2\beta_l^2 r} - (1-r_{g,\infty}) \right) \quad (54)$$

$$= \frac{1}{2} (1-r_{g,\infty})^{-2} |\beta_k \beta_l| + \mathcal{O}(\beta_k^3, \beta_l^3) \quad (55)$$

$$\implies |\mu_{kl,\infty}| = \mathcal{O}(m^{-1}). \quad (56)$$

The last statement follows from the fact that  $(1-r_{g,\infty})^{-2}$  is bounded independent of  $m$  via  $r_{g,\infty} = rh_\infty^2$  with  $r \in (0, 1)$ ,  $h_\infty^2 \in [0, 1]$ . In general, we can write the equilibrium covariance between diploid genotypes as

$$\Upsilon_\infty = \{Cov(\mathbf{Z}_k, \mathbf{Z}_l)\}_{k,l=1}^m \quad (57)$$

$$= \begin{pmatrix} 1 + r\nu_{1,\infty}^2 & & & & & & \text{Symm} \\ 2r\nu_{1,\infty}\nu_{2,\infty} & \ddots & & & & & \\ \vdots & \ddots & \ddots & & & & \\ 2r\nu_{1,\infty}\nu_{m,\infty} & \cdots & 2r\nu_{m-1,\infty}\nu_{m,\infty} & 1 + r\nu_{m,\infty}^2 & & & \end{pmatrix} \quad (58)$$

$$:= D + 2\phi\phi^T, \quad (59)$$

where  $\phi = (\sqrt{r}\nu_{k,\infty})_{k=1}^m$  and  $D = \text{diag}(1 - \phi_k^2)_{k=1}^m$ .

## 2 Assortative mating and the joint distribution of causal variants; novel results

The results presented in the previous section characterize the first and second moments of causal variants under AM but not moments of higher order. In this section, we employ tools from thermodynamics to directly characterize the equilibrium joint distribution of causal variants and thereby obtain bounds for all higher order moments in terms of the number of causal variants  $m$ . Then, based on said bounds, we extend previous results in random matrix theory in order to characterize the limiting spectral distribution of a class of covariance matrices generated from random matrices with dependent entries. Finally, we describe the limiting spectral distribution of the genetic relatedness matrix (GRM) under AM.

### 2.1 A thermodynamic approach

At equilibrium, the variance and covariance of mates' phenotypes are constant across generations. These equilibrium values define the stationary dynamics of assortative mating, and they can be naturally expressed in terms of energy functions suitable for analysis using tools from thermodynamics. This analysis yields a complete characterization of the genetic distribution.

#### 2.1.1 Energy functions and maximum entropy

Begin by defining energy functions

$$E_1(Y^*, Y^{**}) = -Y^{*2} \tag{60}$$

$$E_2(Y^*, Y^{**}) = -Y^{**2} \tag{61}$$

$$E_3(Y^*, Y^{**}) = -Y^*Y^{**}. \tag{62}$$

The intuition from thermodynamics is that states with lower energies are more likely. Thus  $E_3$  expresses that mates' phenotype are positively correlated. The signs for  $E_1$  and  $E_2$  are counterintuitive, but it will be seen below that a countervailing combinatoric effect in going from genotypes to phenotypes leads smaller values of  $Y^{*2}$  and  $Y^{**2}$  to be more likely, as expected. For now, the signs of these functions can be thought of as purely definitional.

In equilibrium, we have fixed values for the mean energies:

$$\begin{aligned} \langle E_1 \rangle &= -\sigma_{y,\infty}^2 \\ \langle E_2 \rangle &= -\sigma_{y,\infty}^2 \\ \langle E_3 \rangle &= -r\sigma_{y,\infty}^2. \end{aligned} \tag{63}$$

Because the dynamics are determined fully at the phenotype level (macroscale), the equilibrium distribution at the genotype level (microscale) will maximize entropy with respect to the macroscale constraints (i.e., Supplementary Equation (63)). Standard thermodynamic analysis thus implies the genotypes obey a Boltzmann distribution. This enables us to characterize exactly the joint genotypic distribution for mates, and thereby to calculate arbitrary moments and other statistics.

### 2.1.2 The joint distribution of genotypes under simplifying assumptions

To simplify the initial analysis, take  $\sigma_e^2 = 0$ , and assume that all SNPs have a minor allele frequency (MAF) of 1/2 and effects  $u_i = \sigma_{g,0}/\sqrt{2m}$ . Thus  $\mathcal{G}_i = \pm 1$  and  $G_i = \pm\sigma_{g,0}/\sqrt{2m}$  for all  $i$ . Denote individuals genetic scores by  $\mathcal{G} = \sum_{i=1}^{2m} G_i$ , noting that, in this simplified case,  $Y = \mathcal{G}$ .

The energy functions can be re-expressed as functions of microstates:

$$\begin{aligned} E_1(\mathcal{G}^*, \mathcal{G}^{**}) &= - \sum_{ij=1}^{2m} G_i^* G_j^* \\ E_2(\mathcal{G}^*, \mathcal{G}^{**}) &= - \sum_{ij=1}^{2m} G_i^{**} G_j^{**} \\ E_3(\mathcal{G}^*, \mathcal{G}^{**}) &= - \sum_{ij=1}^{2m} G_i^* G_j^{**} \end{aligned} \quad (64)$$

Standard thermodynamic analysis then yields the following Boltzmann distribution for mates' genotypes:

$$\mathbb{P}[\mathcal{G}^*, \mathcal{G}^{**}] \propto \exp[-\alpha E_1(\mathcal{G}^*, \mathcal{G}^{**}) - \beta E_2(\mathcal{G}^*, \mathcal{G}^{**}) - \gamma E_3(\mathcal{G}^*, \mathcal{G}^{**})] \quad (65)$$

$$= \exp \left[ \alpha \sum_{ij=1}^{2m} G_i^* G_j^* + \beta \sum_{ij=1}^{2m} G_i^{**} G_j^{**} + \gamma \sum_{ij=1}^{2m} G_i^* G_j^{**} \right]. \quad (66)$$

The parameters  $\alpha$ ,  $\beta$ , and  $\gamma$  are inverse temperatures, with values to be determined.

Mathematically, the inverse temperature parameters correspond to Lagrange multipliers for the hard constraints in Supplementary Equation (63). Typically, these parameters are solved for by substituting the Boltzmann distribution back into the constraint equations (here three constraints for three unknowns). In the present case, there are additional constraints from the assumption of genetic stationarity. Equilibrium at the genetic level implies

$$\mathbb{E}[G_i^* G_j^*] = \mathbb{E}[G_i^* G_j^{**}] = \mathbb{E}[G_i^{**} G_j^{**}] \quad (67)$$

for all  $i \neq j$  (Section 1). This symmetry implies  $\alpha = \beta$  and  $\gamma = 2\beta$ . In more detail, the exponent in Supplementary Equation (66) can be written as  $G^T A G$ , where  $G$  is the concatenation of  $G^*$  and  $G^{**}$  and the entries of  $A$  are all  $\alpha$ ,  $\beta$ , or  $\gamma/2$ . Supplementary Equation (67) implies symmetry under permutation

of the components of  $G$ , which implies all off-diagonal entries of  $A$  must be equal. This strong symmetry argument doesn't hold once we allow heterogeneous allelic effects, but symmetry between the components corresponding to  $G_i^*$  and  $G_i^{**}$ , and likewise between  $G_j^*$  and  $G_j^{**}$ , are sufficient to conclude  $\alpha = \beta$  and  $\gamma = 2\beta$ . Based on this symmetry argument, the genetic distribution reduces to

$$\mathbb{P}[\mathcal{G}^*, \mathcal{G}^{**}] \propto \exp \left[ \beta \left( \sum_{i=1}^{2m} G_i^* + \sum_{i=1}^{2m} G_i^{**} \right)^2 \right]. \quad (68)$$

All that remains to specify this distribution is to solve for  $\beta$ . This is done by requiring the distribution at the phenotype level to satisfy the constraints of Supplementary Equation (63).

### 2.1.3 The joint distribution of phenotypes under simplifying assumptions

We translate the microstate distribution in Supplementary Equation (68) to a macrostate distribution by coarse-graining. For any value of  $Y$ , let  $n_Y$  be the corresponding number of positive alleles,  $|\{i : \mathcal{G}_i = 1\}|$ . We have

$$Y = (n_Y - (2m - n_Y)) \frac{\sigma_{g,0}}{\sqrt{2m}} \quad (69)$$

and therefore

$$n_Y = m + \sqrt{\frac{m}{2}} \frac{Y}{\sigma_{g,0}}. \quad (70)$$

The number of genotypes consistent with  $Y$  is thus

$$\frac{\Gamma(2m+1)}{\Gamma(n_Y+1)\Gamma(2m-n_Y+1)} \approx \frac{2^{2m}}{\sqrt{\pi m}} \exp \left[ -\frac{(n_Y - m)^2}{m} \right] \quad (71)$$

$$= \frac{2^{2m}}{\sqrt{\pi m}} \exp \left[ -\frac{Y^2}{2\sigma_{g,0}^2} \right]. \quad (72)$$

The macrostate probabilities can then be obtained by multiplying the microstate probabilities by these counts:

$$\mathbb{P}[Y^*, Y^{**}] = \sum_{\mathcal{G}^* : \mathcal{G}^* = Y^*} \sum_{\mathcal{G}^{**} : \mathcal{G}^{**} = Y^{**}} \mathbb{P}[\mathcal{G}^*, \mathcal{G}^{**}] \quad (73)$$

$$\propto \frac{2^{2m}}{\sqrt{\pi m}} \exp \left[ -\frac{Y^{*2}}{2\sigma_{g,0}^2} \right] \cdot \frac{2^{2m}}{\sqrt{\pi m}} \exp \left[ -\frac{Y^{**2}}{2\sigma_{g,0}^2} \right] \cdot \exp \left[ \beta (Y^* + Y^{**})^2 \right] \quad (74)$$

$$\propto \exp \left[ \left( \beta - \frac{1}{2\sigma_{g,0}^2} \right) Y^{*2} + \left( \beta - \frac{1}{2\sigma_{g,0}^2} \right) Y^{**2} + 2\beta Y^* Y^{**} \right]. \quad (75)$$

Writing

$$\beta = \frac{r}{2\sigma_{g,0}^2 (1+r)} \quad (76)$$

and

$$\sigma_{Y,\infty}^2 = \frac{\sigma_{g,0}^2}{1-r}, \quad (77)$$

we can re-express the phenotype distribution as

$$\mathbb{P}[Y^*, Y^{**}] \propto \exp \left[ -\frac{1}{2\sigma_{Y,\infty}^2(1-r^2)} (Y^{*2} + Y^{**2} - 2rY^*Y^{**}) \right], \quad (78)$$

or equivalently

$$(Y^*, Y^{**}) \sim \mathcal{N} \left( 0, \sigma_{Y,\infty}^2 \begin{bmatrix} 1 & r \\ r & 1 \end{bmatrix} \right). \quad (79)$$

Supplementary Equation (79) satisfies the constraints of Supplementary Equation (63). Notice that this analysis also yields an expression for the equilibrium phenotypic variance in Supplementary Equation (77), which was taken as fixed but unknown. Altogether there were four unknowns  $(\alpha, \beta, \gamma, \sigma_{Y,\infty}^2)$ , and four constraints from Supplementary Equations 63 and 67 (where  $\mathbb{E}[\mathcal{G}_i^* \mathcal{G}_j^*] = \mathbb{E}[\mathcal{G}_i^{**} \mathcal{G}_j^{**}]$  are  $\langle E_1 \rangle = \langle E_2 \rangle$  are redundant).

#### 2.1.4 Stationarity

Also useful is to verify that the genetic distribution of Supplementary Equation (68) is stationary. Above we used the equilibrium assumption for any two sites, but not for the joint distribution over the full genome. For ease of notation, number loci so that  $\{i, m+i\}$  form a diploid pair, and assume without loss of generality that inherited alleles from each parent come from sites 1 through  $m$ :

$$\tilde{G}_i = \begin{cases} G_i^* & i \leq m \\ G_{i-m}^{**} & i > m. \end{cases} \quad (80)$$

Under these definitions, Supplementary Equation (68) can be rewritten as

$$\mathbb{P}[\mathcal{G}^*, \mathcal{G}^{**}] \propto \exp \left[ \beta \left( \sum_{i=1}^{2m} \tilde{G}_i + \sum_{i=m+1}^{2m} G_i^* + \sum_{i=m+1}^{2m} G_i^{**} \right)^2 \right], \quad (81)$$

and therefore the marginal distribution for the offspring,  $\tilde{\mathcal{G}}$ , is given by

$$\mathbb{P}[\tilde{\mathcal{G}}] \propto \sum_{\{\mathcal{G}_i^*\}_{i=m+1}^{2m}} \sum_{\{\mathcal{G}_i^{**}\}_{i=m+1}^{2m}} \exp \left[ \beta \left( \sum_{i=1}^{2m} \tilde{G}_i + \sum_{i=m+1}^{2m} G_i^* + \sum_{i=m+1}^{2m} G_i^{**} \right)^2 \right]. \quad (82)$$

On the other hand, the marginal distribution for one parent is given by

$$\mathbb{P}[\mathcal{G}^*] \propto \sum_{\{\mathcal{G}_i^{**}\}_{i=1}^{2m}} \exp \left[ \beta \left( \sum_{i=1}^{2m} G_i^* + \sum_{i=1}^m G_i^{**} + \sum_{i=m+1}^{2m} G_i^{**} \right)^2 \right], \quad (83)$$

which is equivalent to the previous expression under the change of variables

$$\left( \{\mathcal{G}_i^*\}_{i=1}^{2m}, \{\mathcal{G}_i^{**}\}_{i=1}^m, \{\mathcal{G}_i^{**}\}_{i=m+1}^{2m} \right) \rightarrow \left( \{\tilde{\mathcal{G}}_i\}_{i=1}^{2m}, \{\mathcal{G}_i^*\}_{i=m+1}^{2m}, \{\mathcal{G}_i^{**}\}_{i=m+1}^{2m} \right). \quad (84)$$

In conclusion, the joint distribution of mates' genotypes given by Supplementary Equation (68) (a) reproduces the dynamics of assortative mating as captured in Supplementary Equation (63) (in particular  $\text{cor}(Y^*, Y^{**}) = r$ ), (b) is fully stationary, and (c) maximizes genotype-level entropy conditioned on the constraints of the phenotype-level dynamics. Therefore Supplementary Equation (68) represents the thermodynamic equilibrium to which the population necessarily converges.

### 2.1.5 Extension to the general case

The above analysis made three simplifying assumptions: no environmental variance, all allele frequencies are all 1/2, and homogeneous allelic effects (i.e., exchangeable loci). These must be removed to obtain a complete model.

To include environmental variance, the above analysis can be applied to the genetic scores,  $\mathcal{G} = \sum_i G_i$ , in place of the phenotypes,  $Y$ . The expression for equilibrium phenotypic variance in Supplementary Equation (77) becomes an expression for equilibrium genetic variance, with the phenotypic correlation ( $r$ ) replaced by the genetic correlation ( $r_{g,\infty}$ ):

$$\sigma_{g,\infty}^2 = \frac{\sigma_{g,0}^2}{1 - r_{g,\infty}}. \quad (85)$$

Using the relationship

$$r_{g,\infty} = r \frac{\sigma_{g,\infty}^2}{\sigma_{g,\infty}^2 + \sigma_e^2}, \quad (86)$$

we can derive the same expression for  $r_{g,\infty}$  (and hence  $\sigma_{g,\infty}^2$  and  $h_\infty^2$ ) as found by previous methods:

$$r_{g,\infty} = \frac{1}{2} \left[ (1 - h_0^2)^{-1} - \sqrt{(1 - h_0^2)^{-2} - 4rh_0^2(1 - h_0^2)^{-1}} \right]. \quad (87)$$

Substituting the genetic correlation for the phenotypic correlation in the expression for inverse temperature

in Supplementary Equation (76), we obtain

$$\beta = \frac{1 + 2rh_0^2 - \sqrt{1 - 4rh_0^2(1 - h_0^2)}}{4\sigma_{g,0}^2(2 + rh_0^2 - h_0^2)}. \quad (88)$$

The analysis above did not explicitly constrain allele frequencies, but the symmetry of Supplementary Equation (68) under sign reversal ( $\mathcal{G} \mapsto -\mathcal{G}$ ) implies all loci have minor allele frequencies (MAF) of 1/2. Arbitrary frequencies values can be imposed by treating them as constraints in the entropy optimization problem, supplementing Supplementary Equation (63). Writing  $p_i$  as the MAF for site  $i$  and continuing to code  $\mathcal{G}_i$  as  $\pm 1$  (this is more convenient and conventional in the present framework than standardizing  $\mathcal{G}_i$ ), we have

$$\langle \mathcal{G}_i \rangle = 1 - 2p_i. \quad (89)$$

These constraints manifest as additional terms in the Boltzmann distribution:

$$\mathbb{P}[\mathcal{G}^*, \mathcal{G}^{**}] \propto \exp \left[ \beta \left( \sum_{i=1}^{2m} G_i^* + \sum_{i=1}^{2m} G_i^{**} \right)^2 + \sum_i \lambda_i (G_i^* + G_i^{**}) \right]. \quad (90)$$

The  $\lambda_i$  coefficients are determined, together with  $\beta$ , by solving for the values that lead the distribution to satisfy Supplementary Equations (63) and (89). The system described by Supplementary Equation (90) is an instance of a spin glass, or generalized Ising model.

Finally, allowing heterogeneous allelic effects,  $u_i$ , further complicates determination of  $\beta$  and  $\{\lambda_i\}$ . The counting strategy of Supplementary Equations (70) and (72) no longer applies, because the genotypes consistent with a given genetic score no longer arise from a binomial distribution (because of heterogeneous  $u_i$ ), and because they no longer all have equal energy (because of heterogeneous MAF or  $\lambda_i$ ). Nevertheless, the expression for the genetic distribution in Supplementary Equation (90) is unchanged (only  $\beta$  and  $\lambda_i$  are changed).

In summary, the complete characterization of the genetic distribution in the general case is given by Supplementary Equation (90), with  $\beta$  and  $\{\lambda_i\}$  determined by the constraints in Supplementary Equations (63) and (89). The former constraint amounts to calculating the joint distribution of genetic scores,

$$\mathbb{P}[\mathcal{G}^*, \mathcal{G}^{**}] \propto \sum_{\mathcal{G}^*: \sum G_i^* = \mathcal{G}^*} \sum_{\mathcal{G}^{**}: \sum G_i^{**} = \mathcal{G}^{**}} \mathbb{P}[\mathcal{G}^*, \mathcal{G}^{**}], \quad (91)$$

using this distribution to calculate the genetic variance ( $\sigma_{g,\infty}^2$ ) and genetic correlation between mates ( $r_{g,\infty}$ ), and finally calculating the phenotypic correlation between mates:

$$r = \frac{r_{g,\infty} (\sigma_{g,\infty}^2 + \sigma_e^2)}{\sigma_{g,\infty}^2}. \quad (92)$$

## 2.2 Moment derivations

Assume a global distribution given by

$$\Pr[\mathcal{G}^*, \mathcal{G}^{**}] \propto \exp[\beta \mathcal{G}^2 + W] \quad (93)$$

where  $\mathcal{G} = \sum_j \mathcal{G}_j u_j$  and  $W = \sum_j \mathcal{G}_j \lambda_j$ . Notice that we're concatenating  $(\mathcal{G}^*, \mathcal{G}^{**})$  into a single vector  $\mathcal{G}$ , and similarly for  $\mathcal{G}$  and  $W$ , for economy of notation.

We have shown above that  $\beta$  is given by

$$\beta = \frac{1 + 2rh_0^2 - \sqrt{1 - 4rh_0^2(1 - h_0^2)}}{4\sigma_{g,0}^2(2 + rh_0^2 - h_0^2)}. \quad (94)$$

The  $\lambda$  parameters are unknown and will be approximated below, based on the constraint

$$\mathbb{E}[\mathcal{G}_i] = q_i. \quad (95)$$

We will make repeated use of the following relation, for various  $X$ :

$$\frac{\partial}{\partial u_i} \mathbb{E}[X] = \frac{\partial}{\partial u_i} \frac{\sum_{\mathcal{G}} X \exp[\beta \mathcal{G}^2 + W]}{\sum_{\mathcal{G}} \exp[\beta \mathcal{G}^2 + W]} \quad (96)$$

$$= \mathbb{E} \left[ \frac{\partial X}{\partial u_i} \right] + \frac{\sum_{\mathcal{G}} X \frac{\partial}{\partial u_i} \exp[\beta \mathcal{G}^2 + W]}{\sum_{\mathcal{G}} \exp[\beta \mathcal{G}^2 + W]} - \mathbb{E}[X] \frac{\sum_{\mathcal{G}} \frac{\partial}{\partial u_i} \exp[\beta \mathcal{G}^2 + W]}{\sum_{\mathcal{G}} \exp[\beta \mathcal{G}^2 + W]} \quad (97)$$

$$= \mathbb{E} \left[ \frac{\partial X}{\partial u_i} \right] + 2\beta \mathbb{E}[\mathcal{G}_i \mathcal{G} X] - 2\beta \mathbb{E}[\mathcal{G}_i \mathcal{G}] \mathbb{E}[X]. \quad (98)$$

### 2.2.1 Deriving $\lambda$ from the first moment

Here we expand  $\mathbb{E}[\mathcal{G}_i]$  in  $u_i$ , holding fixed all other parameters  $(u_{-i}, \lambda, \beta)$ . We then solve for the value of  $\lambda_i$  that yields  $q_i$ .

