SUPPLEMENTARY APPENDIX TO "BAYESIAN METHODS FOR GENETIC ASSOCIATION ANALYSIS WITH HETEROGENEOUS SUBGROUPS: FROM META-ANALYSES TO GENE-ENVIRONMENT INTERACTIONS"

BY XIAOQUAN WEN* AND MATTHEW STEPHENS†

University of Michigan and University of Chicago* †

APPENDIX A: COMPUTING BAYES FACTORS

In this section, we show the detailed calculations of various BFs.

A.1. Computation in ES Model. A particular ES model, describing an alternative hypothesis H_a, is fully specified by setting values for (ϕ, ω) and hyper-parameters $(v_1, \dots, v_S, l_1, m_1, \dots, l_S, m_S)$. Under the contrasting null model H_0, we set $\phi = \omega = 0$ while keeping all other hyper-parameters the same.

Let $\boldsymbol{\beta}_s = (\mu_s, \beta_s)$, $\tau_s = \sigma_s^{-2}$ and $\boldsymbol{\theta} = (\boldsymbol{\beta}_1, \dots, \boldsymbol{\beta}_s, \tau_1, \dots, \tau_s, \bar{b})$, the marginal likelihood under model H_a can be written as

(A.1)

$$P(\boldsymbol{Y}|\boldsymbol{G}, H_a) = \int P(\boldsymbol{Y}|\boldsymbol{G}, \boldsymbol{\theta}, H_a) p(\boldsymbol{\theta}|H_a) d\boldsymbol{\theta}$$

$$= \int \left(\prod_s P(\boldsymbol{y}_s|\boldsymbol{g}_s, \boldsymbol{\beta}_s, \tau_s) \prod_s P(\boldsymbol{\beta}_s|\tau_s, \bar{b}, H_a) \prod_s P(\tau_s|H_a) P(\bar{b}|H_a) \right) d\boldsymbol{\beta}_1 \cdots d\boldsymbol{\beta}_S d\tau_1 \cdots d\tau_S d\bar{b}$$

$$= \int \left(\int \left(\prod_s \int P(\boldsymbol{y}_s|\boldsymbol{g}_s, \boldsymbol{\beta}_s, \tau_s) P(\boldsymbol{\beta}_s|\tau_s, \bar{b}, H_a) d\boldsymbol{\beta}_s \right) p(\bar{b}|H_a) d\bar{b} \right) \prod_s P(\tau_s|H_a) d\tau_1 \cdots d\tau_S$$

Let $\boldsymbol{X}_s = (\boldsymbol{1}\ \boldsymbol{g}_s)$ denote the design matrix of regression model (2.1) of main text for subgroup s, it follows that

$$\text{(A.2)} \quad \begin{aligned} P(\boldsymbol{y}_s|\boldsymbol{g}_s, \boldsymbol{\beta}_s, \tau_s) &= (\frac{2\pi}{\tau_s})^{-n_s/2} \exp\left(-\frac{\tau_s}{2} (\boldsymbol{y}_s - \boldsymbol{X}_s \boldsymbol{\beta}_s)'(\boldsymbol{y}_s - \boldsymbol{X}_s \boldsymbol{\beta}_s\ ) \right) \\ &= (\frac{2\pi}{\tau_s})^{-n_s/2} \exp\left(-\frac{1}{2} (\tilde{\boldsymbol{y}}_s - \boldsymbol{X}_s \boldsymbol{b}_s)'(\tilde{\boldsymbol{y}}_s - \boldsymbol{X}_s \boldsymbol{b}_s\ ) \right), \end{aligned}$$

where $\tilde{\boldsymbol{y}}_s = \sqrt{\tau_s} \boldsymbol{y}_s$ and $\boldsymbol{b}_s = \sqrt{\tau_s} \boldsymbol{\beta}_s = (\sqrt{\tau_s} \mu_s, b_s)$. We further denote

$$\text{(A.3)} \quad \bar{\boldsymbol{b}} = \begin{pmatrix} 0 \\ \bar{b} \end{pmatrix} \quad \text{and} \quad \Phi_s = \begin{pmatrix} v_s^2 & 0 \\ 0 & \phi^2 \end{pmatrix},$$

and write prior distribution $P(\boldsymbol{b}_s|\bar{\boldsymbol{b}}, H_a)$ in following matrix form,

$$\boldsymbol{b}_s|\bar{\boldsymbol{b}}, H_a \sim \mathrm{N}(\bar{\boldsymbol{b}}, \Phi_s). \tag{A.4}$$

We compute the marginal likelihood by sequentially evaluating the following integrals,

(A.5)

$$F_{H_a,s} = \int P(\boldsymbol{y}_s|\boldsymbol{X}_s, \boldsymbol{b}_s, \tau_s) P(\boldsymbol{b}_s|\bar{\boldsymbol{b}}, H_a) d\boldsymbol{b}_s$$
$$= (\frac{2\pi}{\tau_s})^{-n_s/2} |\Phi_s|^{-\frac{1}{2}} \cdot |\boldsymbol{X}_s'\boldsymbol{X}_s + \Phi_s^{-1}|^{-\frac{1}{2}}$$
$$\cdot \exp\left(-\frac{1}{2}\left(\tilde{\boldsymbol{y}}_s'\tilde{\boldsymbol{y}}_s - (\boldsymbol{X}_s'\tilde{\boldsymbol{y}}_s + \Phi_s^{-1}\bar{\boldsymbol{b}})'(\boldsymbol{X}_s'\boldsymbol{X}_s + \Phi_s^{-1})^{-1}(\boldsymbol{X}_s'\tilde{\boldsymbol{y}}_s + \Phi_s^{-1}\bar{\boldsymbol{b}}) + \bar{\boldsymbol{b}}'\Phi_s^{-1}\bar{\boldsymbol{b}}\right)\right).$$

Let $J_{H_a} = \int (\prod_s F_{H_a,s}) P(\bar{b}|H_a) d\bar{b}$; this quantity is also analytically computable by straightforward algebra.