Zeroth term:

$$\mathbb{E}[\mathcal{G}_i]_{u_i=0} = \frac{\sum_{\mathcal{G}_i} \mathcal{G}_i \exp[\mathcal{G}_i \lambda_i] \sum_{\mathcal{G}_{-i}} \exp[\beta \mathcal{G}_{-i}^2 + W_{-i}]}{\sum_{\mathcal{G}_i} \exp[\mathcal{G}_i \lambda_i] \sum_{\mathcal{G}_{-i}} \exp[\beta \mathcal{G}_{-i}^2 + W_{-i}]} \quad (99)$$

$$= \tanh \lambda_i. \quad (100)$$

First derivative:

$$\frac{\partial}{\partial u_i} \mathbb{E}[\mathcal{G}_i] = 2\beta \mathbb{E}[\mathcal{G}] - 2\beta \mathbb{E}[\mathcal{G}_i \mathcal{G}] \mathbb{E}[\mathcal{G}_i]. \quad (101)$$



At  $u_i = 0$ ,  $\mathcal{G}_i$  is independent from  $\mathcal{G}_{-i}$ . Thus the first derivative becomes

$$\frac{\partial}{\partial u_i} \Big|_{u_i=0} \mathbb{E} [\mathcal{G}_i] = 2\beta (1 - \tanh^2 \lambda_i) \mathbb{E} [\mathcal{G}]_{u_i=0}. \quad (102)$$

Second derivative:

$$\frac{\partial^2}{\partial u_i^2} \mathbb{E} [\mathcal{G}_i] = 2\beta \frac{d}{du_i} \mathbb{E} [\mathcal{G}] - 2\beta \mathbb{E} [\mathcal{G}_i \mathcal{G}] \frac{d}{du_i} \mathbb{E} [\mathcal{G}_i] - 2\beta \mathbb{E} [\mathcal{G}_i] \frac{d}{du_i} \mathbb{E} [\mathcal{G}_i \mathcal{G}] \quad (103)$$

$$= 4\beta^2 \mathbb{E} [\mathcal{G}_i \mathcal{G}^2] - 4\beta^2 \mathbb{E} [\mathcal{G}_i] \mathbb{E} [\mathcal{G}^2] + 8\beta^2 \mathbb{E} [\mathcal{G}_i] \mathbb{E} [\mathcal{G}_i \mathcal{G}]^2 - 8\beta^2 \mathbb{E} [\mathcal{G}_i \mathcal{G}] \mathbb{E} [\mathcal{G}]. \quad (104)$$

At  $u_i = 0$  this becomes

$$\frac{\partial^2}{\partial u_i^2} \Big|_{u_i=0} \mathbb{E} [\mathcal{G}_i] = -8\beta^2 \tanh \lambda_i (1 - \tanh^2 \lambda_i) \mathbb{E} [\mathcal{G}]_{u_i=0}^2. \quad (105)$$

Therefore we have the expansion

$$\mathbb{E} [\mathcal{G}_i] = \tanh \lambda_i + 2\beta (1 - \tanh^2 \lambda_i) \mathbb{E} [\mathcal{G}]_{u_i=0} u_i - 4\beta^2 \tanh \lambda_i (1 - \tanh^2 \lambda_i) \mathbb{E} [\mathcal{G}]_{u_i=0}^2 u_i^2 + \mathcal{O}(u_i^3). \quad (106)$$

Next we expand  $\mathbb{E} [\mathcal{G}]$  in terms of  $u_i$ :

$$\frac{\partial}{\partial u_i} \mathbb{E} [\mathcal{G}] = \mathbb{E} [\mathcal{G}_i] + 2\beta \mathbb{E} [\mathcal{G}_i \mathcal{G}^2] - 2\beta \mathbb{E} [\mathcal{G}_i \mathcal{G}] \mathbb{E} [\mathcal{G}]. \quad (107)$$

At  $u_i = 0$  this becomes

$$\frac{\partial}{\partial u_i} \Big|_{u_i=0} \mathbb{E} [\mathcal{G}] = \tanh \lambda_i + 2\beta \tanh \lambda_i \left( \mathbb{E} [\mathcal{G}^2]_{u_i=0} - \mathbb{E} [\mathcal{G}]_{u_i=0}^2 \right). \quad (108)$$

Therefore

$$\mathbb{E} [\mathcal{G}]_{u_i=0} = \mathbb{E} [\mathcal{G}] - \tanh \lambda_i u_i - 2\beta \tanh \lambda_i \left( \mathbb{E} [\mathcal{G}^2] - \mathbb{E} [\mathcal{G}]^2 \right) u_i + \mathcal{O}(u_i^2) \quad (109)$$

where we have dropped the  $u_i = 0$  conditions on the RHS and absorbed the resulting error into the  $\mathcal{O}(u_i^2)$  term.

Substituting Supplementary Equation (109) into Supplementary Equation (106) yields

$$\begin{aligned} \mathbb{E} [\mathcal{G}_i] &= \tanh \lambda_i + 2\beta (1 - \tanh^2 \lambda_i) \mathbb{E} [\mathcal{G}] u_i - 2\beta \tanh \lambda_i (1 - \tanh^2 \lambda_i) u_i^2 \\ &\quad - 4\beta^2 \tanh \lambda_i (1 - \tanh^2 \lambda_i) \mathbb{E} [\mathcal{G}^2] u_i^2 + \mathcal{O}(u_i^3). \end{aligned} \quad (110)$$

Writing  $\tanh \lambda_i = A + Bu_i + Cu_i^2 + \mathcal{O}(u_i^3)$  and substituting  $q_i$  for  $\mathbb{E}[\mathcal{G}_i]$  yields

$$\begin{aligned} q_i &= \mathcal{O}(u_i^3) + A + (B + 2\beta(1 - A^2) \mathbb{E}[\mathcal{G}]) u_i \\ &\quad + (C - 4\beta AB \mathbb{E}[\mathcal{G}] - 2\beta A(1 - A^2) - 4\beta^2 A(1 - A^2) \mathbb{E}[\mathcal{G}^2]) u_i^2 \end{aligned} \quad (111)$$

implying

$$A = q_i \quad (112)$$

$$B = -2\beta(1 - q_i^2) \mathbb{E}[\mathcal{G}] \quad (113)$$

$$C = 2\beta q_i(1 - q_i^2) + 4\beta^2 q_i(1 - q_i^2) (\mathbb{E}[\mathcal{G}^2] - 2\mathbb{E}[\mathcal{G}]^2) \quad (114)$$

and therefore

$$\tanh \lambda_i = q_i - 2\beta(1 - q_i^2) \mathbb{E}[\mathcal{G}] u_i + 2\beta q_i(1 - q_i^2) u_i^2 + 4\beta^2 q_i(1 - q_i^2) (\mathbb{E}[\mathcal{G}^2] - 2\mathbb{E}[\mathcal{G}]^2) u_i^2 + \mathcal{O}(u_i^3). \quad (115)$$

### 2.2.2 Second moments

Here we compute the second order expansion of the second moment in  $u_i$  and  $u_j$  about zero:

$$\mathbb{E}[\mathcal{G}_i \mathcal{G}_j] \Big|_{u_i, u_j=0} = \frac{\sum_{\mathcal{G}_i} \mathcal{G}_i \exp[\mathcal{G}_i \lambda_i] \cdot \sum_{\mathcal{G}_j} \mathcal{G}_j \exp[\mathcal{G}_j \lambda_j] \cdot \sum_{\mathcal{G}_{-ij}} \exp[\beta \mathcal{G}_{-ij}^2 + W_{-ij}]}{\sum_{\mathcal{G}_i} \exp[\mathcal{G}_i \lambda_i] \cdot \sum_{\mathcal{G}_j} \exp[\mathcal{G}_j \lambda_j] \cdot \sum_{\mathcal{G}_{-ij}} \exp[\beta \mathcal{G}_{-ij}^2 + W_{-ij}]} \quad (116)$$

$$= \tanh \lambda_i \tanh \lambda_j, \quad (117)$$

$$\begin{aligned} \frac{\partial}{\partial u_i} \mathbb{E}[\mathcal{G}_i \mathcal{G}_j] \Big|_{u_i, u_j=0} &= \beta^2 \mathbb{E}[\mathcal{G}_i \mathcal{G}_j \mathcal{G}^2] - 4\beta^2 \mathbb{E}[\mathcal{G}_i \mathcal{G}_j] \mathbb{E}[\mathcal{G}^2] + 8\beta^2 \mathbb{E}[\mathcal{G}_i \mathcal{G}_j] \mathbb{E}[\mathcal{G}_i \mathcal{G}^2] - 8\beta^2 \mathbb{E}[\mathcal{G}_i \mathcal{G}] \mathbb{E}[\mathcal{G}_j \mathcal{G}] \Big|_{u_i, u_j=0} \end{aligned} \quad (118)$$

$$= -8\beta^2 \tanh \lambda_i \tanh \lambda_j (1 - \tanh^2 \lambda_i) \mathbb{E}[\mathcal{G}^2]_{u_i, u_j=0}, \quad (119)$$

$$\begin{aligned} \frac{\partial^2}{\partial u_i \partial u_j} \mathbb{E}[\mathcal{G}_i \mathcal{G}_j] \Big|_{u_i, u_j=0} &= 2\beta (1 - \mathbb{E}[\mathcal{G}_i \mathcal{G}_j]^2) + 4\beta^2 \mathbb{E}[\mathcal{G}^2] - 4\beta^2 \mathbb{E}[\mathcal{G}_i \mathcal{G}^2] - 4\beta^2 \mathbb{E}[\mathcal{G}_j \mathcal{G}^2] \\ &\quad + 8\beta^2 \mathbb{E}[\mathcal{G}_i \mathcal{G}] \mathbb{E}[\mathcal{G}_j \mathcal{G}] \mathbb{E}[\mathcal{G}_i \mathcal{G}_j] - 4\beta^2 \mathbb{E}[\mathcal{G}_i \mathcal{G}_j] \mathbb{E}[\mathcal{G}_i \mathcal{G}_j \mathcal{G}^2] \Big|_{u_i, u_j=0} \end{aligned} \quad (120)$$

$$\begin{aligned} &= 2\beta (1 - \tanh^2 \lambda_i \tanh^2 \lambda_j) + 4\beta^2 (1 - \tanh^2 \lambda_i \tanh^2 \lambda_j) \mathbb{E}[\mathcal{G}^2]_{u_i, u_j=0} \\ &\quad + 4\beta^2 (2 \tanh^2 \lambda_i \tanh^2 \lambda_j - \tanh^2 \lambda_i - \tanh^2 \lambda_j) \mathbb{E}[\mathcal{G}^2]_{u_i, u_j=0}^2. \end{aligned} \quad (121)$$

We now have the following expansion of the second moment:

$$\begin{aligned}
\mathbb{E}[\mathcal{G}_i \mathcal{G}_j] = & \mathcal{O}(u_i^3, u_j^3) + \tanh \lambda_i \tanh \lambda_j + 2\beta \tanh \lambda_j (1 - \tanh^2 \lambda_i) \mathbb{E}[\mathcal{G}]_{u_i, u_j=0} u_i \\
& + 2\beta \tanh \lambda_i (1 - \tanh^2 \lambda_j) \mathbb{E}[\mathcal{G}]_{u_i, u_j=0} u_j - 4\beta^2 \tanh \lambda_i \tanh \lambda_j (1 - \tanh^2 \lambda_i) \mathbb{E}[\mathcal{G}]_{u_i, u_j=0}^2 u_i^2 \\
& - 4\beta^2 \tanh \lambda_i \tanh \lambda_j (1 - \tanh^2 \lambda_j) \mathbb{E}[\mathcal{G}]_{u_i, u_j=0}^2 u_j^2 + 2\beta (1 - \tanh^2 \lambda_i \tanh^2 \lambda_j) u_i u_j \\
& + 4\beta^2 (1 - \tanh^2 \lambda_i \tanh^2 \lambda_j) \mathbb{E}[\mathcal{G}^2]_{u_i, u_j=0} u_i u_j \\
& + 4\beta^2 (2 \tanh^2 \lambda_i \tanh^2 \lambda_j - \tanh^2 \lambda_i - \tanh^2 \lambda_j) \mathbb{E}[\mathcal{G}]_{u_i, u_j=0}^2 u_i u_j.
\end{aligned} \tag{122}$$

Generalizing the derivation of Supplementary Equation (109), we have

$$\begin{aligned}
\mathbb{E}[\mathcal{G}]_{u_i, u_j=0} = & \mathbb{E}[\mathcal{G}] - \tanh \lambda_i u_i - \tanh \lambda_j u_j - 2\beta \tanh \lambda_i (\mathbb{E}[\mathcal{G}^2] - \mathbb{E}[\mathcal{G}]^2) u_i \\
& - 2\beta \tanh \lambda_j (\mathbb{E}[\mathcal{G}^2] - \mathbb{E}[\mathcal{G}]^2) u_j + \mathcal{O}(u_i^2, u_j^2).
\end{aligned} \tag{123}$$

Substituting this into Supplementary Equation (122) yields

$$\begin{aligned}
\mathbb{E}[\mathcal{G}_i \mathcal{G}_j] = & \tanh \lambda_i \tanh \lambda_j + 2\beta \tanh \lambda_j (1 - \tanh^2 \lambda_i) \mathbb{E}[\mathcal{G}] u_i + 2\beta \tanh \lambda_i (1 - \tanh^2 \lambda_j) \mathbb{E}[\mathcal{G}] u_j \\
& - 2\beta \tanh \lambda_i \tanh \lambda_j (1 - \tanh^2 \lambda_i) u_i^2 - 4\beta^2 \tanh \lambda_i \tanh \lambda_j (1 - \tanh^2 \lambda_i) \mathbb{E}[\mathcal{G}^2] u_i^2 \\
& - 2\beta \tanh \lambda_i \tanh \lambda_j (1 - \tanh^2 \lambda_j) u_j^2 - 4\beta^2 \tanh \lambda_i \tanh \lambda_j (1 - \tanh^2 \lambda_j) \mathbb{E}[\mathcal{G}^2] u_j^2 \\
& + 2\beta (1 - \tanh^2 \lambda_i) (1 - \tanh^2 \lambda_j) u_i u_j + 4\beta^2 (1 - \tanh^2 \lambda_i) (1 - \tanh^2 \lambda_j) \mathbb{E}[\mathcal{G}^2] u_i u_j + \mathcal{O}(u_i^3, u_j^3).
\end{aligned} \tag{124}$$

Finally, substituting Supplementary Equation (115) yields

$$\mathbb{E}[\mathcal{G}_i \mathcal{G}_j] = q_i q_j + (2\beta + 4\beta^2 \text{var}(\mathcal{G})) (1 - q_i^2) (1 - q_j^2) u_i u_j + \mathcal{O}(u_i^3, u_j^3). \tag{125}$$

This can also be written

$$\text{corr}(\mathcal{G}_i, \mathcal{G}_j) = \alpha_i \alpha_j \tag{126}$$

with

$$\alpha_i = \sqrt{(2\beta + 4\beta^2 \text{var}(\mathcal{G})) (1 - q_i^2)} u_i + \mathcal{O}(u_i^2). \tag{127}$$

### 2.2.3 Third moments

**Three distinct indices** ( $\mathbb{E}[\mathcal{G}_i \mathcal{G}_j \mathcal{G}_k]$ ) We compute the second order expansion of  $\mathbb{E}[\mathcal{G}_i \mathcal{G}_j \mathcal{G}_k]$  in  $u_i, u_j, u_k$  about 0:

Zeroth term:

$$\mathbb{E} [\mathcal{G}_i \mathcal{G}_j \mathcal{G}_k] \Big|_{u_i, u_j, u_k=0} = \tanh \lambda_i \tanh \lambda_j \tanh \lambda_k, \quad (128)$$

$$\frac{\partial}{\partial u_i} \mathbb{E} [\mathcal{G}_i \mathcal{G}_j \mathcal{G}_k] \Big|_{u_i, u_j, u_k=0} = 2\beta \mathbb{E} [\mathcal{G}_j \mathcal{G}_k \mathcal{G}] - 2\beta \mathbb{E} [\mathcal{G}_i \mathcal{G}] \mathbb{E} [\mathcal{G}_i \mathcal{G}_j \mathcal{G}_k] \Big|_{u_i, u_j, u_k=0} \quad (129)$$

$$= 2\beta \tanh \lambda_j \tanh \lambda_k (1 - \tanh^2 \lambda_i) \mathbb{E} [\mathcal{G}] \Big|_{u_i, u_j, u_k=0}, \quad (130)$$

$$\begin{aligned} \frac{\partial^2}{\partial u_i^2} \mathbb{E} [\mathcal{G}_i \mathcal{G}_j \mathcal{G}_k] \Big|_{u_i, u_j, u_k=0} &= 4\beta^2 (\mathbb{E} [\mathcal{G}_i \mathcal{G}_j \mathcal{G}_k \mathcal{G}^2] - \mathbb{E} [\mathcal{G}_i \mathcal{G}_j \mathcal{G}_k] \mathbb{E} [\mathcal{G}^2] \\ &\quad + 2\mathbb{E} [\mathcal{G}_i \mathcal{G}_j \mathcal{G}_k] \mathbb{E} [\mathcal{G}_i \mathcal{G}]^2 - 2\mathbb{E} [\mathcal{G}_i \mathcal{G}] \mathbb{E} [\mathcal{G}_j \mathcal{G}_k \mathcal{G}]) \Big|_{u_i, u_j, u_k=0} \end{aligned} \quad (131)$$

$$= -8\beta^2 \tanh \lambda_i \tanh \lambda_j \tanh \lambda_k (1 - \tanh^2 \lambda_i) \mathbb{E} [\mathcal{G}^2] \Big|_{u_i, u_j, u_k=0}, \quad (132)$$

$$\begin{aligned} \frac{\partial^2}{\partial u_i \partial u_j} \mathbb{E} [\mathcal{G}_i \mathcal{G}_j \mathcal{G}_k] \Big|_{u_i, u_j, u_k=0} &= 2\beta \tanh \lambda_k (1 - \tanh^2 \lambda_i \tanh^2 \lambda_j) + 4\beta^2 \tanh \lambda_k (1 - \tanh^2 \lambda_i \tanh^2 \lambda_j) \mathbb{E} [\mathcal{G}^2] \\ &\quad + 4\beta^2 \tanh \lambda_k (2 \tanh^2 \lambda_i \tanh^2 \lambda_j - \tanh^2 \lambda_i - \tanh^2 \lambda_j) \mathbb{E} [\mathcal{G}^2] \Big|_{u_i, u_j, u_k=0}. \end{aligned} \quad (133)$$

We now have the following expansion of the third moment:

$$\begin{aligned} \mathbb{E} [\mathcal{G}_i \mathcal{G}_j \mathcal{G}_k] &= \mathcal{O}(u_i^3, u_j^3, u_k^3) + \tanh \lambda_i \tanh \lambda_j \tanh \lambda_k \\ &\quad + 2\beta \tanh \lambda_j \tanh \lambda_k (1 - \tanh^2 \lambda_i) \mathbb{E} [\mathcal{G}]_{u_i, u_j, u_k=0} u_i \\ &\quad + 2\beta \tanh \lambda_i \tanh \lambda_k (1 - \tanh^2 \lambda_j) \mathbb{E} [\mathcal{G}]_{u_i, u_j, u_k=0} u_j \\ &\quad + 2\beta \tanh \lambda_i \tanh \lambda_j (1 - \tanh^2 \lambda_k) \mathbb{E} [\mathcal{G}]_{u_i, u_j, u_k=0} u_k \\ &\quad - 4\beta^2 \tanh \lambda_i \tanh \lambda_j \tanh \lambda_k (1 - \tanh^2 \lambda_i) \mathbb{E} [\mathcal{G}^2]_{u_i, u_j, u_k=0} u_i^2 \\ &\quad - 4\beta^2 \tanh \lambda_i \tanh \lambda_j \tanh \lambda_k (1 - \tanh^2 \lambda_j) \mathbb{E} [\mathcal{G}^2]_{u_i, u_j, u_k=0} u_j^2 \\ &\quad - 4\beta^2 \tanh \lambda_i \tanh \lambda_j \tanh \lambda_k (1 - \tanh^2 \lambda_k) \mathbb{E} [\mathcal{G}^2]_{u_i, u_j, u_k=0} u_k^2 \\ &\quad + 2\beta \tanh \lambda_k (1 - \tanh^2 \lambda_i \tanh^2 \lambda_j) u_i u_j \\ &\quad + 4\beta^2 \tanh \lambda_k (1 - \tanh^2 \lambda_i \tanh^2 \lambda_j) \mathbb{E} [\mathcal{G}^2]_{u_i, u_j, u_k=0} u_i u_j \\ &\quad + 4\beta^2 \tanh \lambda_k (2 \tanh^2 \lambda_i \tanh^2 \lambda_j - \tanh^2 \lambda_i - \tanh^2 \lambda_j) \mathbb{E} [\mathcal{G}^2]_{u_i, u_j, u_k=0} u_i u_j \\ &\quad + 2\beta \tanh \lambda_j (1 - \tanh^2 \lambda_i \tanh^2 \lambda_k) u_i u_k \\ &\quad + 4\beta^2 \tanh \lambda_j (1 - \tanh^2 \lambda_i \tanh^2 \lambda_k) \mathbb{E} [\mathcal{G}^2]_{u_i, u_j, u_k=0} u_i u_k \\ &\quad + 4\beta^2 \tanh \lambda_j (2 \tanh^2 \lambda_i \tanh^2 \lambda_k - \tanh^2 \lambda_i - \tanh^2 \lambda_k) \mathbb{E} [\mathcal{G}^2]_{u_i, u_j, u_k=0} u_i u_k \\ &\quad + 2\beta \tanh \lambda_i (1 - \tanh^2 \lambda_j \tanh^2 \lambda_k) u_j u_k \\ &\quad + 4\beta^2 \tanh \lambda_i (1 - \tanh^2 \lambda_j \tanh^2 \lambda_k) \mathbb{E} [\mathcal{G}^2]_{u_i, u_j, u_k=0} u_j u_k \end{aligned}$$

$$+ 4\beta^2 \tanh \lambda_i (2 \tanh^2 \lambda_j \tanh^2 \lambda_k - \tanh^2 \lambda_j - \tanh^2 \lambda_k) \mathbb{E} [\mathcal{G}]_{u_i, u_j, u_k=0}^2 u_j u_k. \quad (134)$$

Generalizing the derivation of Supplementary Equation (109), we have

$$\mathbb{E} [\mathcal{G}]_{u_i, u_j, u_k=0} = \mathbb{E} [\mathcal{G}] - \left(1 + 2\beta \left(\mathbb{E} [\mathcal{G}^2] - \mathbb{E} [\mathcal{G}]^2\right)\right) (\tanh \lambda_i u_i + \tanh \lambda_j u_j + \tanh \lambda_k u_k) + \mathcal{O} (u_i^2, u_j^2, u_k^2). \quad (135)$$

Substituting this into Supplementary Equation (134) yields

$$\begin{aligned} \mathbb{E} [\mathcal{G}_i \mathcal{G}_j \mathcal{G}_k] &= \tanh \lambda_i \tanh \lambda_j \tanh \lambda_k \\ &+ 2\beta \tanh \lambda_j \tanh \lambda_k (1 - \tanh^2 \lambda_i) \mathbb{E} [\mathcal{G}] u_i \\ &+ 2\beta \tanh \lambda_i \tanh \lambda_k (1 - \tanh^2 \lambda_j) \mathbb{E} [\mathcal{G}] u_j \\ &+ 2\beta \tanh \lambda_i \tanh \lambda_j (1 - \tanh^2 \lambda_k) \mathbb{E} [\mathcal{G}] u_k \\ &- 2\beta \tanh \lambda_i \tanh \lambda_j \tanh \lambda_k (1 - \tanh^2 \lambda_i) (1 + 2\beta \mathbb{E} [\mathcal{G}^2]) u_i^2 \\ &- 2\beta \tanh \lambda_i \tanh \lambda_j \tanh \lambda_k (1 - \tanh^2 \lambda_j) (1 + 2\beta \mathbb{E} [\mathcal{G}^2]) u_j^2 \\ &- 2\beta \tanh \lambda_i \tanh \lambda_j \tanh \lambda_k (1 - \tanh^2 \lambda_k) (1 + 2\beta \mathbb{E} [\mathcal{G}^2]) u_k^2 \\ &+ 2\beta \tanh \lambda_k (1 - \tanh^2 \lambda_i) (1 - \tanh^2 \lambda_j) (1 + 2\beta \mathbb{E} [\mathcal{G}^2]) u_i u_j \\ &+ 2\beta \tanh \lambda_j (1 - \tanh^2 \lambda_i) (1 - \tanh^2 \lambda_k) (1 + 2\beta \mathbb{E} [\mathcal{G}^2]) u_i u_k \\ &+ 2\beta \tanh \lambda_i (1 - \tanh^2 \lambda_j) (1 - \tanh^2 \lambda_k) (1 + 2\beta \mathbb{E} [\mathcal{G}^2]) u_j u_k \\ &+ \mathcal{O} (u_i^3, u_j^3, u_k^3). \end{aligned} \quad (136)$$

Substituting Supplementary Equation (115) yields

$$\begin{aligned} \mathbb{E} [\mathcal{G}_i \mathcal{G}_j \mathcal{G}_k] &= \mathcal{O} (u_i^3, u_j^3, u_k^3) + q_i q_j q_k + q_k (1 - q_i^2) (1 - q_j^2) (2\beta + 4\beta^2 \text{var} (\mathcal{G})) u_i u_j \\ &+ q_j (1 - q_i^2) (1 - q_k^2) (2\beta + 4\beta^2 \text{var} (\mathcal{G})) u_i u_k + q_i (1 - q_j^2) (1 - q_k^2) (2\beta + 4\beta^2 \text{var} (\mathcal{G})) u_j u_k. \end{aligned} \quad (137)$$

Notice the nontrivial terms match the expression for the second moment from Supplementary Equation (125). Thus the third central moment is zero to the second degree in  $u$ :

$$\mathbb{E} [(\mathcal{G}_i - q_i) (\mathcal{G}_j - q_j) (\mathcal{G}_k - q_k)] = \mathbb{E} [\mathcal{G}_i \mathcal{G}_j \mathcal{G}_k] - q_i \mathbb{E} [\mathcal{G}_j \mathcal{G}_k] - q_j \mathbb{E} [\mathcal{G}_i \mathcal{G}_k] - q_k \mathbb{E} [\mathcal{G}_i \mathcal{G}_j] + 2q_i q_j q_k \quad (138)$$

$$= \mathcal{O} (u_i^3, u_j^3, u_k^3). \quad (139)$$

For all values of  $u_i, u_j, u_k$ , the distribution of  $\mathcal{G}_i, \mathcal{G}_j, \mathcal{G}_k$  is fully determined by the distribution of  $\mathcal{G}_{-i,j,k}$ , independent of  $m$  (Supplementary Equation (66)). Assuming the higher moments of  $\mathcal{G}_{-i,j,k}$  do not grow

with  $m$ , an assumption we make going forward, we then have

$$\mathbb{E}[(\mathcal{G}_i - q_i)(\mathcal{G}_j - q_j)(\mathcal{G}_k - q_k)] = \mathcal{O}(m^{-3/2}). \quad (140)$$

**Two distinct indices** ( $\mathbb{E}[\mathcal{G}_i^2 \mathcal{G}_j]$ ) Let  $i, j$  be distinct indices. Using  $\mathcal{G}_i^2 = 1$  and substituting Supplementary Equation (125) yields

$$\mathbb{E}[(\mathcal{G}_i - q_i)^2 (\mathcal{G}_j - q_j)] = \mathbb{E}[\mathcal{G}_i^2 \mathcal{G}_j] - 2q_i \mathbb{E}[\mathcal{G}_i \mathcal{G}_j] + q_i^2 \mathbb{E}[\mathcal{G}_j] - q_j \mathbb{E}[\mathcal{G}_i^2] + 2q_i q_j \mathbb{E}[\mathcal{G}_i] - q_i^2 q_j \quad (141)$$

$$= -2q_i (\mathbb{E}[\mathcal{G}_i \mathcal{G}_j] - q_i q_j) \quad (142)$$

$$= -2q_i (1 - q_i^2) (1 - q_j^2) (2\beta + 4\beta^2 \text{var}(\mathcal{G})) u_i u_j + \mathcal{O}(u_i^3, u_j^3) \quad (143)$$

$$= \mathcal{O}(m^{-1}). \quad (144)$$

## 2.2.4 Fourth moments

**Four distinct indices** ( $\mathbb{E}[\mathcal{G}_i \mathcal{G}_j \mathcal{G}_k \mathcal{G}_l]$ ) Let  $i, j, k, l$  be four distinct indices. We again expand  $\mathbb{E}[\mathcal{G}_i \mathcal{G}_j \mathcal{G}_k \mathcal{G}_l]$  in  $u_i, u_j, u_k, u_l$ .

Zeroth term:

$$\mathbb{E}[\mathcal{G}_i \mathcal{G}_j \mathcal{G}_k \mathcal{G}_l] \Big|_{u_i, \dots, u_l=0} = \tanh \lambda_i \tanh \lambda_j \tanh \lambda_k \tanh \lambda_l, \quad (145)$$

$$\frac{\partial}{\partial u_i} \mathbb{E}[\mathcal{G}_i \mathcal{G}_j \mathcal{G}_k \mathcal{G}_l] \Big|_{u_i, \dots, u_l=0} = 2\beta \tanh \lambda_j \tanh \lambda_k \tanh \lambda_l (1 - \tanh^2 \lambda_i) \mathbb{E}[\mathcal{G}]_{u_i, \dots, u_l=0}, \quad (146)$$

$$\frac{\partial^2}{\partial u_i^2} \mathbb{E}[\mathcal{G}_i \mathcal{G}_j \mathcal{G}_k \mathcal{G}_l] \Big|_{u_i, \dots, u_l=0} = -8\beta^2 \tanh \lambda_i \tanh \lambda_j \tanh \lambda_k \tanh \lambda_l (1 - \tanh^2 \lambda_i) \mathbb{E}[\mathcal{G}^2]_{u_i, \dots, u_l=0}, \quad (147)$$

$$\begin{aligned} \frac{\partial^2}{\partial u_i \partial u_j} \mathbb{E}[\mathcal{G}_i \mathcal{G}_j \mathcal{G}_k \mathcal{G}_l] \Big|_{u_i, \dots, u_l=0} &= 2\beta \tanh \lambda_k \tanh \lambda_l (1 - \tanh^2 \lambda_i \tanh^2 \lambda_j) \\ &+ 4\beta^2 \tanh \lambda_k \tanh \lambda_l (1 - \tanh^2 \lambda_i \tanh^2 \lambda_j) \mathbb{E}[\mathcal{G}^2]_{u_i, \dots, u_l=0} \\ &+ 4\beta^2 \tanh \lambda_k \tanh \lambda_l (2 \tanh^2 \lambda_i \tanh^2 \lambda_j - \tanh^2 \lambda_i - \tanh^2 \lambda_j) \mathbb{E}[\mathcal{G}^2]_{u_i, \dots, u_l=0}. \end{aligned} \quad (148)$$

We now have the following expansion of the fourth moment:

$$\begin{aligned} \mathbb{E}[\mathcal{G}_i \mathcal{G}_j \mathcal{G}_k \mathcal{G}_l] &= \mathcal{O}(u_i^3, u_j^3, u_k^3, u_l^3) + \tanh \lambda_i \tanh \lambda_j \tanh \lambda_k \tanh \lambda_l \\ &+ 2\beta \tanh \lambda_j \tanh \lambda_k \tanh \lambda_l (1 - \tanh^2 \lambda_i) \mathbb{E}[\mathcal{G}]_{u_i, \dots, u_l=0} u_i \\ &+ 2\beta \tanh \lambda_i \tanh \lambda_k \tanh \lambda_l (1 - \tanh^2 \lambda_j) \mathbb{E}[\mathcal{G}]_{u_i, \dots, u_l=0} u_j \\ &+ 2\beta \tanh \lambda_i \tanh \lambda_j \tanh \lambda_l (1 - \tanh^2 \lambda_k) \mathbb{E}[\mathcal{G}]_{u_i, \dots, u_l=0} u_k \\ &+ 2\beta \tanh \lambda_i \tanh \lambda_j \tanh \lambda_k (1 - \tanh^2 \lambda_l) \mathbb{E}[\mathcal{G}]_{u_i, \dots, u_l=0} u_l \end{aligned}$$

$$\begin{aligned}
& - 4\beta^2 \tanh \lambda_i \tanh \lambda_j \tanh \lambda_k \tanh \lambda_l (1 - \tanh^2 \lambda_i) \mathbb{E} [\mathcal{G}]_{u_i, \dots, u_l=0}^2 u_i^2 \\
& - 4\beta^2 \tanh \lambda_i \tanh \lambda_j \tanh \lambda_k \tanh \lambda_l (1 - \tanh^2 \lambda_j) \mathbb{E} [\mathcal{G}]_{u_i, \dots, u_l=0}^2 u_j^2 \\
& - 4\beta^2 \tanh \lambda_i \tanh \lambda_j \tanh \lambda_k \tanh \lambda_l (1 - \tanh^2 \lambda_k) \mathbb{E} [\mathcal{G}]_{u_i, \dots, u_l=0}^2 u_k^2 \\
& - 4\beta^2 \tanh \lambda_i \tanh \lambda_j \tanh \lambda_k \tanh \lambda_l (1 - \tanh^2 \lambda_l) \mathbb{E} [\mathcal{G}]_{u_i, \dots, u_l=0}^2 u_l^2 \\
& + 2\beta \tanh \lambda_k \tanh \lambda_l (1 - \tanh^2 \lambda_i \tanh^2 \lambda_j) u_i u_j \\
& + 4\beta^2 \tanh \lambda_k \tanh \lambda_l (1 - \tanh^2 \lambda_i \tanh^2 \lambda_j) \mathbb{E} [\mathcal{G}^2]_{u_i, \dots, u_l=0} u_i u_j \\
& + 4\beta^2 \tanh \lambda_k \tanh \lambda_l (2 \tanh^2 \lambda_i \tanh^2 \lambda_j - \tanh^2 \lambda_i - \tanh^2 \lambda_j) \mathbb{E} [\mathcal{G}]_{u_i, \dots, u_l=0}^2 u_i u_j \\
& + 2\beta \tanh \lambda_j \tanh \lambda_l (1 - \tanh^2 \lambda_i \tanh^2 \lambda_k) u_i u_k \\
& + 4\beta^2 \tanh \lambda_j \tanh \lambda_l (1 - \tanh^2 \lambda_i \tanh^2 \lambda_k) \mathbb{E} [\mathcal{G}^2]_{u_i, \dots, u_l=0} u_i u_k \\
& + 4\beta^2 \tanh \lambda_j \tanh \lambda_l (2 \tanh^2 \lambda_i \tanh^2 \lambda_k - \tanh^2 \lambda_i - \tanh^2 \lambda_k) \mathbb{E} [\mathcal{G}]_{u_i, \dots, u_l=0}^2 u_i u_k \\
& + 2\beta \tanh \lambda_j \tanh \lambda_k (1 - \tanh^2 \lambda_i \tanh^2 \lambda_l) u_i u_l \\
& + 4\beta^2 \tanh \lambda_j \tanh \lambda_k (1 - \tanh^2 \lambda_i \tanh^2 \lambda_l) \mathbb{E} [\mathcal{G}^2]_{u_i, \dots, u_l=0} u_i u_l \\
& + 4\beta^2 \tanh \lambda_j \tanh \lambda_k (2 \tanh^2 \lambda_i \tanh^2 \lambda_l - \tanh^2 \lambda_i - \tanh^2 \lambda_l) \mathbb{E} [\mathcal{G}]_{u_i, \dots, u_l=0}^2 u_i u_l \\
& + 2\beta \tanh \lambda_i \tanh \lambda_l (1 - \tanh^2 \lambda_j \tanh^2 \lambda_k) u_j u_k \\
& + 4\beta^2 \tanh \lambda_i \tanh \lambda_l (1 - \tanh^2 \lambda_j \tanh^2 \lambda_k) \mathbb{E} [\mathcal{G}^2]_{u_i, \dots, u_l=0} u_j u_k \\
& + 4\beta^2 \tanh \lambda_i \tanh \lambda_l (2 \tanh^2 \lambda_j \tanh^2 \lambda_k - \tanh^2 \lambda_j - \tanh^2 \lambda_k) \mathbb{E} [\mathcal{G}]_{u_i, \dots, u_l=0}^2 u_j u_k \\
& + 2\beta \tanh \lambda_i \tanh \lambda_k (1 - \tanh^2 \lambda_j \tanh^2 \lambda_l) u_j u_l \\
& + 4\beta^2 \tanh \lambda_i \tanh \lambda_k (1 - \tanh^2 \lambda_j \tanh^2 \lambda_l) \mathbb{E} [\mathcal{G}^2]_{u_i, \dots, u_l=0} u_j u_l \\
& + 4\beta^2 \tanh \lambda_i \tanh \lambda_k (2 \tanh^2 \lambda_j \tanh^2 \lambda_l - \tanh^2 \lambda_j - \tanh^2 \lambda_l) \mathbb{E} [\mathcal{G}]_{u_i, \dots, u_l=0}^2 u_j u_l \\
& + 2\beta \tanh \lambda_i \tanh \lambda_j (1 - \tanh^2 \lambda_k \tanh^2 \lambda_l) u_k u_l \\
& + 4\beta^2 \tanh \lambda_i \tanh \lambda_j (1 - \tanh^2 \lambda_k \tanh^2 \lambda_l) \mathbb{E} [\mathcal{G}^2]_{u_i, \dots, u_l=0} u_k u_l \\
& + 4\beta^2 \tanh \lambda_i \tanh \lambda_j (2 \tanh^2 \lambda_k \tanh^2 \lambda_l - \tanh^2 \lambda_k - \tanh^2 \lambda_l) \mathbb{E} [\mathcal{G}]_{u_i, \dots, u_l=0}^2 u_k u_l. \quad (149)
\end{aligned}$$

Generalizing the derivation of Supplementary Equation (109), we have

$$\begin{aligned}
\mathbb{E} [\mathcal{G}]_{u_i, \dots, u_l=0} & = \mathcal{O} (u_i^2, u_j^2, u_k^2, u_l^2) + \mathbb{E} [\mathcal{G}] \\
& - \left( 1 + 2\beta \left( \mathbb{E} [\mathcal{G}^2] - \mathbb{E} [\mathcal{G}]^2 \right) \right) (\tanh \lambda_i u_i + \tanh \lambda_j u_j + \tanh \lambda_k u_k + \tanh \lambda_l u_l). \quad (150)
\end{aligned}$$

Substituting this into Supplementary Equation (149) yields

$$\begin{aligned}
\mathbb{E} [\mathcal{G}_i \mathcal{G}_j \mathcal{G}_k \mathcal{G}_l] & = \mathcal{O} (u_i^3, u_j^3, u_k^3, u_l^3) + \tanh \lambda_i \tanh \lambda_j \tanh \lambda_k \tanh \lambda_l \\
& + 2\beta \tanh \lambda_j \tanh \lambda_k \tanh \lambda_l (1 - \tanh^2 \lambda_i) \mathbb{E} [\mathcal{G}] u_i
\end{aligned}$$

$$\begin{aligned}
& + 2\beta \tanh \lambda_i \tanh \lambda_k \tanh \lambda_l (1 - \tanh^2 \lambda_j) \mathbb{E}[\mathcal{G}] u_j \\
& + 2\beta \tanh \lambda_i \tanh \lambda_j \tanh \lambda_l (1 - \tanh^2 \lambda_k) \mathbb{E}[\mathcal{G}] u_k \\
& + 2\beta \tanh \lambda_i \tanh \lambda_j \tanh \lambda_k (1 - \tanh^2 \lambda_l) \mathbb{E}[\mathcal{G}] u_l \\
& - 2\beta \tanh \lambda_i \tanh \lambda_j \tanh \lambda_k \tanh \lambda_l (1 - \tanh^2 \lambda_i) (1 + 2\beta \mathbb{E}[\mathcal{G}^2]) u_i^2 \\
& - 2\beta \tanh \lambda_i \tanh \lambda_j \tanh \lambda_k \tanh \lambda_l (1 - \tanh^2 \lambda_j) (1 + 2\beta \mathbb{E}[\mathcal{G}^2]) u_j^2 \\
& - 2\beta \tanh \lambda_i \tanh \lambda_j \tanh \lambda_k \tanh \lambda_l (1 - \tanh^2 \lambda_k) (1 + 2\beta \mathbb{E}[\mathcal{G}^2]) u_k^2 \\
& - 2\beta \tanh \lambda_i \tanh \lambda_j \tanh \lambda_k \tanh \lambda_l (1 - \tanh^2 \lambda_l) (1 + 2\beta \mathbb{E}[\mathcal{G}^2]) u_l^2 \\
& + 2\beta \tanh \lambda_k \tanh \lambda_l (1 - \tanh^2 \lambda_i) (1 - \tanh^2 \lambda_j) (1 + 2\beta \mathbb{E}[\mathcal{G}^2]) u_i u_j \\
& + 2\beta \tanh \lambda_j \tanh \lambda_l (1 - \tanh^2 \lambda_i) (1 - \tanh^2 \lambda_k) (1 + 2\beta \mathbb{E}[\mathcal{G}^2]) u_i u_k \\
& + 2\beta \tanh \lambda_j \tanh \lambda_k (1 - \tanh^2 \lambda_i) (1 - \tanh^2 \lambda_l) (1 + 2\beta \mathbb{E}[\mathcal{G}^2]) u_i u_l \\
& + 2\beta \tanh \lambda_i \tanh \lambda_l (1 - \tanh^2 \lambda_j) (1 - \tanh^2 \lambda_k) (1 + 2\beta \mathbb{E}[\mathcal{G}^2]) u_j u_k \\
& + 2\beta \tanh \lambda_i \tanh \lambda_k (1 - \tanh^2 \lambda_j) (1 - \tanh^2 \lambda_l) (1 + 2\beta \mathbb{E}[\mathcal{G}^2]) u_j u_l \\
& + 2\beta \tanh \lambda_i \tanh \lambda_j (1 - \tanh^2 \lambda_k) (1 - \tanh^2 \lambda_l) (1 + 2\beta \mathbb{E}[\mathcal{G}^2]) u_k u_l. \tag{151}
\end{aligned}$$