To compute BF of H_a versus H_0 under the ES model, we take limits with respect to hyper-parameters $(v_1, \ldots, v_S, l_1, m_1, \ldots, l_S, m_S)$, such that $u_s^2 \to \infty$ and $l_s, m_s \to 0,\ \forall s$. This yields,

$$\begin{aligned} \mathrm{BF}^{\mathrm{ES}}(\phi, \omega) &= \lim \frac{\int J_{H_a} \prod_s P(\tau_s) d\tau_1 \cdots d\tau_S}{\int J_{H_0} \prod_s P(\tau_s) d\tau_1 \cdots d\tau_S} \\ &= \frac{\int K_{H_a} d\tau_1 \cdots d\tau_S}{\int K_{H_0} d\tau_1 \cdots d\tau_S}. \end{aligned} \tag{A.6}$$

Let us denote

$$\mathrm{RSS}_{0,s} = \boldsymbol{y}_s'\boldsymbol{y}_s - n_s\bar{y}_s^2, \tag{A.7}$$

$$\mathrm{RSS}_{1,s} = \boldsymbol{y}_s'\boldsymbol{y}_s - \boldsymbol{y}_s'\boldsymbol{X}_s(\boldsymbol{X}_s'\boldsymbol{X}_s)^{-1}\boldsymbol{X}_s'\boldsymbol{y}_s, \tag{A.8}$$

$$\delta_s^2 = \frac{1}{\boldsymbol{g}_s'\boldsymbol{g}_s - n_s\bar{g}_s^2}, \tag{A.9}$$

$$\hat{\beta}_s = \frac{\boldsymbol{y}_s'\boldsymbol{g}_s - n_s\bar{y}_s\bar{g}_s}{\boldsymbol{g}_s'\boldsymbol{g}_s - n_s\bar{g}_s^2}, \tag{A.10}$$

$$\zeta^2 = \frac{1}{\sum_s(\delta_s^2 + \phi^2)^{-1}}, \tag{A.11}$$

where $\bar{y}_s$ and $\bar{g}_s$ are the sample means of phenotypes and genotypes in subgroup s. It can be shown that,

$$K_{H_0} = \prod_s \tau_s^{\frac{n_s}{2}-1} \exp\left(-\frac{1}{2}\sum_s \tau_s \cdot \mathrm{RSS}_{0,s}\right), \tag{A.12}$$

and
(A.13)

$$
\begin{aligned}
K_{H_a} =& \sqrt{\frac{\zeta^2}{\zeta^2+\omega^2}} \prod_s \sqrt{\frac{\delta_s^2}{\delta_s^2+\phi^2}} \\
& \cdot \prod_s \tau_s^{\frac{n_s}{2}-1} \exp\left(-\frac{1}{2}\sum_s \tau_s \left(\frac{\phi^2}{\delta_s^2+\phi^2}\cdot \mathrm{RSS}_{1,s} + \frac{\delta_s^2}{\delta_s^2+\phi^2}\cdot \mathrm{RSS}_{0,s}\right)\right) \\
& \cdot \exp\left(\frac{1}{2}\frac{\omega^2\zeta^2}{\zeta^2+\omega^2}\left(\sum_s \frac{\hat\beta_s\sqrt{\tau_s}}{\delta_s^2+\phi^2}\right)^2\right).
\end{aligned}
$$

The multidimensional integral $\int K_{H_a} d\tau_1 \cdots d\tau_S$ generally does not have a simple analytic form (although it can be represented as finite sums of complicated hypergeometric functions). Next, we show two different approximations, both based on Laplace's method, to evaluate this integral. The first approximation is a direct application of Butler and Wood (2002) and the second one yields a simple analytic expression. Although the integral $\int K_{H_0} d\tau_1 \cdots d\tau_S$ can be analytically computed as a gamma function, for computing the BF, we also use Laplace's method to numerically evaluate it (which essentially is applying Sterling's formula) – we find this recipe yields more accurate result for the final BF: in particular, when there is only one subgroup ($S = 1$, where the BF can be analytically computed as in Servin and Stephens (2008)), we obtain the exact result by applying the first Laplace's approximation.

Laplace's method approximates a multivariate integral in the following way,

$$
\int_D h(\boldsymbol{\tau}) e^{g(\boldsymbol{\tau})} d\boldsymbol{\tau} \approx (2\pi)^{S/2} |H_{\hat{\boldsymbol{\tau}}}|^{-1/2} h(\hat{\boldsymbol{\tau}}) e^{g(\hat{\boldsymbol{\tau}})} \tag{A.14}
$$

where $\boldsymbol{\tau}$ is an S-vector,

$$
\hat{\boldsymbol{\tau}} = \arg\max_{\boldsymbol{\tau}} g(\boldsymbol{\tau}), \tag{A.15}
$$

and $|H_{\hat{\boldsymbol{\tau}}}|$ is the absolute value of the determinant of the Hessian matrix of the function g evaluated at $\hat{\boldsymbol{\tau}}$. Note that the factorization of the integrand is rather arbitrary, it only requires that function h is smooth and positively valued and the smooth function g has a unique maximum lying in the interior of D (for detailed discussion, see Butler (2007)).

Our first approach to apply Laplace's method sets $h(\boldsymbol{\tau}) \equiv 1$. Except for some trivial situations (e.g. $S = 1$), the maximization of $\log K_{H_a}$ with respect to $\boldsymbol{\tau}$ is analytically intractable. In practice, we use the Broyden-Fletcher-Goldfarb-Shanno (BFGS2) algorithm, a gradient-based numerical optimization routine (implemented in the GNU Scientific Library), to perform numerical maximization. This procedure leads to $\widehat{\mathrm{BF}^{\mathrm{ES}}}(\phi, \omega)$.