Substituting Supplementary Equation (115) yields

$$\begin{aligned}
\mathbb{E}[\mathcal{G}_i \mathcal{G}_j \mathcal{G}_k \mathcal{G}_l] & = \mathcal{O}(u_i^3, u_j^3, u_k^3, u_l^3) + q_i q_j q_k q_l \\
& + 2\beta q_k q_l (1 - q_i^2) (1 - q_j^2) (1 + 2\beta \text{var}(\mathcal{G}^2)) u_i u_j \\
& + 2\beta q_j q_l (1 - q_i^2) (1 - q_k^2) (1 + 2\beta \text{var}(\mathcal{G}^2)) u_i u_k \\
& + 2\beta q_j q_k (1 - q_i^2) (1 - q_l^2) (1 + 2\beta \text{var}(\mathcal{G}^2)) u_i u_l \\
& + 2\beta q_i q_l (1 - q_j^2) (1 - q_k^2) (1 + 2\beta \text{var}(\mathcal{G}^2)) u_j u_k \\
& + 2\beta q_i q_k (1 - q_j^2) (1 - q_l^2) (1 + 2\beta \text{var}(\mathcal{G}^2)) u_j u_l \\
& + 2\beta q_i q_j (1 - q_k^2) (1 - q_l^2) (1 + 2\beta \text{var}(\mathcal{G}^2)) u_k u_l. \tag{152}
\end{aligned}$$

As with the third moment, the nontrivial terms here match the expression for the second moment from Supplementary Equation (125). Together with Supplementary Equation (139), this implies the fourth central moment is zero to second degree in  $u$ :

$$\begin{aligned}
\mathbb{E}[(\mathcal{G}_i - q_i)(\mathcal{G}_j - q_j)(\mathcal{G}_k - q_k)(\mathcal{G}_l - q_l)] & = \mathbb{E}[\mathcal{G}_i \mathcal{G}_j \mathcal{G}_k \mathcal{G}_l] - q_i \mathbb{E}[(\mathcal{G}_j - q_j)(\mathcal{G}_k - q_k)(\mathcal{G}_l - q_l)] - q_j \mathbb{E}[(\mathcal{G}_i - q_i)(\mathcal{G}_k - q_k)(\mathcal{G}_l - q_l)] \\
& - q_k \mathbb{E}[(\mathcal{G}_i - q_i)(\mathcal{G}_j - q_j)(\mathcal{G}_l - q_l)] - q_l \mathbb{E}[(\mathcal{G}_i - q_i)(\mathcal{G}_j - q_j)(\mathcal{G}_k - q_k)] \\
& - q_i q_j \mathbb{E}[\mathcal{G}_k \mathcal{G}_l] - q_i q_k \mathbb{E}[\mathcal{G}_j \mathcal{G}_l] - q_i q_l \mathbb{E}[\mathcal{G}_k \mathcal{G}_j] - q_j q_k \mathbb{E}[\mathcal{G}_i \mathcal{G}_l] - q_j q_l \mathbb{E}[\mathcal{G}_i \mathcal{G}_k] \\
& - q_k q_l \mathbb{E}[\mathcal{G}_i \mathcal{G}_j] + 5q_i q_j q_k q_l \tag{153}
\end{aligned}$$



$$= \mathcal{O}(u_i^3, u_j^3, u_k^3, u_l^3) = \mathcal{O}(m^{-3/2}). \quad (154)$$

**Mixed fourth moments** Here we consider central fourth moments with repeated indices. These reduce to lower-order moments using  $\mathcal{G}_i^2 = 1$ .

Let  $i, j, k$  be any three distinct indices. Simplifying and then substituting Supplementary Equations (125) and (137) yields

$$\begin{aligned} \mathbb{E} \left[ (\mathcal{G}_i - q_i)^2 (\mathcal{G}_j - q_j) (\mathcal{G}_k - q_k) \right] &= \mathbb{E} [\mathcal{G}_i^2 \mathcal{G}_j \mathcal{G}_k] - 2q_i \mathbb{E} [\mathcal{G}_i \mathcal{G}_j \mathcal{G}_k] + q_i^2 \mathbb{E} [\mathcal{G}_j \mathcal{G}_k] - q_j \mathbb{E} [\mathcal{G}_i^2 \mathcal{G}_k] \\ &\quad + 2q_i q_j \mathbb{E} [\mathcal{G}_i \mathcal{G}_k] - q_k \mathbb{E} [\mathcal{G}_i^2 \mathcal{G}_j] + 2q_i q_k \mathbb{E} [\mathcal{G}_i \mathcal{G}_j] + q_j q_k \mathbb{E} [\mathcal{G}_i^2] - 3q_i^2 q_j q_k \end{aligned} \quad (155)$$

$$= (1 - q_i^2) (1 - q_j^2) (1 - q_k^2) (2\beta + 4\beta^2 \text{var}(\mathcal{G})) u_j u_k + \mathcal{O}(u_i^3, u_j^3, u_k^3) \quad (156)$$

$$= \mathcal{O}(m^{-1}). \quad (157)$$

Now consider  $(\mathcal{G}_i - q_i)^3 (\mathcal{G}_j - q_j)$ :

$$\begin{aligned} \mathbb{E} \left[ (\mathcal{G}_i - q_i)^3 (\mathcal{G}_j - q_j) \right] &= \mathbb{E} [\mathcal{G}_i^3 \mathcal{G}_j] - 3q_i \mathbb{E} [\mathcal{G}_i^2 \mathcal{G}_j] + 3q_i^2 \mathbb{E} [\mathcal{G}_i \mathcal{G}_j] - q_i^3 \mathbb{E} [\mathcal{G}_j] \\ &\quad - q_j \mathbb{E} [\mathcal{G}_i^3] + 3q_i q_j \mathbb{E} [\mathcal{G}_i^2] - 3q_i^2 q_j \mathbb{E} [\mathcal{G}_i] + q_i^3 q_j \end{aligned} \quad (158)$$

$$= (1 + 3q_i^2) (\mathbb{E} [\mathcal{G}_i \mathcal{G}_j] - q_i q_j) \quad (159)$$

$$= (1 - q_i^2) (1 - q_j^2) (1 + 3q_i^2) (2\beta + 4\beta^2 \text{var}(\mathcal{G})) u_i u_j + \mathcal{O}(u_i^3, u_j^3) \quad (160)$$

$$= \mathcal{O}(m^{-1}). \quad (161)$$

Last, consider  $(\mathcal{G}_i - q_i)^2 (\mathcal{G}_j - q_j)^2$ :

$$\begin{aligned} \mathbb{E} \left[ (\mathcal{G}_i - q_i)^2 (\mathcal{G}_j - q_j)^2 \right] &= \mathbb{E} [\mathcal{G}_i^2 \mathcal{G}_j^2] - 2q_j \mathbb{E} [\mathcal{G}_i^2 \mathcal{G}_j] + q_j^2 \mathbb{E} [\mathcal{G}_i^2] - 2q_i \mathbb{E} [\mathcal{G}_i \mathcal{G}_j^2] \\ &\quad + 4q_i q_j \mathbb{E} [\mathcal{G}_i \mathcal{G}_j] - 2q_i q_j^2 \mathbb{E} [\mathcal{G}_i] + q_i^2 \mathbb{E} [\mathcal{G}_j^2] - 2q_i^2 q_j \mathbb{E} [\mathcal{G}_j] + q_i^2 q_j^2 \end{aligned} \quad (162)$$

$$= (1 - q_i^2) (1 - q_j^2) + 4q_i q_j (\mathbb{E} [\mathcal{G}_i \mathcal{G}_j] - q_i q_j) \quad (163)$$

$$= (1 - q_i^2) (1 - q_j^2) + 4q_i q_j (1 - q_i^2) (1 - q_j^2) (2\beta + 4\beta^2 \text{var}(\mathcal{G})) u_i u_j + \mathcal{O}(u_i^3, u_j^3). \quad (164)$$

Therefore the corresponding standardized moment is bound as

$$\mathbb{E} \left[ \frac{(\mathcal{G}_i - q_i)^2 (\mathcal{G}_j - q_j)^2}{1 - q_i^2} \frac{1}{1 - q_j^2} \right] = 1 + 4q_i q_j (2\beta + 4\beta^2 \text{var}(\mathcal{G})) u_i u_j + \mathcal{O}(u_i^3, u_j^3) \quad (165)$$

$$= 1 + \mathcal{O}(m^{-1}). \quad (166)$$

## 2.3 The limiting spectral distribution of covariance matrices for a class of random matrices with dependent entries

Here we establish the limiting spectral distribution for a class of random matrices with dependent entries to which our matrix of standardized genotypes  $Z$  belongs. For simplicity, we will assume that the equilibrium covariances among all causal variants is constant, though our results are readily extended to the general case where  $\mathbb{E}[\{n^{-1}Z^T Z\}_{ij}] = \mathcal{O}(1/m)$  for all  $i, j \neq i$ .

### 2.3.1 Introduction and primary result

For an  $n \times n$  Hermitian matrix  $A$ , we let  $\lambda_1(A), \dots, \lambda_n(A) \in \mathbb{R}$  denote the eigenvalues of  $A$ . The empirical spectral distribution (ESD)  $F^A$  of  $A$  is defined as

$$F^A(x) := \frac{1}{n} |\{1 \leq i \leq n : \lambda_i(A) \leq x\}|, \quad (167)$$

where  $|E|$  denotes the cardinality of the finite set  $E$ . The Marčenko–Pastur distribution  $F_\tau^{\text{MP}}$  with parameter  $\tau > 0$  is the probability distribution function with density

$$f_\tau^{\text{MP}}(x) := \begin{cases} \frac{1}{2\pi\tau x} \sqrt{(b-x)(x-a)}, & \text{if } a \leq x \leq b, \\ 0, & \text{otherwise,} \end{cases} \quad (168)$$

and a point mass  $1 - 1/\tau$  at the origin if  $\tau > 1$ , where  $a := (1 - \sqrt{\tau})^2$  and  $b := (1 + \sqrt{\tau})^2$ .

Let  $z \in \mathbb{R}^p$  be a random vector. We let  $Z$  be an  $n \times p$  matrix whose rows are independent and identically distributed (iid) copies of  $z$ . We will be interested in the eigenvalues of  $\frac{1}{p}ZZ^T$ . We view  $n$  as a large asymptotic parameter tending to infinity;  $p$  and  $m$  will be two additional parameters which we assume tend to infinity with  $n$ . Asymptotic notation (such as  $\mathcal{O}$ ,  $o$ ) will be used under the assumption that  $n, m, p \rightarrow \infty$ .

For the model under consideration, we make the following assumptions regarding the random vector  $z$ . The random vector  $z$  has mean zero and covariance matrix

$$\Sigma := \begin{bmatrix} \Sigma' & 0 \\ 0 & I_{p-m} \end{bmatrix}, \quad (169)$$

where  $\Sigma'$  is an  $m \times m$  matrix with entries

$$\Sigma'_{ij} := \begin{cases} 2\mu(m), & \text{if } i \neq j, \\ 1 + \mu(m), & \text{if } i = j. \end{cases} \quad (170)$$

Here  $I_n$  denotes the  $n \times n$  identity matrix. Additionally, we assume  $\mu(m) = \mathcal{O}(1/m)$ . We also assume the following about the entries of  $z = (z_i)_{i=1}^p$ :

1. There exists a constant  $\kappa > 0$  so that  $\sup_{1 \leq i \leq p} |z_i| \leq \kappa$  with probability 1.
2.  $\mathbb{E}[z_i^2 z_j^2] = 1 + o(1)$  uniformly for all distinct  $i, j$ .
3.  $\mathbb{E}[z_i^3 z_j] = o(1)$  uniformly for all distinct  $i, j$ .
4. One has

$$\mathbb{E}[z_i z_j z_k z_l] = \begin{cases} o\left(\frac{1}{m}\right), & \text{if } |\{i, j, k, l\}| = 4, \\ o\left(\frac{1}{\sqrt{m}}\right), & \text{if } |\{i, j, k, l\}| = 3, \end{cases} \quad (171)$$

uniformly in  $i, j, k, l$ .

**Theorem 1.** *Suppose  $\frac{n}{p} \rightarrow \tau \in (0, \infty)$  as  $n \rightarrow \infty$  and  $m \geq cp$  for some constant  $c > 0$ . Then, under the assumptions above, the ESD  $F^{\frac{1}{p}ZZ^T}$  of  $\frac{1}{p}ZZ^T$  converges almost surely to the Marčenko–Pastur distribution  $F_\tau^{\text{MP}}$  as  $n \rightarrow \infty$ .*

The remainder of this section is devoted to the proof of Theorem 1. Throughout, we assume  $m, n, p, z$ , and  $Z$  satisfy the assumptions of Theorem 1.

### 2.3.2 Notation and overview

$I_n$  denotes the  $n \times n$  identity matrix. Often we will simply write  $I$  when the size can be deduced from context. For a matrix  $A$ , we let  $\|A\|$  denote the spectral (operator) norm of  $A$  and  $\|A\|_2$  denote the Hilbert–Schmidt (Frobenius) norm of  $A$  defined by

$$\|A\|_2 := \sqrt{\text{tr}(AA^*)} = \sqrt{\text{tr}(A^*A)}, \quad (172)$$

where  $A^*$  denotes the conjugate transpose of  $A$ . We will exploit the fact that the spectral norm of  $A$  is an upper bound for the spectral norm of any sub-matrix of  $A$ .

For  $\alpha \in \mathbb{C}^+ := \{w \in \mathbb{C} : \text{Im}(w) > 0\}$ , we define the resolvent of  $\frac{1}{p}ZZ^T$  as

$$R(\alpha) := \left( \frac{1}{p}ZZ^T - \alpha I \right)^{-1}. \quad (173)$$

$Z^{(k)}$  is the  $(n-1) \times p$  matrix constructed from  $Z$  by removing the  $k$ -th row. Let  $R^{(k)}$  be the resolvent of  $\frac{1}{p}Z^{(k)}Z^{(k)T}$ , i.e.,

$$R^{(k)}(\alpha) := \left( \frac{1}{p}Z^{(k)}Z^{(k)T} - \alpha I \right)^{-1}, \quad \alpha \in \mathbb{C}^+. \quad (174)$$

Define  $s_n$  to be the Stieltjes transform of the ESD  $F^{\frac{1}{p}ZZ^T}$ . In other words,

$$s_n(\alpha) := \frac{1}{n} \operatorname{tr} R(\alpha), \quad \alpha \in \mathbb{C}^+. \quad (175)$$

We let  $s$  be the Stieltjes transform of  $F_\tau^{\text{MP}}$ . From Chapter 3 of Bai and Silverstein<sup>7</sup> it follows that  $s$  satisfies the equation

$$s(\alpha) = \frac{1}{1 - \alpha - \tau - \tau \alpha s(\alpha)}, \quad \alpha \in \mathbb{C}^+. \quad (176)$$

We prove Theorem 1 by showing that  $s_n$  converges to  $s$  as  $n$  tends to infinity. This is a standard proof method in random matrix theory (see e.g. Bai and Silverstein<sup>7</sup> for details). However, the dependence between the entries introduces new difficulties in establishing this convergence. The main technical advance in this work involves dealing with the dependence amongst the entries.

In Section 2.3.3 we present the main tools from linear algebra and probability theory that are required for the proof. A concentration inequality is presented in Section 2.3.4 which allows us to pass between the Stieltjes transform and its expected value. Section 2.3.5 contains a version of the Hanson–Wright inequality for the rows of  $Z$ ; this section is the main technical tool for dealing with the dependence amongst the entries. We finally complete the proof of Theorem 1 in Section 2.3.6

### 2.3.3 Tools from probability theory and linear algebra

We begin by introducing some tools from probability theory and linear algebra.

**Lemma 2.** *For any  $\alpha \in \mathbb{C}^+$  and any  $1 \leq k \leq n$ ,*

$$\left| \operatorname{tr} R(\alpha) - \operatorname{tr} R^{(k)}(\alpha) \right| \leq \frac{1}{\operatorname{Im}(\alpha)}. \quad (177)$$

*Proof.* This follows from equation (A.1.12) in Bai and Silverstein<sup>7</sup>. □

**Lemma 3.** *If  $\mu$  is a probability measure supported on  $[0, \infty)$  and  $s_\mu$  is its Stieltjes transform defined by*

$$s_\mu(\alpha) := \int_0^\infty \frac{1}{x - \alpha} d\mu(x), \quad \alpha \in \mathbb{C}^+, \quad (178)$$

*then  $|s_\mu(\alpha)| \leq \frac{1}{\operatorname{Im}(\alpha)}$  and*

$$\operatorname{Im} s_\mu(\alpha) > 0, \quad \operatorname{Im}(\alpha s_\mu(\alpha)) \geq 0 \quad (179)$$

*for all  $\alpha \in \mathbb{C}^+$ .*

*Proof.* The first bound follows from the triangle inequality:

$$|s_\mu(\alpha)| \leq \int_0^\infty \frac{1}{|x - \alpha|} d\mu(x) \leq \int_0^\infty \frac{1}{\operatorname{Im}\alpha} d\mu(x) = \frac{1}{\operatorname{Im}\alpha}. \quad (180)$$

For the other bounds, since

$$s_\mu(\alpha) = \int_0^\infty \frac{x - \bar{\alpha}}{|x - \alpha|^2} d\mu(x), \quad (181)$$

it follows that  $\operatorname{Im}s_\mu(\alpha) > 0$ . Similarly,

$$\operatorname{Im}(\alpha s_\mu(\alpha)) = \operatorname{Im}(\alpha) \int_0^\infty \frac{x}{|x - \alpha|^2} d\mu(x) \geq 0, \quad (182)$$

which completes the proof.  $\square$

**Lemma 4 (Stability).** *Suppose  $\mu$  is a probability measure supported on  $[0, \infty)$  and  $s_\mu$  is its Stieltjes transform defined in Supplementary Equation (178). If*

$$s_\mu(\alpha) = \frac{1}{1 - \alpha - \tau - \tau\alpha s_\mu(\alpha)} + \varepsilon \quad (183)$$

for some  $\alpha \in \mathbb{C}^+$ , some  $\varepsilon \in \mathbb{C}$ , and some  $\tau > 0$ , then

$$|s_\mu(\alpha) - s(\alpha)| \leq \frac{|\varepsilon|}{\operatorname{Im}\alpha} |1 - \alpha - \tau - \tau\alpha s_\mu(\alpha)|, \quad (184)$$

where  $s$  is the Stieltjes transform of the Marčenko–Pastur distribution function  $F_y^{\text{MP}}$  (see Supplementary Equation (176)).

*Proof.* Using Supplementary Equations (176) and (183), we have

$$-\tau\alpha s_\mu^2(\alpha) + s_\mu(\alpha)(1 - \alpha - \tau) - 1 = \varepsilon(1 - \alpha - \tau - \tau\alpha s_\mu(\alpha)) \quad (185)$$

and

$$-\tau\alpha s^2(\alpha) + s(\alpha)(1 - \alpha - \tau) - 1 = 0. \quad (186)$$

Subtracting the two equations gives

$$|s_\mu(\alpha) - s(\alpha)| |-\tau\alpha(s_\mu(\alpha) + s(\alpha)) + 1 - \alpha - \tau| = |\varepsilon| |1 - \alpha - \tau - \tau\alpha s_\mu(\alpha)|. \quad (187)$$

In view of Lemma 3,

$$|-\tau\alpha(s_\mu(\alpha) + s(\alpha)) + 1 - \alpha - \tau| \geq \tau \operatorname{Im}(\alpha(s_\mu(\alpha) + s(\alpha))) + \operatorname{Im}(\alpha) \geq \operatorname{Im}(\alpha), \quad (188)$$

and the conclusion follows.  $\square$

Throughout the proof, we will utilize the resolvent identity:

$$A^{-1} - B^{-1} = A^{-1}(B - A)B^{-1}, \quad (189)$$

which holds for invertible matrices  $A$  and  $B$ . We will also use the bound

$$\|(A - \alpha I)^{-1}\| \leq \frac{1}{\text{Im}(\alpha)} \quad (190)$$

which follows from the spectral theorem and holds for any Hermitian matrix  $A$  and any  $\alpha \in \mathbb{C}^+$ . We will also need the following bound.

**Lemma 5.** *Let  $A$  be an  $n \times p$  matrix. Then*

$$A^*(AA^* - \alpha I)^{-1}A = (A^*A - \alpha I)^{-1}A^*A \quad (191)$$

for all  $\alpha \in \mathbb{C}^+$ . In addition,

$$\|A^*(AA^* - \alpha I)^{-1}A\| \leq 1 + \frac{|\alpha|}{\text{Im}\alpha} \quad (192)$$

for all  $\alpha \in \mathbb{C}^+$ .

*Proof.* For  $|\alpha| > \|AA^*\|$  the Neumann series for the resolvent  $(AA^* - \alpha I)^{-1}$  gives

$$(AA^* - \alpha I)^{-1} = -\frac{1}{\alpha}I - \sum_{k=1}^{\infty} \frac{(AA^*)^k}{\alpha^{k+1}}. \quad (193)$$

It follows that

$$A^*(AA^* - \alpha I)^{-1}A = -\frac{A^*A}{\alpha} - \sum_{k=1}^{\infty} \frac{(A^*A)^k A^*A}{\alpha^{k+1}} = (A^*A - \alpha I)^{-1}A^*A. \quad (194)$$

This establishes Supplementary Equation (191) for  $\alpha \in \mathbb{C}^+$  with  $|\alpha| > \|AA^*\|$ . The identity can be extended to all of  $\mathbb{C}^+$  by analytic continuation since the entries of the resolvent are analytic in  $\mathbb{C}^+$ . Alternatively, Supplementary Equation (191) can be derived using the singular value decomposition for  $A$ .

The bound in Supplementary Equation (192) can be established using Supplementary Equation (191). Indeed, by Supplementary Equation (191)

$$A^*(AA^* - \alpha I)^{-1}A = (A^*A - \alpha I)^{-1}A^*A \quad (195)$$

$$= (A^*A - \alpha I)^{-1}(A^*A - \alpha I + \alpha I) \quad (196)$$

$$= I + \alpha(A^*A - \alpha I)^{-1}. \quad (197)$$

We conclude that

$$\|A^*(AA^* - \alpha I)^{-1}A\| \leq 1 + \frac{\alpha}{\text{Im}(\alpha)} \quad (198)$$

by the triangle inequality and Supplementary Equation (190).  $\square$

We will also take advantage of the bound

$$|\text{tr}(A)| \leq \text{rank}(A)\|A\|, \quad (199)$$

which holds for any square matrix  $A$  and follows from Corollary (A.12) in Bai and Silverstein<sup>7</sup>.

### 2.3.4 Concentration

Recall the Stieltjes transform  $s_n$  defined in Supplementary Equation (175). The following results show that the Stieltjes transform  $s_n$  concentrates around its expected value  $\mathbb{E}[s_n]$ .

**Lemma 6.** *For any  $\alpha \in \mathbb{C}^+$  and any  $t > 0$ ,*

$$\mathbb{P}(|s_n(\alpha) - \mathbb{E}[s_n(\alpha)]| > t) \leq C \exp(-cnt^2(\text{Im}\alpha)^2), \quad (200)$$

where  $C, c > 0$  are absolute constants.

*Proof.* Let  $\tilde{Z}^{(k)}$  be formed from the matrix  $Z$  by replacing the  $k$ -th row with zeros. Let  $\tilde{R}^{(k)}$  be the resolvent of  $\tilde{Z}^{(k)}\tilde{Z}^{(k)T}$  defined by

$$\tilde{R}^{(k)}(\alpha) := \left(\tilde{Z}^{(k)}\tilde{Z}^{(k)T} - \alpha I\right)^{-1}, \quad \alpha \in \mathbb{C}^+. \quad (201)$$

Then for any  $1 \leq k \leq n$ ,

$$\left|\text{tr} R(\alpha) - \text{tr} \tilde{R}^{(k)}(\alpha)\right| \leq \left|\text{tr} R(\alpha) - \text{tr} R^{(k)}(\alpha)\right| + \left|\text{tr} R^{(k)}(\alpha) - \text{tr} \tilde{R}^{(k)}(\alpha)\right|. \quad (202)$$

The first term on the right-hand side can be bounded using Lemma 2. The second term can be controlled using the fact that  $Z^{(k)}Z^{(k)T}$  and  $\tilde{Z}^{(k)}\tilde{Z}^{(k)T}$  have the same eigenvalues except  $\tilde{Z}^{(k)}\tilde{Z}^{(k)T}$  has one additional zero eigenvalue. In other words,

$$\left|\text{tr} R^{(k)}(\alpha) - \text{tr} \tilde{R}^{(k)}(\alpha)\right| = \left|\frac{1}{\alpha}\right| \leq \frac{1}{\text{Im}(\alpha)}. \quad (203)$$

We conclude that

$$\left|\text{tr} R(\alpha) - \text{tr} \tilde{R}^{(k)}(\alpha)\right| \leq \frac{2}{\text{Im}(\alpha)} \quad (204)$$

for any  $1 \leq k \leq n$ . Now since  $\text{tr } R(\alpha)$  is a function of the independent random vectors  $z_1, \dots, z_n$ , McDiarmid's inequality (see<sup>8</sup> or Theorem 2.1.10 in<sup>9</sup>) implies that

$$\mathbb{P}(|\text{tr } R(\alpha) - \mathbb{E} \text{tr } R(\alpha)| > t) \leq C \exp\left(-c(\text{Im}\alpha)^2 \frac{t^2}{n}\right) \quad (205)$$

for any  $t > 0$ , where  $C, c > 0$  are absolute constants. The claim now follows by rescaling.  $\square$

**Lemma 7.** *Define the domain*

$$S := \left\{ \alpha \in \mathbb{C}^+ : 1 \geq \text{Im}(\alpha) \geq \frac{\log n}{n^{1/4}}, -n \leq \text{Re}(\alpha) \leq n \right\}. \quad (206)$$

One has

$$\mathbb{P}\left(\sup_{\alpha \in S} |s_n(\alpha) - \mathbb{E}[s_n(\alpha)]| > 2n^{-1/4}\right) \leq C \exp(-c \log^2 n), \quad (207)$$

where  $C, c > 0$  are absolute constants.

We will utilize an  $\varepsilon$ -net in order to prove Lemma 7.s

**Definition (Nets).** Let  $X$  be a subset of  $\mathbb{C}$  and  $\varepsilon > 0$ . A subset  $\mathcal{N}$  of  $X$  is called an  $\varepsilon$ -net of  $X$  if every point  $x \in X$  can be approximated within  $\varepsilon$  by some point  $y \in \mathcal{N}$ , i.e. so that  $|x - y| \leq \varepsilon$ .

*Proof of Lemma 7.* By Lemma 6,

$$\sup_{\alpha \in S} \mathbb{P}\left(|s_n(\alpha) - \mathbb{E}[s_n(\alpha)]| > n^{-1/4}\right) \leq C' \exp(-c' \log^2 n), \quad (208)$$

where  $C', c' > 0$  are absolute constants. The goal is to extend this bound to simultaneously hold for all  $\alpha \in S$ . We will use a net to do so. Let  $\mathcal{N}$  be a  $n^{-1}$ -net of  $S$ . A simple volume argument (see for instance<sup>10</sup>) shows that  $\mathcal{N}$  can be chosen so that

$$|\mathcal{N}| \leq C_0 n^{C_0} \quad (209)$$

for an absolute constant  $C_0 > 0$ . By the union bound and Supplementary Equation (208), we find

$$\mathbb{P}\left(\sup_{\alpha \in \mathcal{N}} |s_n(\alpha) - \mathbb{E}[s_n(\alpha)]| > n^{-1/4}\right) \leq C \exp(-c \log^2 n), \quad (210)$$

where  $C, c > 0$  are absolute constants.

We now wish to extend this bound from  $\mathcal{N}$  to  $S$ . To this end, suppose there exists  $\alpha \in S$  so that  $|s_n(\alpha) - \mathbb{E}s_n(\alpha)| > 2n^{-1/4}$ . Then there exists  $\alpha' \in \mathcal{N}$  such that  $|\alpha - \alpha'| \leq \frac{1}{n}$ . Thus, by the resolvent identity in Supplementary Equation (189), Supplementary Equation (199), and Supplementary Equation (190), we have

$$\left| \frac{1}{n} \text{tr } R(\alpha) - \frac{1}{n} \text{tr } R(\alpha') \right| \leq |\alpha - \alpha'| \|R(\alpha)\| \|R(\alpha')\| \leq \frac{|\alpha - \alpha'|}{\text{Im}(\alpha) \text{Im}(\alpha')}. \quad (211)$$



The same bound holds for the expected value (by the same argument) and so we deduce that

$$|(s_n(\alpha) - \mathbb{E}[s_n(\alpha)]) - (s_n(\alpha') - \mathbb{E}[s_n(\alpha')])| \leq 2 \frac{\sqrt{n}}{\log^2 n} |\alpha - \alpha'| < n^{-1/4} \quad (212)$$

for  $n$  sufficiently large since  $|\alpha - \alpha'| \leq n^{-1}$ . Here we bounded the imaginary parts of  $\alpha$  and  $\alpha'$  from below by  $n^{-1/4} \log n$  using the definition of the set  $S$ . Since  $|s_n(\alpha) - \mathbb{E}s_n(\alpha)| > 2n^{-1/4}$ , it follows that  $|s_n(\alpha') - \mathbb{E}s_n(\alpha')| > n^{-1/4}$ . Summarizing, we have shown that

$$\mathbb{P}\left(\sup_{\alpha \in S} |s_n(\alpha) - \mathbb{E}[s_n(\alpha)]| > 2n^{-1/4}\right) \leq \mathbb{P}\left(\sup_{\alpha \in \mathcal{N}} |s_n(\alpha) - \mathbb{E}[s_n(\alpha)]| > n^{-1/4}\right). \quad (213)$$

Thus, the claim follows from Supplementary Equation (210).  $\square$

### 2.3.5 Hanson–Wright inequality

Recall that  $z$  is a random vector in  $\mathbb{R}^p$  with mean zero, and the rows of  $Z$  are iid copies of  $z$ . In addition,  $\Sigma$  is the covariance matrix of  $z$ . The Hanson–Wright inequality is a concentration inequality which shows that the quadratic form  $z^T A z$  concentrates around its expected value, where  $A$  here is an arbitrary deterministic  $p \times p$  matrix. If the entries of  $z$  were independent, the bound would follow from more classical results such as<sup>11,12</sup>. In the model under consideration the entries of  $z$  are dependent, and this dependence requires different techniques to deal with the quadratic form  $z^T A z$ . We begin first with its expected value.

**Lemma 8.** *For a deterministic  $p \times p$  matrix  $A$ ,*

$$\frac{1}{p} \mathbb{E}[z^T A z] = \frac{1}{p} \text{tr}(A \Sigma) = \frac{1}{p} \text{tr}(A) + \mathcal{O}(m^{-1} \|A\|). \quad (214)$$

*Proof.* We denote  $z = (z_i)_{i=1}^p$ . Then

$$\mathbb{E}[z^T A z] = \sum_{i,j=1}^p A_{ij} \mathbb{E}[z_i z_j] = \sum_{i,j=1}^p A_{ij} \Sigma_{ij} = \text{tr}(A \Sigma^T) = \text{tr}(A \Sigma). \quad (215)$$

This establishes the first equality in Supplementary Equation (214).

We now turn to the second equality in Supplementary Equation (214). Recall that

$$\Sigma = \begin{bmatrix} \Sigma' & 0 \\ 0 & I_{p-m} \end{bmatrix}, \quad (216)$$

where  $\Sigma'$  is the  $m \times m$  matrix defined in Supplementary Equation (170). We similarly decompose  $A$  in block

form as

$$A = \begin{bmatrix} A_1 & A_2 \\ A_3 & A_4 \end{bmatrix}, \quad (217)$$

where  $A_1$  is  $m \times m$  and  $A_4$  is  $(p - m) \times (p - m)$ . Then  $\text{tr}(A\Sigma) = \text{tr}(A_1\Sigma') + \text{tr}(A_4)$ . We rewrite  $\Sigma' = 2\mu(m)J + (1 - \mu(m))I$ , where  $J$  is the all-ones matrix. So, we obtain

$$\text{tr}(A_1\Sigma') = (1 - \mu(m))\text{tr}(A_1) + 2\mu(m)\text{tr}(A_1J). \quad (218)$$

Since  $J$  is rank one, we can use Supplemental Equation (199) to bound  $|\text{tr}(A_1J)| \leq \|A_1\| \|J\| \leq \|A\| \|J\|$ . A simple computation shows that  $\|J\| = m$ , and using  $\mu(m) = \mathcal{O}(m^{-1})$ , we conclude that

$$\frac{1}{p}\text{tr}(A_1\Sigma') = \frac{1}{p}\text{tr}(A_1) + \mathcal{O}(m^{-1}\|A\|), \quad (219)$$

where we used that  $m \leq p$ . Putting these bounds together yields

$$\frac{1}{p}\text{tr}(A\Sigma) = \frac{1}{p}\text{tr}(A_1) + \frac{1}{p}\text{tr}(A_4) + \mathcal{O}(m^{-1}\|A\|) = \frac{1}{p}\text{tr}(A) + \mathcal{O}(m^{-1}\|A\|), \quad (220)$$

which completes the proof.  $\square$

Our main concentration bound is the following.

**Lemma 9.** *For a deterministic  $p \times p$  complex symmetric matrix  $A$  with  $\|A\| = \mathcal{O}(1)$ ,*

$$\mathbb{E} \left| \frac{1}{p}z^T Az - \frac{1}{p}\mathbb{E}[z^T Az] \right| = o(1). \quad (221)$$

*Proof.* By the Cauchy–Schwarz inequality, it suffices to show that

$$\mathbb{E} \left| \frac{1}{p}z^T Az - \frac{1}{p}\mathbb{E}[z^T Az] \right|^2 = o(1). \quad (222)$$

Identifying the left-hand side as the variance of  $\frac{1}{p}z^T Az$  and applying Lemma 8, it suffices to show that

$$\mathbb{E} \left| \frac{1}{p}z^T Az \right|^2 = \left| \frac{1}{p}\text{tr}(A) \right|^2 + o(1). \quad (223)$$

We expand the left-hand side as

$$\mathbb{E} \left| \frac{1}{p}z^T Az \right|^2 = \frac{1}{p^2} \sum_{i,j,k,l=1}^p A_{ij} \overline{A_{kl}} \mathbb{E}[z_i z_j z_k z_l], \quad (224)$$

where we have denoted  $z = (z_i)_{i=1}^p$ . Decompose the sum as

$$\frac{1}{p^2} \sum_{i,j,k,l=1}^p A_{ij} \overline{A_{kl}} \mathbb{E}[z_i z_j z_k z_l] = S_1 + S_2 + S_3 + S_4, \quad (225)$$

where

$$S_r := \frac{1}{p^2} \sum A_{ij} \overline{A_{kl}} \mathbb{E}[z_i z_j z_k z_l] \quad (226)$$

and the sum is over all  $1 \leq i, j, k, l \leq p$  such that  $|\{i, j, k, l\}| = r$ .

We will show that  $S_2$  contains the main contribution and the remaining terms  $S_r$ ,  $r \neq 2$  can be appropriately bounded. We begin by bounding  $S_3$  and  $S_4$  together. Let  $\Lambda$  be a  $p^2 \times p^2$  matrix, indexed by  $\{(i, j) : 1 \leq i, j \leq p\}$  and defined by

$$\Lambda_{(i,j),(k,l)} = \begin{cases} \mathbb{E}[z_i z_j z_k z_l] & \text{if } |\{i, j, k, l\}| \geq 3, \\ 0, & \text{otherwise.} \end{cases} \quad (227)$$

Then

$$S_3 + S_4 = \frac{1}{p^2} \sum_{i,j,k,l=1}^p A_{ij} \overline{A_{kl}} \Lambda_{(i,j),(k,l)} \quad (228)$$

can be identified as a quadratic form in the matrix  $\Lambda$ . Thus, we can bound (see for instance<sup>13</sup>)

$$|S_3 + S_4| \leq \frac{1}{p^2} \|A\|_2^2 \|\Lambda\|_2. \quad (229)$$

Trivially, by Supplemental Equation (199), we have the bound

$$\frac{1}{p} \|A\|_2^2 \leq \|A\|^2. \quad (230)$$

Additionally, by assumption (4) and the fact that  $m \geq cp$ ,

$$\frac{1}{p^2} \|\Lambda\|_2^2 = \frac{1}{p^2} \sum_{|\{i,j,k,l\}| \geq 3} |\mathbb{E}[z_i z_j z_k z_l]|^2 = o(1). \quad (231)$$

Thus, we conclude that  $|S_3 + S_4| = o(1)$ .

For  $S_1$ , we easily bound

$$|S_1| \leq \frac{1}{p^2} \sum_{i=1}^p \mathbb{E}[z_i^4] |A_{ii}|^2 \leq \frac{\kappa^4}{p} \|A\|^2 = o(1) \quad (232)$$

using assumption 1 on page 27.

It remains to control the terms from  $S_2$ . When  $i = j$  and  $k = l \neq i$ , the sum in  $S_2$  gives (using assumption 2 on page 27)

$$\frac{1}{p^2} \sum A_{ij} \overline{A_{kl}} \mathbb{E}[z_i z_j z_k z_l] = \left( \frac{1}{p} \text{tr}(A) \right)^2 + o(1), \quad (233)$$

which is the main contribution. Thus, we need to show that the other terms in  $S_2$  give negligible contribution. Indeed, we have

$$\left| S_2 - \left( \frac{1}{p} \operatorname{tr} A \right)^2 \right| \leq C \frac{1}{p^2} \|A\|_2^2 + o(1) \quad (234)$$

by assumptions 2 and 3 on page 27. Here  $C > 0$  is a constant. Using Supplemental Equation (230), we conclude that

$$\left| S_2 - \left( \frac{1}{p} \operatorname{tr} A \right)^2 \right| = o(1), \quad (235)$$

and the proof is complete.  $\square$

### 2.3.6 Proof of Theorem 1

We are now in a position to complete the proof of Theorem 1. We let  $z_1, \dots, z_n$  denote the rows of  $Z$  (which are iid copies of  $z$ ). Set  $\tau_n := \frac{n}{p}$  and recall that  $\tau_n \rightarrow \tau \in (0, \infty)$ .

By the Schur complement formula (see for instance<sup>7</sup>, it follows that

$$s_n(\alpha) = \frac{1}{n} \operatorname{tr} R(\alpha) \quad (236)$$

$$= \frac{1}{n} \sum_{k=1}^n \frac{1}{\frac{1}{p} z_k z_k^T - \alpha - \frac{1}{p^2} z_k Z^{(k)T} R^{(k)}(\alpha) Z^{(k)} z_k^T} = \frac{1}{n} \sum_{k=1}^n \frac{1}{a_k}, \quad (237)$$

where

$$a_k := \frac{1}{p} z_k z_k^T - \alpha - \frac{1}{p^2} z_k Z^{(k)T} R^{(k)}(\alpha) Z^{(k)} z_k^T. \quad (238)$$

Moreover, it follows (see for example page 472 of<sup>7</sup>) that

$$|a_k| \geq \operatorname{Im}(\alpha). \quad (239)$$

So far we have considered arbitrary  $\alpha \in \mathbb{C}^+$ , but we now restrict ourselves to

$$\alpha \in Q := \{w \in \mathbb{C}^+ : \operatorname{Re}(w) = 0, 1/2 \leq \operatorname{Im}(w) \leq 1\}. \quad (240)$$

We have, for any  $\alpha \in Q$ ,

$$\left| \mathbb{E} s_n(\alpha) - \frac{1}{n} \sum_{k=1}^n \frac{1}{\mathbb{E}[a_k]} \right| \leq \frac{1}{n} \sum_{k=1}^n \mathbb{E} \left| \frac{1}{a_k} - \frac{1}{\mathbb{E}[a_k]} \right| \quad (241)$$

$$\leq \frac{1}{n} \sum_{k=1}^n \mathbb{E} \left[ \frac{|a_k - \mathbb{E}[a_k]|}{(\operatorname{Im}(\alpha))^2} \right] \quad (242)$$

$$\leq 4\mathbb{E} [|a_1 - \mathbb{E}[a_1]|], \quad (243)$$

where we used Supplemental Equation (239) and the fact that  $\text{Im}(\alpha) \geq 1/2$  for any  $\alpha \in Q$ . In the last inequality we also used that the random variables  $a_k$  are identically distributed since the rows of  $Z$  are iid.