Alternatively, we apply Laplace's method by factoring the integrand in such a way that g can be analytically maximized. This approach results in a closed-form approximation. More specifically, we factor K_{H_a} into

$$K_{H_a} = h(\tau_1, \ldots, \tau_S) e^{g(\tau_1, \ldots, \tau_S)}, \tag{A.16}$$

where

$$\begin{aligned} h(\tau_1, \ldots, \tau_S) = & \sqrt{\frac{\zeta^2}{\zeta^2 + \omega^2}} \prod_s \sqrt{\frac{\delta_s^2}{\delta_s^2 + \phi^2}} \\ & \cdot \prod_s \exp\left(-\frac{1}{2} \sum_s \frac{\delta_s^2}{\delta_s^2 + \phi^2} \cdot \tau_s \cdot (\mathrm{RSS}_{0,s} - \mathrm{RSS}_{1,s})\right) \\ & \cdot \exp\left(\frac{1}{2} \frac{\omega^2 \zeta^2}{\zeta^2 + \omega^2} \left(\sum_s \frac{\hat{\beta}_s \sqrt{\tau_s}}{\delta_s^2 + \phi^2}\right)^2\right) \end{aligned} \tag{A.17}$$

and

$$e^{g(\tau_1, \ldots, \tau_S)} = \prod_s \tau_s^{\frac{n_s}{2} - 1} \cdot \exp\left(-\frac{1}{2} \sum_s \tau_s \cdot \mathrm{RSS}_{1,s}\right). \tag{A.18}$$

It is straightforward to show that the unique maximum of $g(\tau_1, \ldots, \tau_S)$ is attained at

$$\hat{\tau}_s = \frac{n_s - 2}{\mathrm{RSS}_{1,s}}, \;\; s = 1, \ldots, S, \tag{A.19}$$

which coincides with the REML estimate of τ_s in subgroup-level regression model (2.1) of main text. Similarly, we factor K_{H_0} into

(A.20)

$$K_{H_0} = \exp\left(-\frac{1}{2} \sum_s \tau_s \left(\mathrm{RSS}_{0,s} - \mathrm{RSS}_{1,s}\right)\right) \cdot \prod_s \tau_s^{\frac{n_s}{2} - 1} \cdot \exp\left(-\frac{1}{2} \sum_s \tau_s \cdot \mathrm{RSS}_{1,s}\right),$$

and expand it around $\hat{\tau}_s$ as well.

Following the notations in section 2.3 of main text and noting the relationship between t and F statistics in the simple linear regression,

$$T_s^2 = \hat{\tau}_s \cdot (\mathrm{RSS}_{0,s} - \mathrm{RSS}_{1,s}), \tag{A.21}$$

Applying (A.14) results in

(A.22)
$$\mathrm{BF}^{\mathrm{ES}}(\phi,\omega) = \sqrt{\frac{\zeta^2}{\zeta^2+\omega^2}}\exp\left(\frac{\mathcal{T}_{\mathrm{ES}}^2}{2}\frac{\omega^2}{\zeta^2+\omega^2}\right)\cdot\prod_s\left(\sqrt{\frac{\delta_s^2}{\delta_s^2+\phi^2}}\exp\left(\frac{T_s^2}{2}\frac{\phi^2}{\delta_s^2+\phi^2}\right)\right) \cdot\left(1+O\left(\frac{1}{n_s}\right)\right).$$

In conclusion, we have obtained

(A.23)
$$\mathrm{ABF}^{\mathrm{ES}}(\phi,\omega) = \sqrt{\frac{\zeta^2}{\zeta^2+\omega^2}}\exp\left(\frac{\mathcal{T}_{\mathrm{ES}}^2}{2}\frac{\omega^2}{\zeta^2+\omega^2}\right)\prod_s\left(\sqrt{\frac{\delta_s^2}{\delta_s^2+\phi^2}}\exp\left(\frac{T_s^2}{2}\frac{\phi^2}{\delta_s^2+\phi^2}\right)\right).$$

Remarks. Note, in case $\tau_1,\ldots,\tau_S$ are known, we can directly compute the exact BF using

$$\mathrm{BF}^{\mathrm{ES}}(\phi,\omega) = \lim\frac{J_{H_a}}{J_{H_0}} \tag{A.24}$$

without evaluating the multi-dimensional integrals in (A.6). In this particular case, it is easy to show that the exact BF has the exact functional form as in (A.23), only with all the $\hat{\tau}_s$'s replaced by the corresponding true values of τ_s's.

Finally, we give the proof of Proposition 4.1:

PROOF. The derivation above serves as a proof. An alternative proof can be obtained by noting that the REML estimate of $\hat{\boldsymbol{\tau}}$ asymptotically converges to the true value of $\boldsymbol{\tau}$ with probability 1. From the remarks above, by applying continuous mapping theorem, we conclude that $\mathrm{ABF}^{\mathrm{ES}}(\phi,\omega)$ converges to $\mathrm{BF}^{\mathrm{ES}}(\phi,\omega)$ with probability 1. □

A.2. Computation in EE Model. The procedure for computing BF assuming an EE model is essentially the same, we omit repeating the details but only show the final results of the BF of an EE model H_b, specified by (ψ,w), versus the null model H_0,

$$\mathrm{BF}^{\mathrm{EE}}(\psi,w) = \frac{\int K_{H_b}d\tau_1\cdots d\tau_S}{\int K_{H_0}d\tau_1\cdots d\tau_S}. \tag{A.25}$$

The expression of K_{H_0} remains the same as (A.12). We denote

$$\eta^2 = \left(\sum_s\frac{\tau_s}{\delta_s^2+\tau_s\psi^2}\right)^{-1}. \tag{A.26}$$