Letting

$$A := I - \frac{1}{p} Z^{(1)T} R^{(1)}(\alpha) Z^{(1)}, \quad (244)$$

we note that

$$a_1 - \mathbb{E}[a_1] = \frac{1}{p} z_1 A z_1^T - \frac{1}{p} \mathbb{E}[z_1 A z_1^T]. \quad (245)$$

By Supplemental Equation (192),  $\|A\| \leq 3$  for all  $\alpha \in Q$ . In addition,  $A$  is independent of  $z_1$ . Let  $\mathbb{E}_1[\cdot]$  denote the expectation with respect to only the first row  $z_1$ . Our goal is to now show that

$$\mathbb{E}[|a_1 - \mathbb{E}[a_1]|] = o(1) \quad (246)$$

uniformly for  $\alpha \in Q$ . Indeed, using the triangle inequality and Lemma 9 (by conditioning on the matrix  $A$ ), we see that

$$\mathbb{E}[|a_1 - \mathbb{E}[a_1]|] \leq \mathbb{E}[|a_1 - \mathbb{E}_1[a_1]|] + \mathbb{E}[|\mathbb{E}_1[a_1] - \mathbb{E}[a_1]|] = o(1) + \mathbb{E}[|\mathbb{E}_1[a_1] - \mathbb{E}[a_1]|] \quad (247)$$

uniformly for  $\alpha \in Q$ . To bound the second term we note that

$$\mathbb{E}[|\mathbb{E}_1[a_1] - \mathbb{E}[a_1]|] = \mathbb{E}\left[\left|\frac{1}{p} \text{tr} A - \mathbb{E}\left[\frac{1}{p} \text{tr} A\right]\right|\right] + o(1) \quad (248)$$

due to Lemma 8 and Fubini's theorem. By the cyclic property of the trace

$$|\text{tr} A - \mathbb{E}[\text{tr} A]| = |\alpha| \left| \text{tr} R^{(1)}(\alpha) - \mathbb{E}[\text{tr} R^{(1)}(\alpha)] \right| \quad (249)$$

since

$$\frac{1}{p} R^{(1)} Z^{(1)} Z^{(1)T} = R^{(1)} \left( \frac{1}{p} Z^{(1)} Z^{(1)T} - \alpha I + \alpha I \right) = I + \alpha R^{(1)}. \quad (250)$$

So by Lemmas 2 and 7, we see that

$$\mathbb{E}[|\mathbb{E}_1[a_1] - \mathbb{E}[a_1]|] \leq \tau_n \mathbb{E}[|s_n(\alpha) - \mathbb{E}[s_n(\alpha)]|] + o(1) = o(1) \quad (251)$$

uniformly for  $\alpha \in Q$ . Putting together the bounds from above, we obtain (246).

Combining Supplemental Equation (246) with Supplemental Equation (243), we conclude that

$$\left| \mathbb{E}[s_n(\alpha)] - \frac{1}{\mathbb{E}[a_1]} \right| = \left| \mathbb{E}[s_n(\alpha)] - \frac{1}{n} \sum_{k=1}^n \frac{1}{\mathbb{E}[a_k]} \right| = o(1) \quad (252)$$

uniformly for  $\alpha \in Q$ . Here we have again exploited the fact that the random variables  $a_k$  are identically distributed.

In view of Lemma 8 and Fubini's theorem,

$$\frac{1}{p}\mathbb{E}[z_1Az_1^T] = \frac{1}{p}\mathbb{E}[\text{tr} A] + o(1) \quad (253)$$

uniformly for  $\alpha \in Q$ . Using Lemma 2, the cyclic property of the trace, and Supplemental Equation (250), we obtain

$$\frac{1}{p}\mathbb{E}[z_1Az_1^T] = 1 - \frac{1}{p^2}\mathbb{E}\left[\text{tr} R^{(1)}(\alpha)Z^{(1)}Z^{(1)T}\right] + o(1) \quad (254)$$

$$= 1 - \tau_n - \frac{\alpha}{p}\mathbb{E}[\text{tr} R^{(1)}(\alpha)] + o(1) \quad (255)$$

$$= 1 - \tau_n - \frac{\alpha}{p}\mathbb{E}[\text{tr} R(\alpha)] + o(1) \quad (256)$$

$$= 1 - \tau_n - \alpha\tau_n\mathbb{E}[s_n(\alpha)] + o(1) \quad (257)$$

uniformly for  $\alpha \in Q$ . Thus, since  $\tau_n \rightarrow \tau$ ,

$$\mathbb{E}[a_1] = 1 - \alpha - \tau - \alpha\tau\mathbb{E}[s_n(\alpha)] + o(1) \quad (258)$$

uniformly for  $\alpha \in Q$ . Returning to Supplemental Equation (252), we can express the Stieltjes transform of the empirical spectral measure as

$$\mathbb{E}[s_n(\alpha)] = \frac{1}{\mathbb{E}[a_1]} + o(1) = \frac{1}{1 - \alpha - \tau - \alpha\tau\mathbb{E}[s_n(\alpha)]} + o(1) \quad (259)$$

uniformly for  $\alpha \in Q$ , where we used Supplemental Equation (239) and Lemma 3 to remove the  $o(1)$  error term from the denominator.

We are now in a position to apply Lemma 4. Indeed, Lemma 4 gives

$$\sup_{\alpha \in Q} |\mathbb{E}[s_n(\alpha)] - s(\alpha)| = o(1). \quad (260)$$

Combined with Lemma 7 and the Borel–Cantelli lemma, this implies that

$$\sup_{\alpha \in Q} |s_n(\alpha) - s(\alpha)| \rightarrow 0 \quad (261)$$

almost surely as  $n \rightarrow \infty$ . Since  $s_n$  and  $s$  are analytic in  $\mathbb{C}^+$  and satisfy  $|s_n(\alpha)| \leq \frac{1}{\text{Im}(\alpha)}$  and  $|s(\alpha)| \leq \frac{1}{\text{Im}(\alpha)}$  (by Lemma 3), Vitali's convergence theorem (see page 168 in Tao<sup>9</sup> or Lemma 2.14 in Bai and Silverstein<sup>7</sup>)

implies that almost surely

$$s_n(\alpha) \rightarrow s(\alpha) \quad (262)$$

as  $n \rightarrow \infty$  for each  $\alpha \in \mathbb{C}^+$ . In other words, we have shown that the Stieltjes transform of the ESD of  $\frac{1}{p}ZZ^T$  converges almost surely to the Stieltjes transform of  $F_\tau^{\text{MP}}$ . This implies the almost sure convergence of the ESD by Theorem B.9 in Bai and Silverstein<sup>7</sup>. The proof of Theorem 1 is complete.

## 2.4 The limiting spectral distribution of the genetic relatedness matrix

We now demonstrate that the ESD of the GRM  $m^{-1}ZZ^T$  converges almost surely to  $F_\tau^{\text{MP}}$ . All that remains is to show that the elements of  $Z$  conform to the assumptions needed for Theorem 1. We restate those assumptions here:

1. There exists a constant  $\kappa > 0$  so that  $\sup_{1 \leq i \leq m} |z_i| \leq \kappa$  with probability 1. We note that with a bit of extra work, this can be generalized to the case where  $\sup_{1 \leq i \leq m} |z_i| = o(\log(m))$ .
2.  $\mathbb{E}[z_i^2 z_j^2] = 1 + o(1)$  uniformly for all distinct  $i, j$ .
3.  $\mathbb{E}[z_i^3 z_j] = o(1)$  uniformly for all distinct  $i, j$ .
4. One has

$$\mathbb{E}[z_i z_j z_k z_l] = \begin{cases} o\left(\frac{1}{m}\right), & \text{if } |\{i, j, k, l\}| = 4, \\ o\left(\frac{1}{\sqrt{m}}\right), & \text{if } |\{i, j, k, l\}| = 3, \end{cases} \quad (263)$$

uniformly in  $i, j, k, l$ .

First, 1 is ensured by choosing a minor allele frequency cutoff  $\delta > 0$  independent of sample size, as is common practice. Because each  $\mathcal{Z} = 2^{-1/2}(\mathcal{G}_1 + \mathcal{G}_2)$  is the sum of two standardized Bernoullis, we then have

$$\sup_{1 \leq i \leq m} |\mathcal{Z}_i| \leq \sqrt{2} \frac{1 - \delta}{\sqrt{\delta(1 - \delta)}}. \quad (264)$$

Thus, setting  $\delta = 2(2 + \kappa^2)^{-1}$  ensures  $\sup_{1 \leq i \leq m} |\mathcal{Z}_i| \leq \kappa$ .

### 2.4.1 From diploid genotypes to haploid genotypes

In what follows we establish that Theorem 1 can be applied directly to the GRM under assortative mating.

For a given collection of diploid sites indexed  $\alpha_1, \dots, \alpha_N$ , denote

$$\mathcal{S}_{\alpha_1 | \dots | \alpha_N} := \sup_{\{\iota_1, \dots, \iota_N\}} \left| \mathbb{E} \left[ \prod_{i=1}^N \mathcal{G}_{\iota_i}^{\alpha_i} \right] \right|. \quad (265)$$

Then  $|\mathbb{E}[\mathcal{Z}_i^2 \mathcal{Z}_{j \neq i}^2]|$  is a homogeneous fourth degree polynomial bounded by

$$|\mathbb{E}[\mathcal{Z}_i^2 \mathcal{Z}_{j \neq i}^2]| \leq \mathcal{S}_{2|2} + 2\mathcal{S}_{1|1|2} + \mathcal{S}_{1|1|1|1}, \quad (266)$$

Likewise, we have

$$|\mathbb{E}[\mathcal{Z}_i^3 \mathcal{Z}_{j \neq i}]| \leq \mathcal{S}_{1|3} + 3\mathcal{S}_{1|1|2}, \quad (267)$$

$$|\mathbb{E}[\mathcal{Z}_i^2 \mathcal{Z}_{j \neq i} \mathcal{Z}_{k \neq i,j}]| \leq 2\mathcal{S}_{1|1|2} + 2\mathcal{S}_{1|1|1|1}, \quad (268)$$

$$|\mathbb{E}[\mathcal{Z}_i \mathcal{Z}_{j \neq i} \mathcal{Z}_{k \neq i,j} \mathcal{Z}_{l \neq i,j,k}]| \leq 4\mathcal{S}_{1|1|1|1}. \quad (269)$$

Thus, we need simply need to establish equivalent bounds on the moments of haploid genotypes:

1.  $\mathbb{E}[\mathcal{G}_i \mathcal{G}_{j \neq i} \mathcal{G}_{k \neq i,j} \mathcal{G}_{l \neq i,j,k}] = o\left(\frac{1}{m}\right),$
2.  $\mathbb{E}[\mathcal{G}_i^2 \mathcal{G}_{j \neq i} \mathcal{G}_{k \neq i,j}] = o\left(\frac{1}{\sqrt{m}}\right),$
3.  $\mathbb{E}[\mathcal{G}_i^3 \mathcal{G}_{j \neq i}] = o(1),$
4.  $\mathbb{E}[\mathcal{G}_i^2 \mathcal{G}_{j \neq i}^2] = 1 + o(1),$

all which we demonstrated in Section 2.1.

### 3 Impact of assortative mating on marker-based heritability estimators

#### 3.1 Haseman-Elston regression

Denote the lower triangular components of the phenotypic and genotypic sample covariance matrices as

$$\mathcal{Y} = \text{vec}(\{\{\sigma_y^{-2} y y^T\}_{i,j:i < j}\}), \quad \mathcal{S} = \text{vec}(\{\{p^{-1} \mathcal{Z} \mathcal{Z}^T\}_{i,j:i < j}\}). \quad (270)$$

The HE regression heritability estimator is obtained by regressing  $\mathcal{Y}$  on to  $\mathcal{S}$ :

$$\hat{h}_{\text{HE}}^2 = \frac{\widehat{\text{Cov}}(\mathcal{Y}, \mathcal{S})}{\widehat{\text{Var}}(\mathcal{S})}. \quad (271)$$

Under exchangeable, loci, the same is true for each of the haploid standardized genotypes at casual variants, which we denote  $\{\mathcal{G}_k\}_{k=1}^{2m}$ , where  $Z_k = 2^{-1/2}(\mathcal{G}_{2k} + \mathcal{G}_{2k+1})$ ,  $k = 1, \dots, m$ . We assume that all genotypes at non-causal variants are independent.



For each allelic effect  $u_k$  define  $\phi_k = \sqrt{r}\nu_{k,\infty}$  such that each correlation between haploid variants is  $\mu_{kl,\infty} = \phi_k\phi_l$ . The variance/covariance matrix of the standardized diploid genetic effects is thus

$$\mathbb{E}(m^{-1}\mathcal{Z}^T\mathcal{Z}) = \text{diag}\{1 - \phi_k^2\}_{k=1}^m + 2\phi\phi^T \quad (272)$$

$$:= \Upsilon_\infty. \quad (273)$$

Now consider the element  $\mathcal{S}_{ij}$  of  $\mathcal{S}$  corresponding to the average similarity at  $m$  diploid loci among the  $i^{\text{th}}$  and  $j^{\text{th}}$  unrelated individuals:

$$\mathcal{S}_{ij} = \frac{1}{m} \sum_{k=1}^m \mathcal{Z}_{i,k} \mathcal{Z}_{j,k}. \quad (274)$$

Because these individuals are unrelated,  $\mathcal{Z}_{i,k}$  and  $\mathcal{Z}_{j,l}$  are independent for all  $k, l$  and thus  $\mathbb{E}[\mathcal{S}_{ij}] = 0$ . The variance is then simply computed as

$$\text{Var}(\mathcal{S}_{ij}) = m^{-2} \sum_{k=1}^m \sum_{l=1}^m \mathbb{E}[\mathcal{Z}_{i,k} \mathcal{Z}_{j,k} \mathcal{Z}_{i,l} \mathcal{Z}_{j,l}] \quad (275)$$

$$= m^{-2} \text{tr}[\Upsilon_\infty^2]. \quad (276)$$

Turning our attention to  $\mathcal{Y}_{ij}$  and marginalizing over independent error terms, we see that

$$\hat{\sigma}_{y,\infty}^2 \mathcal{Y}_{ij} = \sum_{k,l=1}^m \mathcal{Z}_{i,k} \mathcal{Z}_{j,l} u_k u_l, \quad (277)$$

has zero expectation. The covariance between  $\mathcal{Y}_{ij}$  and  $\mathcal{S}_{ij}$  is then

$$\text{Cov}(\mathcal{Y}_{ij}, \mathcal{S}_{ij}) = \hat{\sigma}_{y,\infty}^{-2} m^{-1} u^T \Upsilon_\infty^2 u. \quad (278)$$

Thus, the expected value of the HE estimator is computed

$$\mathbb{E}[\hat{h}_{\text{HE}}^2] = \frac{\text{Cov}(\mathcal{Y}, \mathcal{S})}{\text{Var}(\mathcal{S})} \quad (279)$$

$$= \frac{m^{-1} u^T \Upsilon_\infty^2 u}{\sigma_{y,\infty}^2 m^{-2} \text{tr}[\Upsilon_\infty^2]} \left( \frac{\sigma_{g,\infty}^2}{\sigma_{g,\infty}^2} \right) \quad (280)$$

$$= \frac{m \cdot u^T \Upsilon_\infty^2 u}{\sigma_{g,\infty}^2 \cdot \text{tr}[\Upsilon_\infty^2]} h_\infty^2. \quad (281)$$

Note that when  $\Upsilon_\infty = I$ , as is expected under random mating, we simply have  $\mathbb{E}[\hat{h}_{\text{HE}}^2] = h_\infty^2 = h_\emptyset^2$ ; i.e. HE regression is unbiased under random mating. We can further simplify the above if we assume exchangeable loci (i.e., for all  $k, l$ ,  $u_k u_l = m^{-1} \sigma_{g,0}^2$ ). In this case, define  $\phi \in (-1, 1)^m$  by  $\phi_k \equiv r^{1/2} \nu_\infty := \phi$  for all  $k$ . We

can then write  $\Upsilon_\infty = D + 2\phi\phi^T$ , where  $D = \text{diag}\{1 - \phi_k^2\}_{k=1}^m$ . We compute  $u^T \Upsilon_\infty^2 u$  by first noting that

$$u^T \Upsilon_\infty^2 u = u^T (D^2 + 2\phi\phi^T D + 2D\phi\phi^T + 4\phi\phi^T \phi\phi^T) u \quad (282)$$

$$= u^T D^2 u + 4u^T \phi\phi^T D u + 4\phi^T \phi u^T \phi\phi^T u. \quad (283)$$

Computing the above terms individually yields

$$u^T D^2 u = \sigma_{g,0}^2 (1 - \phi^2)^2, \quad (284)$$

$$u^T \phi\phi^T D u = m\sigma_{g,0}^2 \phi^2 (1 - \phi^2), \quad (285)$$

$$\phi^T \phi u^T \phi\phi^T u = m^2 \sigma_{g,0}^2 \phi^4, \quad (286)$$

$$\implies \text{Cov}(\mathcal{Y}_{ij}, \mathcal{S}_{ij}) = \hat{\sigma}_{y,\infty}^{-2} m^{-1} \sigma_{g,0}^2 (1 - \phi^2 + 2m\phi^2)^2, \quad (287)$$

Likewise, we compute the denominator

$$\text{Var}(\mathcal{S}_{ij}) = m^{-1} (1 + 2\phi^2 + 4m\phi^4). \quad (288)$$

Thus, under exchangeable loci,

$$\mathbb{E}[\hat{h}_{\text{HE}}^2] = \frac{\sigma_{g,0}^2 (1 - \phi^2 + 2m\phi^2)^2}{\sigma_{g,\infty}^2 (1 + 2\phi^2 + 4m\phi^4)} h_\infty^2 \quad (289)$$

where  $\phi^2 = m\sigma_{y,\infty}^2 (8r\sigma_{g,0}^2)^{-1} \left( \sqrt{4r\sigma_{g,0}^2 m^{-1} \sigma_{y,\infty}^{-2} + (1 - r_{g,\infty})^2} - (1 - r_{g,\infty}) \right)^2$ . Substituting in panmictic parameters and taking the limit as  $m \rightarrow \infty$  yields

$$\mathbb{E}[\hat{h}_{\text{HE}}^2] \approx \frac{\sigma_{g,0}^2}{\sigma_{g,\infty}^2} h_\infty^2 = \frac{h_\infty^2}{1 - r_{g,\infty}}. \quad (290)$$

This gives us a way to estimate the panmictic and equilibrium heritabilities as a function of the HE estimator when  $r$  is known:

$$\hat{h}_0^2 := \mathbb{E}[h_0^2 | \hat{h}_{\text{HE}}^2] = \frac{\hat{h}_{\text{HE}}^2}{1 + 2r\hat{h}_{\text{HE}}^2 + r(r-1)\hat{h}_{\text{HE}}^4}, \quad (291)$$

$$\hat{h}_\infty^2 := \mathbb{E}[h_\infty^2 | \hat{h}_{\text{HE}}^2] = \frac{\hat{h}_{\text{HE}}^2}{1 + r\hat{h}_{\text{HE}}^2}. \quad (292)$$

We can then employ the delta method to approximate the standard errors:

$$\text{se}[\hat{h}_0^2] \approx \text{se}[\hat{h}_{\text{HE}}^2] \frac{1 + r\hat{h}_{\text{HE}}^4 - r^2\hat{h}_{\text{HE}}^4}{\left(1 - r(\hat{h}_{\text{HE}}^2 - 2)\hat{h}_{\text{HE}}^2 + r^2\hat{h}_{\text{HE}}^4\right)^2}, \quad (293)$$

$$\text{se} \left[ \hat{h}_{\infty}^2 \right] \approx \frac{\text{se} \left[ \hat{h}_{\text{HE}}^2 \right]}{1 + r \hat{h}_{\text{HE}}^2}. \quad (294)$$

## 3.2 Residual maximum likelihood

### 3.2.1 Notation and problem statement

Consider the model

$$\begin{aligned} y &= X\beta + \tilde{Z}u + e, \\ u &\stackrel{i.i.d.}{\sim} \mathcal{N}(0, \sigma_g^2), \quad e \stackrel{i.i.d.}{\sim} \mathcal{N}(0, \sigma_e^2), \end{aligned} \quad (295)$$

or, equivalently,

$$\begin{aligned} (y|Z) &\sim \mathcal{MVN}(X\beta, \sigma_e^2 V_{\hat{\gamma}}), \\ V_{\hat{\gamma}} &= \hat{\gamma} \tilde{Z} \tilde{Z}^T + I_n, \quad \hat{\gamma} = \sigma_g^2 / \sigma_e^2. \end{aligned} \quad (296)$$

Above,  $Z \in \mathbb{R}^{n \times p}$  and  $u \in \mathbb{R}^p$  are independent of  $e \in \mathbb{R}^n$  and  $X \in \mathbb{R}^{n \times c}$ ,  $\beta \in \mathbb{R}^c$  are deterministic. The variance components  $\sigma_g^2$ ,  $\sigma_e^2$ , and thus their ratio  $\hat{\gamma}$ , are strictly positive.  $\tilde{Z} \tilde{Z}^T = p^{-1} Z Z^T$ , the matrix we've been concerned with, is sub-Gaussian with independent rows and, by our results from Section 2, is such that its empirical spectral density  $F^{\tilde{Z} \tilde{Z}^T} \xrightarrow{\text{a.s.}} F_{\tau}^{\text{MP}}$ , where  $\tau = n/p$ . For simplicity, we assume  $m = p$  (all variants are causal) and  $\text{Cov}(Z_{\cdot k}, Z_{\cdot l}) \equiv \mu = \mathcal{O}(m^{-1})$  for all  $k \neq l$  (exchangeable loci), though our results will hold as long as  $m/p \rightarrow c \in (0, 1]$  and  $\sup_{k \neq l} \text{Cov}(Z_{\cdot k}, Z_{\cdot l}) = \mathcal{O}(m^{-1})$ . In what follows,  $\hat{\gamma}$  will denote true parameter value,  $\hat{\gamma}$  the REML estimator, and  $\gamma > 0$  an arbitrary value.