It can be shown
(A.27)

$$
\begin{aligned}
K_{H_b} = & \sqrt{\frac{\eta^2}{\eta^2 + w^2}} \prod_s \sqrt{\frac{\delta_s^2}{\delta_s^2 + \tau_s \psi^2}} \\
& \cdot \prod_s \tau_s^{\frac{n_s}{2} - 1} \exp\left(-\frac{1}{2} \sum_s \tau_s \left(\frac{\tau_s \psi^2}{\delta_s^2 + \tau_s \psi^2} \cdot \mathrm{RSS}_{1,s} + \frac{\delta_s^2}{\delta_s^2 + \tau_s \psi^2} \cdot \mathrm{RSS}_{0,s} \right) \right) \\
& \cdot \exp\left(\frac{1}{2} \frac{w^2}{\eta^2 + w^2} \frac{\left(\sum_s \frac{\tau_s}{\delta_s^2 + \tau_s \psi^2} \cdot \hat{\beta}_s \right)^2}{\eta^2} \right).
\end{aligned}
$$

We use the similar numerical procedure to obtain $\widehat{\mathrm{BF}^{\mathrm{EE}}}(\psi, w)$ as in the ES model.

To derive $\mathrm{ABF}^{\mathrm{EE}}$, we factor K_{H_b} into

$$
K_{H_b} = h(\tau_1, \ldots, \tau_S) e^{g(\tau_1, \ldots, \tau_S)}, \tag{A.28}
$$

where,

$$
\begin{aligned}
h(\tau_1, \ldots, \tau_S) = & \sqrt{\frac{\eta^2}{\eta^2 + w^2}} \prod_s \sqrt{\frac{\delta_s^2}{\delta_s^2 + \tau_s \psi^2}} \\
& \cdot \prod_s \exp\left(-\frac{1}{2} \sum_s \frac{\tau_s \delta_s^2}{\delta_s^2 + \tau_s \psi^2} \cdot (\mathrm{RSS}_{0,s} - \mathrm{RSS}_{1,s}) \right) \\
& \cdot \exp\left(\frac{1}{2} \frac{w^2}{\eta^2 + w^2} \frac{\left(\sum_s \frac{\tau_s}{\delta_s^2 + \tau_s \psi^2} \cdot \hat{\beta}_s \right)^2}{\eta^2} \right)
\end{aligned} \tag{A.29}
$$

and

$$
e^{g(\tau_1, \ldots, \tau_S)} = \prod_s \tau_s^{\frac{n_s}{2} - 1} \cdot \exp\left(-\frac{1}{2} \sum_s \tau_s \cdot \mathrm{RSS}_{1,s} \right). \tag{A.30}
$$

Again, function $g(\tau_1, \ldots, \tau_S)$ is maximized at

$$
\hat{\tau}_s = \frac{n_s - 2}{\mathrm{RSS}_{1,s}}, \; s = 1, \ldots, S. \tag{A.31}
$$

We denote

$$d_s^2 = \frac{1}{\hat{\tau}_s}\delta_s^2 = \frac{\hat{\sigma}_s^2}{\boldsymbol{g}_s'\boldsymbol{g}_s - n_s\bar{g}_s^2}, \tag{A.32}$$

$$T_s^2 = \frac{\hat{\beta}_s}{d_s^2}, \tag{A.33}$$

$$\xi^2 = \left(\sum_s \frac{\hat{\tau}_s}{\delta_s^2 + \hat{\tau}_s\psi^2}\right)^{-1} = \frac{1}{\sum_s (d_s^2 + \psi^2)^{-1}}, \tag{A.34}$$

$$\hat{\bar{\beta}} = \frac{\sum_s (d_s^2 + \psi^2)^{-1}\hat{\beta}_s}{\sum_s (d_s^2 + \psi^2)^{-1}}, \tag{A.35}$$

and

$$\mathcal{T}_{\text{EE}}^2 = \frac{\hat{\bar{\beta}}^2}{\xi^2} = \frac{\left(\sum_s \frac{\hat{\beta}}{d_s^2 + \psi^2}\right)^2}{\eta^2}. \tag{A.36}$$

Using a similar procedure as in the ES model, we obtain

$$\text{ABF}^{\text{EE}}(\psi, w) = \sqrt{\frac{\xi^2}{\xi^2 + w^2}} \exp\left(\frac{\mathcal{T}_{\text{EE}}^2}{2}\frac{w^2}{\xi^2 + w^2}\right) \prod_s \left(\sqrt{\frac{d_s^2}{d_s^2 + \psi^2}} \exp\left(\frac{T_s^2}{2}\frac{\psi^2}{d_s^2 + \psi^2}\right)\right). \tag{A.37}$$

As discussed in **Remarks** of section A.1, if $\tau_1, \ldots, \tau_S$ are known, the exact BF of the EE model has the same function form as in (A.37), with $\hat{\tau}_s$'s replaced by the corresponding τ_s's.

A.3. Computation using CEFN Priors. Using curved exponential family normal prior, the computation of the BFs is slightly different than what we show in the previous sections. Here, we use the ES model as a demonstration, the procedure for the EE model is very similar.

To compute the BF of a CEFN-ES model defined by parameters (k, ω) vs. the null model, we can carry out the same and exact calculation up to (A.5). However, due to the nature of the CEFN prior, we can no longer perform analytic calculation to integrate out $\bar{b}$. Instead, we exchange the order of integrations by first analytically approximating the multi-dimensional integration with respect to $\tau_1, \ldots, \tau_S$ using the second procedure of Laplace's method described in previous sections. As a result, we obtain the approximate BF as a one-dimensional integral

(A.38)

$$\mathrm{ABF}^{\mathrm{ES}}_{\mathrm{CEFN}}(k,\omega) = \frac{1}{\sqrt{2\pi}\omega}\prod_s \left(\frac{\mathrm{RSS}_{0,s}}{\mathrm{RSS}_{1,s}}\right)^{n_s/2} \int_{-\infty}^{\infty} \prod_s \sqrt{\frac{\delta_s^2}{\delta_s^2 + k\bar{b}^2}}$$
$$\cdot \exp\left[-\frac{1}{2}\left((\sum_s \frac{1}{\delta_s^2+k\bar{b}^2} + \frac{1}{\omega^2})\bar{b}^2 - 2\sum_s(\frac{\hat{b}_s}{\delta_s^2+k\bar{b}^2})\bar{b} + \sum_s \frac{\delta_s^2}{\delta_s^2+k\bar{b}^2}T_s^2\right)\right] d\bar{b}.$$

We then apply an adaptive Gaussian quadrature method, QAGI (implemented in GNU scientific library), to numerically evaluate this integral. Essentially, this method first maps the integrand to the semi-open interval $[0,1)$ using the transformation $y = (1-\bar{b})/\bar{b}$, then apply the standard adaptive Gaussian quadrature routine for the finite interval integration.

For the EE model with CEFN prior, the final one-dimensional integral can be shown as

(A.39)

$$\mathrm{ABF}^{\mathrm{EE}}_{\mathrm{CEFN}}(k,w) = \frac{1}{\sqrt{2\pi}w}\prod_s \left(\frac{\mathrm{RSS}_{0,s}}{\mathrm{RSS}_{1,s}}\right)^{n_s/2} \int_{-\infty}^{\infty} \prod_s \sqrt{\frac{d_s^2}{d_s^2 + k\bar{\beta}^2}}$$
$$\cdot \exp\left[-\frac{1}{2}\left((\sum_s \frac{1}{d_s^2+k\bar{\beta}^2} + \frac{1}{w^2})\bar{\beta}^2 - 2\sum_s(\frac{\hat{\beta}_s}{d_s^2+k\bar{\beta}^2})\bar{\beta} + \sum_s \frac{d_s^2}{d_s^2+k\bar{\beta}^2}T_s^2\right)\right] d\bar{\beta}.$$

APPENDIX B: BF FOR GENERALIZED LINEAR MODELS

In this section, we show the computation of BF for generalized linear models. As a useful application in genetic association study, these results can be directly applied to case-control data where a logistic link function is typically used.

For an appropriate link function $g(\cdot)$, we modify the subgroup-level linear model (2.1) of main text into

$$\mathrm{E}(\boldsymbol{y}_s) = g^{-1}\left(\mu_s \mathbf{1} + \beta_s \boldsymbol{g}_s\right). \tag{B.1}$$

Let us denote $\boldsymbol{\beta}_s = (\mu_s, \beta_s)$. The key component in our computation is to approximate subgroup-level log-likelihood function $l(\boldsymbol{\beta}_s)$ with a quadratic form expanding around its maximum likelihood estimate, i.e.

$$\log P(\boldsymbol{y}_s|\boldsymbol{g}_s,\boldsymbol{\beta}_s) = l(\boldsymbol{\beta}_s) \simeq l(\hat{\boldsymbol{\beta}}_s) - \frac{1}{2}(\boldsymbol{\beta}_s - \hat{\boldsymbol{\beta}}_s)' I_s(\hat{\boldsymbol{\beta}}_s)(\boldsymbol{\beta}_s - \hat{\boldsymbol{\beta}}_s), \tag{B.2}$$

where $I_s(\hat{\boldsymbol{\beta}}_s) = \begin{pmatrix} i_{\hat{\mu}_s\hat{\mu}_s} & i_{\hat{\mu}_s\hat{\beta}_s} \\ i_{\hat{\beta}_s\hat{\mu}_s} & i_{\hat{\beta}_s\hat{\beta}_s} \end{pmatrix}$ is the expected Fisher information evaluated at $\hat{\boldsymbol{\beta}}_s$. We note that

$$\gamma_s^2 := \mathrm{Var}(\hat{\beta}_s) = (i_{\hat{\beta}_s\hat{\beta}_s} - i_{\hat{\beta}_s\hat{\mu}_s} i^{-1}_{\hat{\mu}_s\hat{\mu}_s} i_{\hat{\mu}_s\hat{\beta}_s})^{-1} \tag{B.3}$$

is the estimated asymptotic variance of MLE $\hat{\beta}_s$.

Given approximate log-likelihood function (B.2) and a model H_c specified by (ψ, w), the prior distribution for $\boldsymbol{\beta}_s$ is given by

$$\boldsymbol{\beta}_s | \bar{\boldsymbol{\beta}}, H_c \sim \mathrm{N}(\bar{\boldsymbol{\beta}}, \Psi_s), \tag{B.4}$$

where

$$\bar{\boldsymbol{\beta}} = \begin{pmatrix} 0 \\ \bar{\beta} \end{pmatrix} \quad \text{and} \quad \Psi_s = \begin{pmatrix} v_s^2 & 0 \\ 0 & \psi^2 \end{pmatrix}. \tag{B.5}$$

It follows that

$$\begin{aligned} F_{H_c,s} = & \int P(\boldsymbol{y}_s | \boldsymbol{g}_s, \boldsymbol{\beta}_s) P(\boldsymbol{\beta}_s | \bar{\boldsymbol{\beta}}, H_c) d\boldsymbol{\beta}_s \\ = & \exp\big(l(\hat{\boldsymbol{\beta}}_s)\big) \cdot |\Psi_s|^{-\frac{1}{2}} \cdot |I_s + \Psi_s^{-1}|^{-\frac{1}{2}} \cdot \exp\left(-\frac{1}{2}\hat{\boldsymbol{\beta}}_s'\big(I_s - I_s(I_s + \Psi_s^{-1})^{-1} I_s\big)\hat{\boldsymbol{\beta}}_s\right) \\ & \cdot \exp\left(-\frac{1}{2}\Big(\ \bar{\boldsymbol{\beta}}'\big(\Psi_s^{-1} - \Psi_s^{-1}(I_s + \Psi_s^{-1})^{-1}\Psi_s^{-1}\big)\bar{\boldsymbol{\beta}} - \bar{\boldsymbol{\beta}}'\eta_s - \eta_s'\bar{\boldsymbol{\beta}}\Big)\ \right), \end{aligned} \tag{B.6}$$

with $\eta_s = \Psi^{-1}(I_s + \Psi^{-1})^{-1} I_s \hat{\boldsymbol{\beta}}_s$.