Let  $A^T : \mathbb{R}^n \rightarrow (\text{col } X)^{\perp}$  such that  $A^T X = 0$ ,  $A^T A = I_{n-c}$  and define the following quantities for any value of  $\gamma > 0$ :

- $\tilde{y} = y - X\beta$ , the mean-deviated outcome vector
- $\Sigma_{\gamma} = A^T V_{\gamma} A = I_{n-c} + \gamma A^T \tilde{Z} \tilde{Z}^T A$ , the transformed marginal covariance
- $\zeta = A^T \tilde{Z}$ , the transformed standardized genotypes
- $P_{\gamma} = A \Sigma_{\gamma}^{-1} A^T$ , the Schur complement of  $X^T V_{\gamma}^{-1} X$  in  $\begin{pmatrix} V_{\gamma}^{-1} & X^T V_{\gamma}^{-1} \\ V_{\gamma}^{-1} X & X^T V_{\gamma}^{-1} X \end{pmatrix}$
- $\Delta(\gamma) = \sigma_e^{-2} y^T \left( \frac{P_{\gamma} \tilde{Z} \tilde{Z}^T P_{\gamma}}{\text{tr}[P_{\gamma} \tilde{Z} \tilde{Z}^T]} - \frac{P_{\gamma}^2}{\text{tr}[P_{\gamma}]} \right) y = \sigma_e^{-2} \tilde{y}^T \left( \frac{P_{\gamma} \tilde{Z} \tilde{Z}^T P_{\gamma}}{\text{tr}[P_{\gamma} \tilde{Z} \tilde{Z}^T]} - \frac{P_{\gamma}^2}{\text{tr}[P_{\gamma}]} \right) \tilde{y}$ , the ‘‘stationarity function’’.

The REML estimator  $\hat{\gamma}$  satisfies

$$\sigma_e^{-2} \frac{y^T P_{\hat{\gamma}} \tilde{Z} \tilde{Z}^T P_{\hat{\gamma}} y}{\text{tr}[P_{\hat{\gamma}} \tilde{Z} \tilde{Z}^T]} = \sigma_e^{-2} \frac{y^T P_{\hat{\gamma}}^2 y}{\text{tr}[P_{\hat{\gamma}}]}, \quad (297)$$

or, equivalently, it is the root of the stationarity function  $\Delta(\hat{\gamma}) = 0$ . Choosing  $\hat{\gamma}$  such that  $\Delta(\hat{\gamma}) = 0$  and setting  $\hat{\sigma}_e^2 = \frac{y^T P_{\hat{\gamma}}^2 y}{\text{tr}[P_{\hat{\gamma}}]}$ ,  $\hat{\sigma}_g^2 = \hat{\sigma}_e^2 \hat{\gamma}$  yields the variance component estimators that maximize the likelihood of

$$(A^T y | Z) \sim \mathcal{MVN}(0, \sigma_e^2 \Sigma_{\hat{\gamma}}). \quad (298)$$

Our goal is to demonstrate that  $\hat{\gamma} \xrightarrow{P} \hat{\gamma}$  as  $n, p \rightarrow \infty$ ,  $n/p \rightarrow \tau$ . For simplicity, we address the exchangeable loci case where each  $u_k \equiv \pm m^{-1/2} \sigma_{g,0}$  for  $k = 1, \dots, m$ , though a similar analysis is possible under a random SNP effect model given the moments of  $u|Z$ .

### 3.2.2 The limiting spectral distribution of $\zeta \zeta^T = A^T \tilde{Z} \tilde{Z}^T A$

From your previous results, we know that the ESD of the  $n \times p$  matrix  $\tilde{Z} = p^{1/2} Z$  is such that  $F^{\tilde{Z} \tilde{Z}^T} \xrightarrow{\text{a.s.}} F_{\tau}^{\text{MP}}$ . Here we show that the ESD of the  $(n-c) \times p$  matrix  $\zeta = A^T \tilde{Z}$ , is such that  $F^{\zeta \zeta^T} \xrightarrow{\text{a.s.}} F_{\tau}^{\text{MP}}$  as well. We first establish some intermediate results.

**Lemma 10.** *Suppose  $Q$  is a Hermitian matrix. If each eigenvalue of  $Q$  is either 0 or 1, then  $Q$  is an orthogonal projection.*

*Proof.* By supposition  $Q$  is Hermitian, so it suffices to show that  $Q^2 = Q$ . By the spectral theorem, we can write

$$Q = U \Lambda U^*, \quad (299)$$

where  $U$  is a unitary matrix and  $\Lambda$  is a diagonal matrix whose diagonal entries are the eigenvalues of  $Q$ . This implies that  $\Lambda^2 = \Lambda$  since the eigenvalues of  $Q$  can only be 0 or 1 by assumption. It follows that

$$Q^2 = U \Lambda U^* U \Lambda U^* = U \Lambda^2 U^* = Q, \quad (300)$$

and the proof is complete.  $\square$

**Lemma 11.** *Suppose  $Q$  is an  $n \times n$  orthogonal projection matrix. If zero is an eigenvalue of  $Q$  with multiplicity  $m$ , then  $\text{rank}(I_n - Q) = m$ .*

*Proof.* By the spectral theorem, we can decompose

$$Q = U \Lambda U^*, \quad (301)$$

where  $U$  is an  $n \times n$  unitary matrix and  $\Lambda$  is an  $n \times n$  diagonal matrix whose diagonal entries are the eigenvalues of  $P$  (and so must be either 0 or 1). Hence,

$$I_n - Q = U(I_n - \Lambda)U^*. \quad (302)$$

Since  $U$  and  $U^*$  have full rank, it follows that

$$\text{rank}(I_n - Q) = \text{rank}(I_n - \Lambda). \quad (303)$$

Since  $I_n - \Lambda$  is a diagonal Hermitian matrix, it follows that the rank of  $I_n - \Lambda$  is just the number of nonzero diagonal entries of  $I_n - \Lambda$ . In other words, the rank of  $I_n - \Lambda$  is simply the number of diagonal entries of  $\Lambda$  that are zero.  $\square$

**Lemma 12.** *Let  $A$  be an  $n \times m$  matrix and  $B$  an  $m \times n$  matrix. Then the non-trivial eigenvalues of  $AB$  are the same as the non-trivial eigenvalues of  $BA$  (counting algebraic multiplicity).*

*Proof.* The lemma will follow as an application of Sylvester's determinant identity (also called the Weinstein-Aronszajn identity): if  $C$  and  $D$  are matrices of size  $n \times m$  and  $m \times n$  respectively, then

$$\det(I_n + CD) = \det(I_m + DC). \quad (304)$$

This identity can be found on pg. 271 of Pozrikidis<sup>14</sup>.

We now use Sylvester's determinant identity to complete the proof of the lemma. Assume without loss of generality that  $n \geq m$ . Then, for  $z \neq 0$ ,

$$\det(zI_n - AB) = z^n \det(I_n - z^{-1}AB) = z^n \det(I_m - z^{-1}BA) = z^{n-m} \det(zI_m - BA). \quad (305)$$

We conclude that  $AB$  and  $BA$  have the same characteristic polynomials up to a factor of  $z^{n-m}$ . The factor  $z^{n-m}$  corresponds to the trivial eigenvalues at zero.  $\square$

We now return to the problem of determining the limiting spectral distribution of  $F\zeta\zeta^T$ .

**Lemma 13.** *If  $A$  is an  $n \times (n - c)$  matrix such that  $A^T A = I_{n-c}$ , then*

$$\|F^{p-1}A^T Z Z^T A - F^{p-1}Z Z^T\| \leq 3\frac{c}{n}. \quad (306)$$

*Proof.* The eigenvalues of  $A^T A$  are all one. By Lemma 12, we conclude that  $AA^T$  has  $n - c$  eigenvalues that are one and  $c$  eigenvalues that are zero. It follows from Lemma 10 that  $P := AA^T$  is an orthogonal projection matrix (since it is clearly Hermitian).

We now consider the eigenvalues of  $\frac{1}{p}A^T Z Z^T A$ . By Lemma 12, the eigenvalues of  $\frac{1}{p}A^T Z Z^T A$  are the same as the eigenvalues of  $\frac{1}{p}AA^T Z Z^T = \frac{1}{p}Q Z Z^T$ , except the latter matrix has  $c$  additional zero eigenvalues. Hence, it follows that

$$\|F^{p-1}A^T Z Z^T A - F^{p-1}P Z Z^T\| \leq 2\frac{c}{n}. \quad (307)$$

Since  $Q = Q^2 = Q^T Q$ , by another application of Lemma 12, the eigenvalues of  $\frac{1}{p} Q Z Z^T$  are the same as the eigenvalues of  $\frac{1}{p} Q Z Z^T Q^T$ , counting multiplicity (this latter matrix is nicer to work with because it is Hermitian). In other words,

$$\|F^{p-1} Q Z Z^T - F^{p-1} Q Z Z^T Q^T\| = 0. \quad (308)$$

To complete the proof, we now apply Theorem A.44 from <sup>7</sup>, which yields

$$\|F^{p-1} Q Z Z^T Q^T - F^{p-1} Z Z^T\| \leq \frac{1}{n} \text{rank}(Q Z - Z). \quad (309)$$

By properties of the rank,

$$\text{rank}(Q Z - Z) = \text{rank}((Q - I_n)Z) \leq \text{rank}(Q - I_n). \quad (310)$$

Finally, by Lemma 11, we see that  $\text{rank}(Q - I_n) = c$ . Therefore, putting together the previous bounds (and applying the triangle inequality), we conclude that

$$\|F^{p-1} A^T Z Z^T A - F^{p-1} Z Z^T\| \leq 3 \frac{c}{n}. \quad (311)$$

□

Finally, combining Theorem 1 with the above lemma immediately yields  $F^{\zeta \zeta^T} \xrightarrow{\text{a.s.}} F_\tau^{\text{MP}}$ .

### 3.2.3 Spectral functions of $\zeta \zeta^T$

Here we introduce some notation that will simplify future computations. For fixed  $\gamma, \tau > 0$ , and for non-negative integers  $k, l$  such that  $k \leq l$ , denote the continuous bounded functions

$$\begin{aligned} \psi_{(k,l)} &: [0, \infty) \rightarrow [0, \infty), \\ \psi_{(k,l)} &: x \mapsto \frac{x^k}{(\gamma x + 1)^l}. \end{aligned} \quad (312)$$

Likewise, define their integrals with respect to the MP-law  $F_\tau^{\text{MP}}$ ,  $\tau > 0$  by

$$\Psi_{(j,l)} = \int \psi_{(k,l)} dF_\tau^{\text{MP}} \quad (313)$$

$$= \frac{1}{2\pi\tau} \int_{a_\tau}^{b_\tau} \psi_{(k,l)}(x) x^{-1} \sqrt{(x - a_\tau)(b_\tau - x)} dx, \quad (314)$$

where  $a_\tau = (1 - \sqrt{\tau})^2$ ,  $b_\tau = (1 + \sqrt{\tau})^2$ . Note that  $0 < \Psi_{(k,l)} < \infty$ .

### 3.2.4 Conditional expectation of the stationarity function

Recall that  $y = \tilde{Z}u + e$ , where  $u$  is deterministic with  $u^T u = \sigma_{g,0}^2$ ,  $\max_k |u_k| = \alpha m^{-1}$  for some  $\alpha$  independent of  $m$  and  $e \stackrel{i.i.d.}{\sim} \mathcal{N}(0, \sigma_e^2)$ , and, additionally, that  $\zeta = A^T \tilde{Z}$ ,  $P_\gamma = A \Sigma_\gamma^{-1} A^T$ ,  $A^T A = I$ , and  $\hat{\gamma} = \sigma_{g,0}^2 / \sigma_e^2$ . We consider the conditional expectation of  $\Delta(\gamma)$  given  $Z$ :

$$\mathbb{E}[\Delta(\gamma)|Z] = \sigma_e^{-2} u^T \left( \frac{\Sigma_\gamma^{-1} \zeta \zeta^T \Sigma_\gamma^{-1} \zeta \zeta^T}{\text{tr}[\Sigma_\gamma^{-1} \zeta \zeta^T]} - \frac{\Sigma_\gamma^{-2} \zeta \zeta^T}{\text{tr}[\Sigma_\gamma^{-1}]} \right) u + \left( \frac{\text{tr}[\Sigma_\gamma^{-1} \zeta \zeta^T \Sigma_\gamma^{-1}]}{\text{tr}[\Sigma_\gamma^{-1} \zeta \zeta^T]} - \frac{\text{tr}[\Sigma_\gamma^{-2}]}{\text{tr}[\Sigma_\gamma^{-1}]} \right) \quad (315)$$

$$= \sigma_e^{-2} \text{tr} \left[ \left( \frac{\Sigma_\gamma^{-1} \zeta \zeta^T \Sigma_\gamma^{-1} \zeta \zeta^T}{\text{tr}[\Sigma_\gamma^{-1} \zeta \zeta^T]} - \frac{\Sigma_\gamma^{-2} \zeta \zeta^T}{\text{tr}[\Sigma_\gamma^{-1}]} \right) u u^T \right] + \left( \frac{\text{tr}[\Sigma_\gamma^{-2} \zeta \zeta^T]}{\text{tr}[\Sigma_\gamma^{-1} \zeta \zeta^T]} - \frac{\text{tr}[\Sigma_\gamma^{-2}]}{\text{tr}[\Sigma_\gamma^{-1}]} \right). \quad (316)$$

Write the spectral decomposition  $\zeta \zeta^T = Q \Lambda Q^T$ , such that  $\Sigma_\gamma = Q(I + \gamma \Lambda) Q^T$ , and recall that  $u^T u = \sigma_{g,0}^2$ . Under our exchangeable loci assumption, we have

$$\mathbb{E}[\Delta(\gamma)|Z] = \hat{\gamma} \left( \frac{\text{tr}[(I + \gamma \Lambda)^{-1} \Lambda (I + \gamma \Lambda)^{-1} \Lambda]}{\text{tr}[\Sigma_\gamma^{-1} \zeta \zeta^T]} - \frac{\text{tr}[(I + \gamma \Lambda)^{-2} \Lambda]}{\text{tr}[\Sigma_\gamma^{-1}]} \right) + \left( \frac{\text{tr}[\Sigma_\gamma^{-2} \zeta \zeta^T]}{\text{tr}[\Sigma_\gamma^{-1} \zeta \zeta^T]} - \frac{\text{tr}[\Sigma_\gamma^{-2}]}{\text{tr}[\Sigma_\gamma^{-1}]} \right). \quad (317)$$

Thus, we can write  $\Delta(\gamma)$  exclusively in terms of the true parameter value  $\hat{\gamma}$  and spectral functions with known limits:

$$\mathbb{E}[\Delta(\gamma)|Z] \xrightarrow{P} \hat{\gamma} \left( \frac{\Psi_{(2,2)}}{\Psi_{(1,1)}} - \frac{\Psi_{(1,2)}}{\Psi_{(0,1)}} \right) + \left( \frac{\Psi_{(1,2)}}{\Psi_{(1,1)}} - \frac{\Psi_{(0,2)}}{\Psi_{(0,1)}} \right) := h(\gamma). \quad (318)$$

Let  $w = \hat{\gamma}/\gamma$ . Using a computer algebra system to compute the above integrals and defining  $\chi_\gamma = \sqrt{\gamma^2(\tau-1)^2 + 2\gamma(\tau+1) + 1}$ , the above quantity can be written as

$$\begin{aligned} & w\gamma \left( \frac{\Psi_{(2,2)}}{\Psi_{(1,1)}} - \frac{\Psi_{(1,2)}}{\Psi_{(0,1)}} \right) + \left( \frac{\Psi_{(1,2)}}{\Psi_{(1,1)}} - \frac{\Psi_{(0,2)}}{\Psi_{(0,1)}} \right) \\ &= \left( \frac{4\gamma(-(\sqrt{\tau+1})|1-\sqrt{\tau}|_{\chi_\gamma+\gamma(\tau-1)^2+\tau+1})}{\gamma(\sqrt{\tau+1})|1-\sqrt{\tau}|_{-\chi_\gamma+1}} \right)^{+4} - \frac{4w(-\chi_\gamma+\gamma\tau+\gamma+1)^2}{(-\chi_\gamma+\gamma\tau+\gamma+1)(-\gamma(\sqrt{\tau+1})|1-\sqrt{\tau}|_{\chi_\gamma-1})} \\ & \quad + \frac{w}{-\chi_\gamma+\gamma\tau+\gamma+1} (2\gamma(\sqrt{\tau-1})^2(\chi_\gamma-2\gamma(\sqrt{\tau+1})^2-3)+2\gamma(\sqrt{\tau+1})^2(\chi_\gamma-3)+8(\chi_\gamma-1)) (4\chi_\gamma)^{-1} \end{aligned} \quad (319)$$

$$:= \frac{C_1 + wC_2}{4\chi_\gamma}, \quad \text{where } C_2/C_1 = -1, \quad (320)$$

where  $h$  is strictly monotone in  $\gamma$  and  $h(\gamma) = 0 \iff \hat{\gamma}/\gamma = 1$ .

Under the more general case, dropping the assumption of exchangeable loci, we can use the following lemma to demonstrate that  $\mathbb{E}[\Delta(\gamma)|Z] \xrightarrow{P} \tilde{h}(\gamma)$  with  $|\tilde{h}(\gamma)| \leq |h(\gamma)|$ :

**Lemma 14.** *Let  $A = A^T \in \mathbb{R}^{m \times m}$ ,  $B = B^T \succeq 0$ . Then*

$$0 \leq \text{tr}[AB] \leq \text{tr}[B] \lambda_{\max}[A]. \quad (321)$$

*Proof.* This is an immediate corollary of Theorem 1 in<sup>15</sup>. □

Denote the rank one matrix  $A = uu^T$  and observe that

$$\sigma_e^{-2}Au = \sigma_e^{-2}u^Tuu = \sigma_e^{-2}\sigma_{g,0}^2u = \dot{\gamma}u. \quad (322)$$

Thus  $\sigma_{g,0}^2$  is the only non-zero eigenvalue of  $A$  and  $\sigma_e^{-2}\text{tr}[AB] \leq \dot{\gamma}\text{tr}[B]$ . All of the matrices in Supplementary Equation (13) (besides the outer product  $uu^T$ ) are symmetric positive definite, and we have

$$\mathbb{E}[\Delta(\gamma)|Z] \leq \dot{\gamma} \left( \frac{\text{tr} \Sigma_\gamma^{-1} \zeta \zeta^T \Sigma_\gamma^{-1} \zeta \zeta^T}{\text{tr} [\Sigma_\gamma^{-1} \zeta \zeta^T]} - \frac{\text{tr} \Sigma_\gamma^{-2} \zeta \zeta^T}{\text{tr} [\Sigma_\gamma^{-1}]} \right) + \left( \frac{\text{tr} [\Sigma_\gamma^{-2} \zeta \zeta^T]}{\text{tr} [\Sigma_\gamma^{-1} \zeta \zeta^T]} - \frac{\text{tr} [\Sigma_\gamma^{-2}]}{\text{tr} [\Sigma_\gamma^{-1}]} \right). \quad (323)$$

In what follows, we will assume exchangeable loci, though the above inequality can be used to demonstrate consistency under any set of effects  $u$  for which  $\frac{d}{d\gamma}\mathbb{E}[\Delta(\gamma)|Z]$  can be bounded away from zero in some neighborhood of  $\dot{\gamma}$ .