Under contrasting null model H_0, the parameter space is restricted to $\beta_s = 0$, for $\boldsymbol{\beta}_s$ satisfies this restriction

$$(\boldsymbol{\beta}_s - \hat{\boldsymbol{\beta}}_s)' I_s(\hat{\boldsymbol{\beta}}_s)(\boldsymbol{\beta}_s - \hat{\boldsymbol{\beta}}_s) = i_{\hat{\mu}_s\hat{\mu}_s} \cdot (\mu_s - \hat{m}_s)^2 + \frac{\hat{\beta}_s^2}{\gamma_s^2}, \tag{B.7}$$

where $\hat{m}_s = \hat{\mu}_s + \frac{i_{\hat{\mu}_s\hat{\beta}_s}}{i_{\hat{\mu}_s\hat{\mu}_s}}\hat{\beta}_s$. It can be shown that

$$\begin{aligned} F_{H_0,s} = & \int P(\boldsymbol{y}_s | \boldsymbol{g}_s, \boldsymbol{\beta}_s) P(\boldsymbol{\beta}_s | \bar{\boldsymbol{\beta}}, H_0) d\boldsymbol{\beta}_s \\ = & \exp\big(l(\hat{\boldsymbol{\beta}}_s)\big) \cdot v_s^{-1}(i_{\hat{\mu}\hat{\mu}} + v_s^{-2})^{-\frac{1}{2}} \cdot \exp(-\frac{\hat{\beta}_s^2}{2\gamma_s^2}) \\ & \cdot \exp\left(-\frac{1}{2}\big(\ \hat{m}_s' i_{\hat{\mu}\hat{\mu}} \hat{m}_s - (i_{\hat{\mu}\hat{\mu}}\hat{m}_s)'(i_{\hat{\mu}\hat{\mu}} + v_s^{-2})^{-1}(i_{\hat{\mu}\hat{\mu}}\hat{m}_s)\ \big)\right). \end{aligned} \tag{B.8}$$

Finally, we compute

$$\mathrm{ABF}(\psi, w) = \lim \frac{\int (\prod_s F_{H_c,s}) P(\bar{\beta} | H_c)\ d\bar{\beta}}{\prod_s F_{H_0,s}}, \tag{B.9}$$

where the limit is taken as $v_s \to \infty$, $\forall s$. By straightforward algebra, we obtain the following final result

(B.10)
$$\mathrm{BF}(\psi, w) \approx \mathrm{ABF}(\psi, w) := \mathrm{ABF}_{\mathrm{single}}(\mathcal{Z}_{\mathrm{cc}}^2, \xi; w) \cdot \prod_s \mathrm{ABF}_{\mathrm{single}}(Z_s^2, \gamma_s; \psi),$$

where

$$\mathrm{ABF}_{\mathrm{single}}(Z_s^2, \gamma_s; \psi) = \sqrt{\frac{\gamma_s^2}{\gamma_s^2 + \psi^2}} \exp\left(\frac{Z_s^2}{2} \frac{\psi^2}{\gamma_s^2 + \psi^2}\right), \tag{B.11}$$

$$\mathrm{ABF}_{\mathrm{single}}^{\mathrm{CC}}(\mathcal{Z}_{\mathrm{cc}}^2, \xi; w) = \sqrt{\frac{\xi^2}{\xi^2 + w^2}} \exp\left(\frac{\mathcal{Z}_{\mathrm{cc}}^2}{2} \frac{w^2}{\xi^2 + w^2}\right), \tag{B.12}$$

(B.13)

and

$$\gamma_s^2 := \mathrm{se}(\hat{\beta}_s)^2, \tag{B.14}$$

$$Z_s^2 = \frac{\hat{\beta}_s^2}{\gamma_s^2}, \tag{B.15}$$

$$\hat{\bar{\beta}} = \frac{\sum_s (\gamma_s^2 + \psi^2)^{-1} \hat{\beta}_s}{\sum_s (\gamma_s^2 + \psi^2)^{-1}}, \tag{B.16}$$

$$\xi^2 := \mathrm{se}(\hat{\bar{\beta}})^2 = \frac{1}{\sum_s (\gamma_s^2 + \psi^2)^{-1}}, \tag{B.17}$$

$$\mathcal{Z}_{\mathrm{cc}}^2 = \frac{\hat{\bar{\beta}}^2}{\zeta^2}. \tag{B.18}$$

APPENDIX C: SMALL SAMPLE SIZE CORRECTION FOR APPROXIMATE BAYES FACTORS

The accuracy of $\mathrm{ABF}^{\mathrm{ES}}$ relies on the sample sizes in subgroups: when sample sizes are small in some subgroups, the approximation may become inaccurate. In particular, we consider the behavior of the approximate BF when the null hypothesis is true. A valid BF has the property that

$$\mathrm{E}(\mathrm{BF}|H_0) = 1, \tag{C.1}$$

where the expectation is taken with respect to the data distribution ($\boldsymbol{Y}$ in our settings) under the null model. This is because,

$$\mathrm{E}(\mathrm{BF}|H_0) = \int \frac{P(\boldsymbol{Y}|H_1)}{P(\boldsymbol{Y}|H_0)} \cdot P(\boldsymbol{Y}|H_0)\, d\boldsymbol{Y} = 1. \tag{C.2}$$

Unfortunately, when sample sizes are small, (C.1) can be violated when the approximate BF in used (as the expected value is strictly greater than 1) and this essentially indicates that the approximation becomes inaccurate.

To demonstrate a violation of (C.1), we consider the special case of one single subgroup. The approximate BF assuming the ES model with parameters (ϕ, ω) is given by

$$\text{ABF}^{\text{ES}}_{\text{single}}(\phi, \omega) = \sqrt{1-\lambda}\exp(\frac{\lambda}{2}T_s^2), \tag{C.3}$$

and,

$$\log\left(\text{ABF}^{\text{ES}}_{\text{single}}(\phi, \omega)\right) = \frac{1}{2}\log(1-\lambda) + \frac{\lambda}{2}T_s^2, \tag{C.4}$$

where $\lambda = \frac{\phi^2+\omega^2}{\phi^2+\omega^2+\delta_s^2}$ and takes values from $[0, 1]$. Under H_0, T_s follows t-distribution with $n_s - 2$ degree of freedom and

$$\text{E}(T_s^2|H_0) = \frac{n_s-2}{n_s-4} > 1. \tag{C.5}$$

Now consider the continuous function

$$f(\lambda) = \frac{1}{\lambda}\log(\frac{1}{1-\lambda}) \tag{C.6}$$

for $\lambda \in [0, 1]$, it can be shown that

$$\lim_{\lambda\to 0} f(\lambda) = 1 \tag{C.7}$$

$$\lim_{\lambda\to 1} f(\lambda) = \infty. \tag{C.8}$$

Hence, there must exist values of $\lambda \in (0, 1)$, such that

$$1 < f(\lambda) < \text{E}(T_s^2|H_0). \tag{C.9}$$

Consequently, by Jensen's inequality, for those λ values

$$\log\left(\text{E}\left(\text{ABF}^{\text{ES}}_{\text{single}}|H_0\right)\right) \geq \text{E}\left(\log\left(\text{ABF}^{\text{ES}}_{\text{single}}\right)|H_0\right) > 0. \tag{C.10}$$

This shows property (C.1) does not generally hold for approximate BFs, and when sample size n_s is small, the inaccuracy may become severe.

We now propose a simple correction procedure for small sample sizes, which ensures the resulting approximation satisfies property (C.1). Specifically, we modify expression (4.9) of main text into the following form

$$\text{A}^*\text{BF}^{\text{ES}}(\phi, \omega) = \text{ABF}^{\text{ES}}_{\text{single}}(q(\mathcal{T}_{\text{es}}{}^2), \zeta; \omega) \cdot \prod_s \text{ABF}^{\text{ES}}_{\text{single}}(q_s(T_s^2), \delta_s; \phi), \tag{C.11}$$

where the function q_s denotes a one-to-one quantile transformation from a t-distribution with $n_s - 2$ degree of freedom to a standard normal distribution, and the function q is defined as

$$q(\mathcal{T}_{\rm ES})^2 = \frac{\hat{\bar{b}}_{\rm cor}^2}{\zeta^2}, \tag{C.12}$$

where

$$\hat{\bar{b}}_{\rm cor} = \frac{\sum_s (\delta_s^2 + \phi^2)^{-1} \delta_s q_s(T_s)}{\sum_s (\delta_s^2 + \phi^2)^{-1}}. \tag{C.13}$$

Note, the quantile transformation functions q_s and q converge to the identity mappings as $n_s \to \infty$ and the asymptotic property of expression (4.9) of main text is preserved. The numerical performance of this correction is demonstrated in appendix D.

To show the corrected version of approximate BF satisfying (C.1) , we note that $\text{ABF}^{\rm ES}$ depends on data $\boldsymbol{Y}$ only through T_s (δ_s depends on genotype data but not $\boldsymbol{Y}$). Further, from **Remark** in appendix A.1, we also notice the approximation becomes an exact BF (for which property (C.1) is guaranteed) if estimated error variance terms $\hat{\sigma}_s^2$'s are replaced by their corresponding true values. When the true error variances are plugged in, under the H_0, T_s's instead follow standard norm distributions. It is therefore sufficient to satisfy property (C.1) by quantile transforming each individual T_s in (4.9) of main text from the t-distribution to standard normal distribution. In essence, the correction can be viewed as a general strategy of providing a better point estimate of residual errors, therefore the similar strategy also likely improves the accuracy of approximate BF when EE model or CEFN model is used.

APPENDIX D: NUMERICAL ACCURACY OF BF EVALUATIONS

In this section, we evaluate the numerical accuracy of various approximation methods for computing the BFs.

We use the dataset from population eQTL study (Stranger et al. (2007)) discussed in section 3.3 of main text for this purpose. For each of the 8,427 genes examined, we select the top associated *cis*-SNP based on the values of $\widehat{\text{BF}}_{\rm av}^{\rm ES}$ and re-calculate the BF directly based on (A.6) using a general adaptive Gaussian quadrature procedure (Note, because of its high computational cost in numerically evaluating multi-dimensional integrals, this numerical recipe does not apply in general practice). We treat these results as the "truth" and make comparison with $\widehat{\text{BF}}_{\rm av}^{\rm ES}$ and $\text{ABF}_{\rm av}^{\rm ES}$ (with and without small sample corrections). Moreover, we convert various numerical results of BFs to log 10 scale and compute Root Mean Squared Errors (RMSE) for each approximation.