### 3.2.5 Asymptotic variance of the stationarity function

Split the stationarity functions into two components:

$$\Delta(\gamma) = \underbrace{\frac{\sigma_e^{-2}y^T P_\gamma \tilde{Z} \tilde{Z}^T P_\gamma y}{\text{tr} [P_\gamma \tilde{Z} \tilde{Z}^T]}}_{:=\text{LHS}(\gamma)} - \underbrace{\frac{\sigma_e^{-2}y^T P_\gamma^2 y}{\text{tr} [P_\gamma]}}_{:=\text{RHS}(\gamma)}. \quad (324)$$

Consider the conditional variance of the stationarity function  $\Delta(\gamma)$  given  $Z$ :

$$\begin{aligned} \text{Var}(\Delta(\gamma)|Z) &= \mathbb{E}[\Delta(\gamma)^2|Z] - \mathbb{E}[\Delta(\gamma)|Z]^2 \\ &= (\mathbb{E}[\text{LHS}(\gamma)^2|Z] + \mathbb{E}[\text{RHS}(\gamma)^2|Z] - 2\mathbb{E}[\text{LHS}(\gamma)\text{RHS}(\gamma)|Z]) - \mathbb{E}[\Delta(\gamma)|Z]^2. \end{aligned} \quad (325)$$

To compute the above, we make use of the following lemma regarding the products of quadratic forms:

**Lemma 15.** *Let  $x \sim \mathcal{MVN}(0, I\sigma^2)$  and let  $A, B$  denote symmetric real matrices conformable with  $x$ . Then*

$$\mathbb{E}[x^T Ax \cdot x^T Bx] = \text{tr}[A]\text{tr}[B] + 2\text{tr}[AB]. \quad (326)$$

*Proof.* This is an immediate corollary of Lemma 2 in Bao and Ullah<sup>16</sup>. □

Now, computing the above terms comprising  $\text{Var}(\Delta(\gamma)|Z)$  individually, we have

$$\sigma_e^{-2}u^T \left( \frac{\Sigma_\gamma^{-1} \zeta \zeta^T \Sigma_\gamma^{-1} \zeta \zeta^T}{\text{tr} [\Sigma_\gamma^{-1} \zeta \zeta^T]} - \frac{\Sigma_\gamma^{-2} \zeta \zeta^T}{\text{tr} [\Sigma_\gamma^{-1}]} \right) u + \left( \frac{\text{tr} [\Sigma_\gamma^{-1} \zeta \zeta^T \Sigma_\gamma^{-1}]}{\text{tr} [\Sigma_\gamma^{-1} \zeta \zeta^T]} - \frac{\text{tr} [\Sigma_\gamma^{-2}]}{\text{tr} [\Sigma_\gamma^{-1}]} \right) \quad (327)$$



First noting that

$$\mathbb{E}[\text{LHS}(\gamma)|Z]^2 = \sigma_e^{-4} \left( \frac{u^T \Sigma_\gamma^{-1} \zeta \zeta^T \Sigma_\gamma^{-1} \zeta \zeta^T u + \text{tr} [\Sigma_\gamma^{-1} \zeta \zeta^T \Sigma_\gamma^{-1}]}{\text{tr} [\Sigma_\gamma^{-1} \zeta \zeta^T]} \right)^2 \quad (328)$$

$$\begin{aligned} &= \frac{\text{tr} [\zeta^T \Sigma_\gamma^{-1} \zeta \zeta^T \Sigma_\gamma^{-1} \zeta u u^T]^2 + \text{tr} [\Sigma_\gamma^{-1} \zeta \zeta^T \Sigma_\gamma^{-1}]^2}{\sigma_e^4 \text{tr} [\Sigma_\gamma^{-1} \zeta \zeta^T]} \\ &+ \frac{2 \text{tr} [\zeta^T \Sigma_\gamma^{-1} \zeta \zeta^T \Sigma_\gamma^{-1} \zeta u u^T] \text{tr} [\Sigma_\gamma^{-1} \zeta \zeta^T \Sigma_\gamma^{-1}]}{\sigma_e^4 \text{tr} [\Sigma_\gamma^{-1} \zeta \zeta^T]}, \end{aligned} \quad (329)$$

we see that

$$\mathbb{E}[\text{LHS}(\gamma)^2|Z] = \sigma_e^{-4} \mathbb{E} \left[ \left( \frac{u^T \zeta^T \Sigma_\gamma^{-1} \zeta \zeta^T \Sigma_\gamma^{-1} \zeta u + e^T P_\gamma \tilde{Z} \tilde{Z}^T P_\gamma e}{\text{tr} [P_\gamma \tilde{Z} \tilde{Z}^T]} \right)^2 \middle| Z \right] \quad (330)$$

$$\begin{aligned} &= \frac{(u^T \zeta^T \Sigma_\gamma^{-1} \zeta \zeta^T \Sigma_\gamma^{-1} \zeta u)^2 + \mathbb{E}[(e^T P_\gamma \tilde{Z} \tilde{Z}^T P_\gamma e)^2|Z]}{\sigma_e^4 \text{tr} [P_\gamma \tilde{Z} \tilde{Z}^T]} \\ &+ \frac{2(u^T \zeta^T \Sigma_\gamma^{-1} \zeta \zeta^T \Sigma_\gamma^{-1} \zeta u) \mathbb{E}[e^T P_\gamma \tilde{Z} \tilde{Z}^T P_\gamma e|Z]}{\sigma_e^4 \text{tr} [P_\gamma \tilde{Z} \tilde{Z}^T]} \end{aligned} \quad (331)$$

$$\begin{aligned} &= \frac{\text{tr} [\zeta^T \Sigma_\gamma^{-1} \zeta \zeta^T \Sigma_\gamma^{-1} \zeta u u^T]^2 + \text{tr} [\Sigma_\gamma^{-1} \zeta \zeta^T \Sigma_\gamma^{-1}]^2 + 2 \text{tr} [(\Sigma_\gamma^{-1} \zeta \zeta^T \Sigma_\gamma^{-1})^2]}{\sigma_e^4 \text{tr} [\Sigma_\gamma^{-1} \zeta \zeta^T]^2} \\ &+ \frac{2 \text{tr} [\zeta^T \Sigma_\gamma^{-1} \zeta \zeta^T \Sigma_\gamma^{-1} \zeta u u^T] \text{tr} [\Sigma_\gamma^{-1} \zeta \zeta^T \Sigma_\gamma^{-1} \zeta \zeta^T]}{\sigma_e^4 \text{tr} [\Sigma_\gamma^{-1} \zeta \zeta^T]^2} \end{aligned} \quad (332)$$

$$= \mathbb{E}[\text{LHS}(\gamma)|Z]^2 + \frac{2 \text{tr} [\Sigma_\gamma^{-2} \zeta \zeta^T \Sigma_\gamma^{-2} \zeta \zeta^T]}{\sigma_e^4 \text{tr} [\Sigma_\gamma^{-1} \zeta \zeta^T]^2}. \quad (333)$$

Similarly, we have that

$$\mathbb{E}[\text{RHS}(\gamma)^2|Z] = \mathbb{E}[\text{RHS}(\gamma)|Z]^2 + \frac{2 \text{tr} [\Sigma_\gamma^{-4}]}{\sigma_e^4 \text{tr} [\Sigma_\gamma^{-1}]^2}. \quad (334)$$

Further, noting that,

$$\mathbb{E}[\text{LHS}(\gamma)|Z] \cdot \mathbb{E}[\text{RHS}(\gamma)|Z] = \frac{(\text{tr} [\zeta^T \Sigma_\gamma^{-1} \zeta \zeta^T \Sigma_\gamma^{-1} \zeta u u^T] + \text{tr} [\Sigma_\gamma^{-2} \zeta \zeta^T]) (\text{tr} [\zeta^T \Sigma_\gamma^{-2} \zeta u u^T] + \text{tr} [\Sigma_\gamma^{-2}])}{\sigma_e^4 \text{tr} [\Sigma_\gamma^{-1} \zeta \zeta^T] \text{tr} [\Sigma_\gamma^{-1}]}, \quad (335)$$

we can write

$$\mathbb{E}[\text{LHS}(\gamma)\text{RHS}(\gamma)|Z] = \mathbb{E}[\text{LHS}(\gamma)|Z] \cdot \mathbb{E}[\text{RHS}(\gamma)|Z] + \frac{2 \text{tr} [\Sigma_\gamma^{-4} \zeta \zeta^T]}{\sigma_e^4 \text{tr} [\Sigma_\gamma^{-1} \zeta \zeta^T] \text{tr} [\Sigma_\gamma^{-1}]}. \quad (336)$$

All together, we have

$$\mathbb{E}[\text{LHS}(\gamma)^2|Z] = \mathbb{E}[\text{LHS}(\gamma)|Z]^2 + \frac{2 \text{tr} [(\Sigma_\gamma^{-2} \zeta \zeta^T \Sigma_\gamma^{-2} \zeta \zeta^T)]}{\sigma_e^4 \text{tr} [\Sigma_\gamma^{-1} \zeta \zeta^T]^2}, \quad (337)$$

$$\mathbb{E}[\text{RHS}(\gamma)^2|Z] = \mathbb{E}[\text{RHS}(\gamma)|Z]^2 + \frac{2 \text{tr} [\Sigma_\gamma^{-4}]}{\sigma_e^4 \text{tr} [\Sigma_\gamma^{-1}]^2}, \quad (338)$$

$$\mathbb{E}[\text{LHS}(\gamma)\text{RHS}(\gamma)|Z] = \mathbb{E}[\text{LHS}(\gamma)|Z] \cdot \mathbb{E}[\text{RHS}(\gamma)|Z] + \frac{2\text{tr}[\Sigma_\gamma^{-4}\zeta\zeta^T]}{\sigma_e^4\text{tr}[\Sigma_\gamma^{-1}\zeta\zeta^T]\text{tr}[\Sigma_\gamma^{-1}]}, \quad (339)$$

$$\begin{aligned} \mathbb{E}[\Delta(\gamma)|Z]^2 &= \sigma_e^4 (\mathbb{E}[\text{RHS}(\gamma)|Z]^2 + \mathbb{E}[\text{LHS}(\gamma)|Z]^2 \\ &\quad - 2\mathbb{E}[\text{LHS}(\gamma)|Z] \cdot \mathbb{E}[\text{RHS}(\gamma)|Z]). \end{aligned} \quad (340)$$

Canceling terms, we thus have

$$\frac{1}{2\sigma_e^4} \text{Var}(\Delta(\gamma)|Z) = \mathbb{E}[\text{LHS}(\gamma)^2|Z] + \mathbb{E}[\text{RHS}(\gamma)^2|Z] - 2\mathbb{E}[\text{LHS}(\gamma)\text{RHS}(\gamma)|Z] - \mathbb{E}[\Delta(\gamma)|Z]^2. \quad (341)$$

$$= \frac{\text{tr}[\Sigma_\gamma^{-2}\zeta\zeta^T\Sigma_\gamma^{-2}\zeta\zeta^T]}{\text{tr}[\Sigma_\gamma^{-1}\zeta\zeta^T]^2} + \frac{\text{tr}[\Sigma_\gamma^{-4}]}{\text{tr}[\Sigma_\gamma^{-1}]^2} - 2\frac{\text{tr}[\Sigma_\gamma^{-4}\zeta\zeta^T]}{\text{tr}[\Sigma_\gamma^{-1}\zeta\zeta^T]\text{tr}[\Sigma_\gamma^{-1}]}. \quad (342)$$

We will show that these terms vanish with high probability. First note that

$$\frac{\text{tr}[\Sigma_\gamma^{-2}\zeta\zeta^T\Sigma_\gamma^{-2}\zeta\zeta^T]}{\text{tr}[\Sigma_\gamma^{-1}\zeta\zeta^T]^2} = \frac{(n-c)^{-2}\text{tr}[\Sigma_\gamma^{-2}\zeta\zeta^T\Sigma_\gamma^{-2}\zeta\zeta^T]}{(n-c)^{-2}\text{tr}[\Sigma_\gamma^{-1}\zeta\zeta^T]^2}. \quad (343)$$

The trace term in the numerator is a continuous, bounded spectral function of the eigenvalues of  $\zeta\zeta^T$ :

$$\text{tr}[\Sigma_\gamma^{-2}\zeta\zeta^T\Sigma_\gamma^{-2}\zeta\zeta^T] = \sum \lambda_i \psi_{(2,4)}(\lambda_i) \quad (344)$$

$$\implies (n-c)^{-2}\text{tr}[\Sigma_\gamma^{-2}\zeta\zeta^T\Sigma_\gamma^{-2}\zeta\zeta^T] = (n-c)^{-1} \left( (n-c)^{-1}\text{tr}[\Sigma_\gamma^{-2}\zeta\zeta^T\Sigma_\gamma^{-2}\zeta\zeta^T] \right) \quad (345)$$

$$\xrightarrow{P} (n-c)^{-1}\Psi_{(2,4)}. \quad (346)$$

On the other hand, we've already seen that  $(n-c)^{-1}\text{tr}[\Sigma_\gamma^{-1}\zeta\zeta^T] \xrightarrow{P} \Psi_{(1,1)}$ . Again applying Slutsky's theorem, we have

$$\frac{\text{tr}[\Sigma_\gamma^{-2}\zeta\zeta^T\Sigma_\gamma^{-2}\zeta\zeta^T]}{\text{tr}[\Sigma_\gamma^{-1}\zeta\zeta^T]^2} \xrightarrow{P} (n-c)^{-1} \frac{\Psi_{(2,4)}}{\Psi_{(1,1)}^2} = o_p(1). \quad (347)$$

By a similar argument, we have that

$$\frac{\text{tr}[\Sigma_\gamma^{-4}]}{\text{tr}[\Sigma_\gamma^{-1}]^2} \xrightarrow{P} (n-c)^{-1} \frac{\Psi_{(0,4)}}{\Psi_{(0,1)}^2} = o_p(1), \quad (348)$$

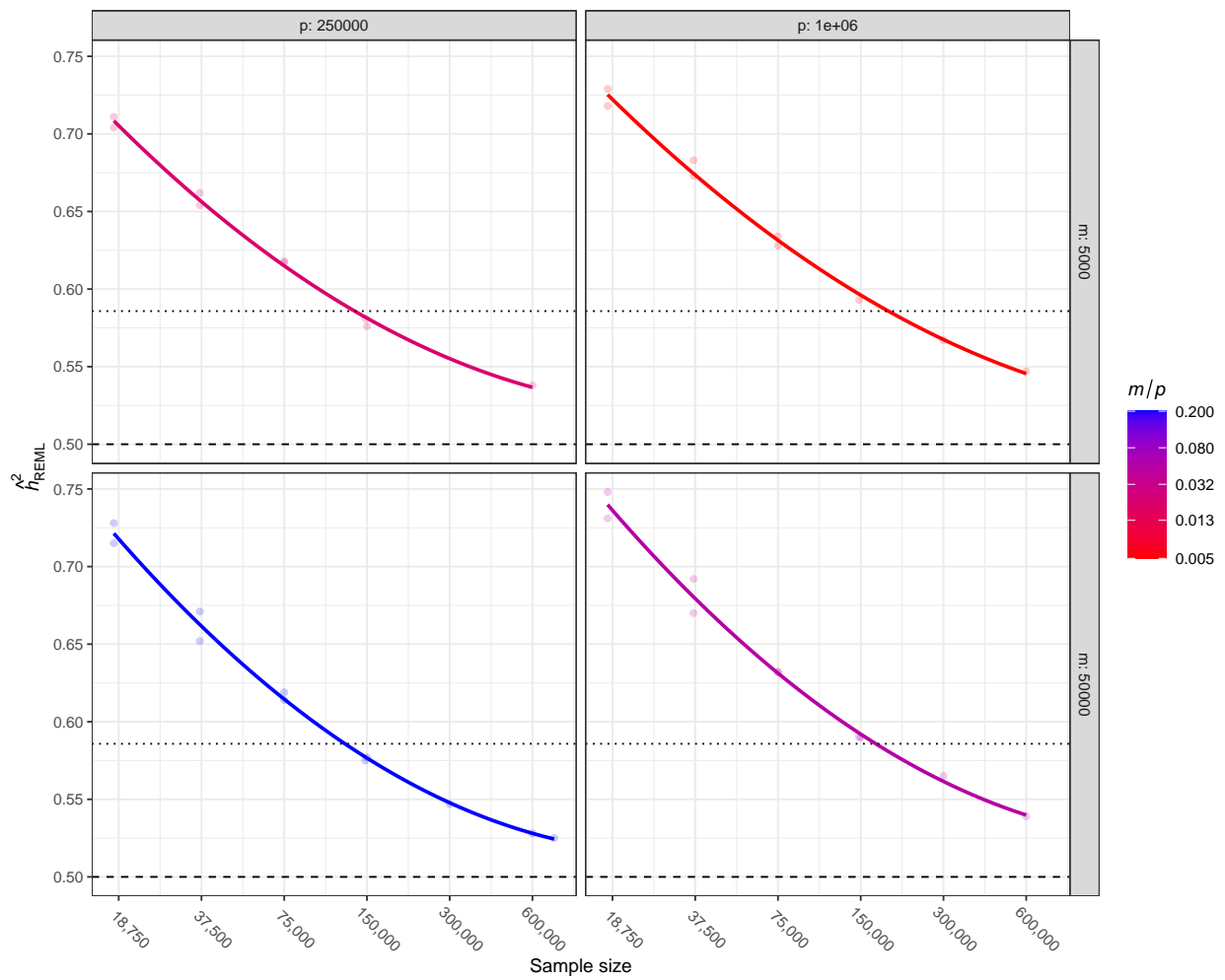
$$\frac{\text{tr}[\Sigma_\gamma^{-4}\zeta\zeta^T]}{\text{tr}[\Sigma_\gamma^{-1}\zeta\zeta^T]\text{tr}[\Sigma_\gamma^{-1}]} \xrightarrow{P} (n-c)^{-1} \frac{\Psi_{(1,4)}}{\Psi_{(1,1)}\Psi_{(0,1)}} = o_p(1), \quad (349)$$

thus yielding  $\text{Var}(\Delta(\gamma)|Z) \xrightarrow{P} 0$ . Now, applying the conditional form of Chebyshev's inequality, we have  $\forall \epsilon > 0$ ,

$$P(|\Delta(\gamma) - \mathbb{E}[\Delta(\gamma)|Z]| \geq \epsilon|Z) \leq \epsilon^{-2}\text{Var}(\Delta(\gamma)|Z) \xrightarrow{P} 0, \quad (350)$$

i.e., that  $\Delta(\gamma) - \mathbb{E}[\Delta(\gamma)|Z] \xrightarrow{P} 0$ . Together with the result of the previous section ( $|\mathbb{E}[\Delta(\gamma)|Z]| \xrightarrow{P} h(\gamma)$  with  $h(\gamma) = 0$  iff  $\gamma = \hat{\gamma}$ ), this implies that  $\hat{\gamma} \xrightarrow{P} \hat{\gamma}$ .

# Supplementary Figures



Supplementary Figure 1:  $\hat{h}_{\text{REML}}^2$  as a function of sample size ( $n$ ), the total number of SNPs included in the model ( $p$ ), and the number of causal variants ( $m$ ) in synthetic data. We observe convergence of  $\hat{h}_{\text{REML}}^2$  toward  $h_0^2$  for all values of  $m/p$ .

## Supplementary References

1. Fisher, R. A. The Correlation between Relatives on the Supposition of Mendelian Inheritance. en. *Earth and Environmental Science Transactions of The Royal Society of Edinburgh* **52**, 399–433 (1919/ed).
2. Crow, J. F. & Felsenstein, J. The Effect of Assortative Mating on the Genetic Composition of a Population. eng. *Eugenics Quarterly* **15**, 85–97 (June 1968).
3. Nagylaki, T. The Correlation between Relatives with Assortative Mating. eng. *Annals of human genetics* **42**, 131–137 (July 1978).
4. Nagylaki, T. Assortative Mating for a Quantitative Character. en. *Journal of Mathematical Biology* **16**, 57–74 (Dec. 1982).
5. Karlin, S. & Raper, J. Preferential Mating in Symmetric Multilocus Systems: Stability Conditions of the Central Equilibrium. en. *Genetics* **100**, 137–147 (Jan. 1982).
6. Gimelfarb, A. Quantitative Characters under Assortative Mating: Gametic Model. *Theoretical Population Biology* **25**, 312–330 (June 1984).
7. Bai, Z. & Silverstein, J. W. *Spectral Analysis of Large Dimensional Random Matrices* (Springer, 2010).
8. McDiarmid, C. *On the Method of Bounded Differences* en. /core/books/surveys-in-combinatorics-1989/on-the-method-of-bounded-differences/AABA597B562BDA7D89C6077E302694FB. Aug. 1989.
9. Tao, T. *Topics in Random Matrix Theory* (2011).
10. O’Rourke, S. & Williams, N. Pairing between Zeros and Critical Points of Random Polynomials with Independent Roots. *Transactions of the American Mathematical Society* **371**, 2343–2381 (2019).
11. Wright, F. T. A Bound on Tail Probabilities for Quadratic Forms in Independent Random Variables Whose Distributions Are Not Necessarily Symmetric. *Ann. Probability* **1**, 1068–1070 (1973).
12. Rudelson, M. & Vershynin, R. Hanson-Wright Inequality and Sub-Gaussian Concentration. EN. *Electronic Communications in Probability* **18** (2013).
13. Bobkov, S. G., Chistyakov, G. P. & Götze, F. Normal Approximation for Weighted Sums under a Second Order Correlation Condition. *arXiv:1906.09063 [math]*. arXiv: [1906.09063](https://arxiv.org/abs/1906.09063) [math] (June 2019).
14. Pozrikidis, C. *An Introduction to Grids, Graphs, and Networks* (Oxford University Press, 2014).
15. Fang, Y., Loparo, K. & Feng, X. Inequalities for the Trace of Matrix Product. *IEEE Transactions on Automatic Control* **39**, 2489–2490 (Dec. 1994).
16. Bao, Y. & Ullah, A. Expectation of Quadratic Forms in Normal and Nonnormal Variables with Applications. *Journal of Statistical Planning and Inference* **140**, 1193–1205 (2010).