The results of the numerical evaluation for the ES model shown in Table 1 and Figure 1. Although the sample sizes in each subgroup are quite small in this dataset (41 Europeans, 59 Asians and 41 Africans), the numerical results of $\widehat{\mathrm{BF}}^{\mathrm{ES}}_{\mathrm{av}}$ are almost identical to the results obtained from the adaptive Gaussian quadrature procedure (RMSE $= 1.2 \times 10^{-4}$ in log 10 scale). As expected, the approximate BF, $\mathrm{ABF}^{\mathrm{ES}}_{\mathrm{av}}$, has the worst numerical performance, mainly due to the small sample sizes in this dataset. Nevertheless, the ranking of the SNPs by $\mathrm{ABF}^{\mathrm{ES}}_{\mathrm{av}}$ is quite consistent with true values (rank correlation $= 0.99$). Figure 1 suggests that under small sample situations, $\mathrm{ABF}^{\mathrm{ES}}_{\mathrm{av}}$ tends to over-evaluate the true value and this over-evaluation can become quite severe when the true values are extremely large. On the other hand, the proposed small sample size correction method seems very effective: with this simple correction, the resulting $\mathrm{A^*BF}^{\mathrm{ES}}_{\mathrm{av}}$ are quite accurate comparing with the true values.

We also perform a similar experiment for the EE model using the same dataset with five levels of $\sqrt{\psi^2 + w^2}$ values: 0.1, 0.2, 0.4, 0.8, 1.6, and seven degrees of heterogeneities characterized by ψ^2/w^2 values: $0, 1/4, 1/2, 1, 2, 4, \infty$, and we assign these 35 grid values equal prior weight. The results are similar with the case in the EE model and shown in Table 2.

	$\log_{10}(\widehat{\mathrm{BF}}^{\mathrm{ES}}_{\mathrm{av}})$	$\log_{10}(\mathrm{ABF}^{\mathrm{ES}}_{\mathrm{av}})$	$\log_{10}(\mathrm{A^*BF}^{\mathrm{ES}}_{\mathrm{av}})$
RMSE	1.2×10^{-4}	4.95	0.14

TABLE 1

Numerical accuracy of three approximations for evaluating BFs under the ES model. $\widehat{\mathrm{BF}}^{\mathrm{ES}}_{\mathrm{av}}$ *is based on the first approximation of Laplace's method discussed in appendix A,* $\mathrm{ABF}^{\mathrm{ES}}_{\mathrm{av}}$ *is computed using expression (4.9) of main text and* $\mathrm{A^*BF}^{\mathrm{ES}}_{\mathrm{av}}$ *is based on (C.11) which is corrected for small sample sizes.*

	$\log_{10}(\widehat{\mathrm{BF}}^{\mathrm{EE}}_{\mathrm{av}})$	$\log_{10}(\mathrm{ABF}^{\mathrm{EE}}_{\mathrm{av}})$	$\log_{10}(\mathrm{A^*BF}^{\mathrm{EE}}_{\mathrm{av}})$
RMSE	4.1×10^{-4}	5.03	0.09

TABLE 2

Numerical accuracy of three approximations for evaluating BFs under the EE model.

REFERENCES

BUTLER, R. (2007). *Saddlepoint Approximations with Applications*, 1st ed. Cambridge University Press.

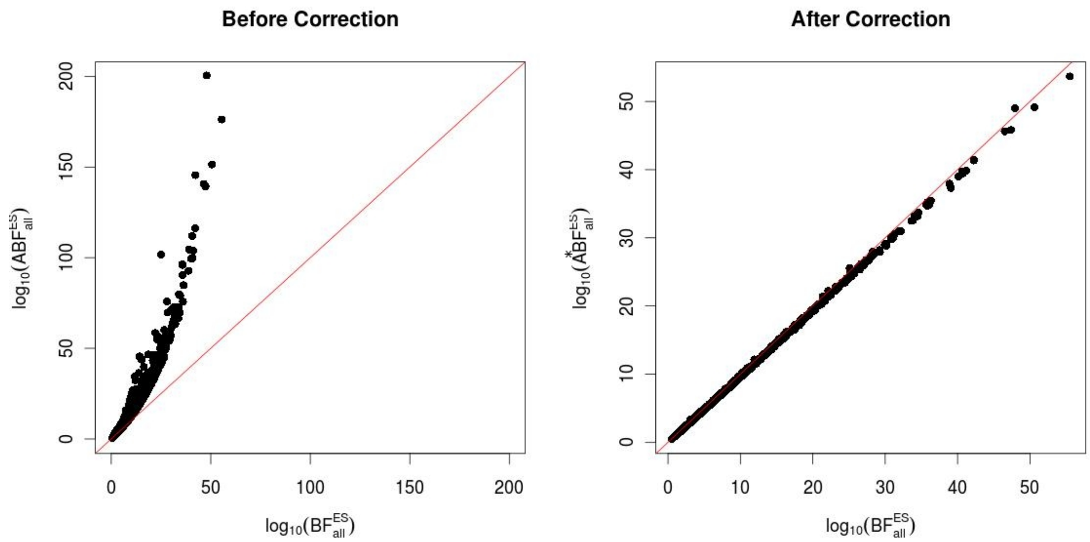

Fig 1. *Comparison of approximate BFs before and after applying small sample size corrections.*

Butler, R. and Wood, A. (2002). Laplace approximations for hypergeometric functions with matrix argument. *The Annals of Statistics* **30** 1155–1177.

Servin, B. and Stephens, M. (2008). Imputation-based Analysis of Association Studies: Candidate Regions and Quantitative Traits. *PLoS Genetics* **3(7)** e114.

Stranger, B., Nica, A., Forrest, M., Dimas, A., Bird, C. et al. (2007). Population genomics of human gene expression. *Nature genetics* **39** 1217–24.

Department of Biostatistics
University of Michigan
1415 Washington Heights,
Ann Arbor, MI 48109
USA
E-mail: xwen@umich.edu

Department of Statistics
and
Department of Human Genetics
University of Chicago
5734 S. University Avenue,
Chicago, IL 60637
USA
E-mail: mstephens@uchicago.edu