Article:
Walters, K. and Yaacob, H. (2023) Bayesian multivariant fine mapping using the Laplace prior. Genetic Epidemiology, 47 (3). pp. 249-260. ISSN: 0741-0395

https://doi.org/10.1002/gepi.22517

Received: 5 July 2022 | Revised: 13 January 2023 | Accepted: 24 January 2023

DOI: 10.1002/gepi.22517

RESEARCH ARTICLE

Genetic Epidemiology

OFFICIAL JOURNAL INTERNATIONAL GENETIC EPIDEMIOLOGY SOCIETY www.geneticepi.org

WILEY

Bayesian multivariant fine mapping using the Laplace prior

Kevin Walters[1] | **Hannuun Yaacob**[1,2]

[1]School of Mathematics and Statistics, University of Sheffield, Sheffield, UK

[2]Department of Economics and Applied Statistics, Faculty of Business and Economics, Universiti Malaya, Kuala Lumpur, Malaysia

Correspondence
Kevin Walters, School of Mathematics and Statistics, University of Sheffield, Sheffield, UK.
Email: k.walters@sheffield.ac.uk

Funding information
Ministry of Higher Education, Malaysia; Universiti Malaya

Abstract

Currently, the only effect size prior that is routinely implemented in a Bayesian fine-mapping multi-single-nucleotide polymorphism (SNP) analysis is the Gaussian prior. Here, we show how the Laplace prior can be deployed in Bayesian multi-SNP fine mapping studies. We compare the ranking performance of the posterior inclusion probability (PIP) using a Laplace prior with the ranking performance of the corresponding Gaussian prior and FINEMAP. Our results indicate that, for the simulation scenarios we consider here, the Laplace prior can lead to higher PIPs than either the Gaussian prior or FINEMAP, particularly for moderately sized fine-mapping studies. The Laplace prior also appears to have better worst-case scenario properties. We reanalyse the iCOGS case–control data from the CASP8 region on Chromosome 2. Even though this study has a total sample size of nearly 90,000 individuals, there are still some differences in the top few ranked SNPs if the Laplace prior is used rather than the Gaussian prior. R code to implement the Laplace (and Gaussian) prior is available at https://github.com/Kevin-walters/lapmapr.

KEYWORDS

Bayesian, fine mapping, Laplace prior, multi-SNP

1 | INTRODUCTION

The Gaussian distribution is the most frequently used effect size (log odds ratio) prior in Bayesian fine-mapping studies (Benner et al., 2016; Chen et al., 2015; Spencer et al., 2016; Wakefield, 2009). This distribution is deemed appropriate as an effect size prior because it is symmetric around zero and has quickly decaying tails. The appeal of the Gaussian prior is that, when used alongside a Gaussian likelihood (Wakefield, 2009), it yields a Gaussian posterior distribution for the effect size. As a result of the tractability of the marginal likelihood, Bayes factors are easily calculated with a zero mean Gaussian effect size prior. In the univariate case, this gives the so-called Wakefield Bayes factor or other tractable Bayes factor that use a hierarchical Gaussian prior (Spencer et al., 2015). For a review of genetic fine-mapping that includes Bayesian approaches, see Hutchinson et al. (2020).

The Gaussian distribution is not the only distribution that has been considered for the effect size prior in univariate analyses. The *t* distribution (Marchini & Howie, 2010), the normal-gamma distribution (Alenazi et al., 2019; Boggis et al., 2016) and the Laplace distribution

(Hoggart et al., 2008; Walters et al., 2021) have all been considered. Walters et al. (2019) assessed the fit of the Gaussian and Laplace priors using Bayesian model selection. They analysed the estimated effect sizes of a large number of independent breast cancer genome-wide association study (GWAS) top hits and found that the Laplace prior gave a higher model posterior probability than the Gaussian prior. Following on from the findings of Walters et al. (2019), Walters et al. (2021) explored the performance of the Laplace prior (compared to the Gaussian prior) in Bayesian fine mapping. They only considered univariate analyses and examined the causal SNP ranks via receiver-operating characteristic (ROC) curves and credible set sizes. They found that there are many scenarios where the Laplace prior outperformed the Gaussian prior in terms of causal SNP rank, although the differences were often not large. The downside of the Laplace prior was that it lead to larger credible set sizes than the Gaussian prior. To the best of our knowledge, no consideration has been given to the use of non-Gaussian priors in a fully Bayesian multi-SNP fine mapping analysis. This may be because of the lack of tractability of the posterior distribution when non-Gaussian priors are used. Calculating the marginal likelihood with a Laplace prior requires either numerical integration or a Monte Carlo approach, which will make it slower than implementing the Gaussian prior. The question of interest is to what extent this increase in computing time is made up for by gains in fine-mapping performance.

Several authors have implemented multi-SNP fine mapping approaches using the Gaussian prior (Benner et al., 2016; Bottolo, 2010; Chen et al., 2015). There are two main approaches used in multi-SNP fine mapping. The first is to enumerate over all possible models, usually with a restriction on the maximum number of causal SNPs allowed in the model (Chen et al., 2015). The second is to use a stochastic search algorithm to explore the model space (Benner et al., 2016; Bottolo, 2010). The latter approach gives an approximation to the former, at considerable computational savings, making it possible to consider a larger number of causal SNPs than is possible with enumerative methods. We consider modest numbers of causal SNPs in the region to be fine-mapped and so the exhaustive approach is an acceptable approach.

We compare the performance of both Gaussian and Laplace priors in multi-SNP Bayesian fine-mapping via ranking of SNP marginal posterior inclusion probabilities (PIPs). The PIP was recommended by Chen et al. (2015) as a suitable measure because selecting SNPs using marginal PIPs maximises the expected number of causal SNPs for a fixed number of SNPs selected. We provide detailed calculations of the marginal likelihood needed to obtain the model posterior probability with the Laplace prior. The calculations for the Gaussian prior are similar to those in Chen et al. (2015). We also compare our results with those of FINEMAP (Benner et al., 2016). Benner et al. (2016) showed that FINEMAP yielded single SNP PIPs that were generally very similar to those of CAVIARBF Chen et al. (2015). The results of CAVIARBF should be almost identical to our Gaussian prior approach because they both enumerate over all possible models (up to a maximum model size) with the same Gaussian prior. For these reasons, we do not include the CAVIARBF performance.

2 | MATERIALS AND METHODS

We need to calculate the posterior probability for a specific model for both the Gaussian and Laplace prior case. We start by detailing all the components that are common to the two priors and then show how to calculate the posterior probability for a specific model for the Laplace case. We associate an indicator vector, $\mathbf{c}$, with each model, $\boldsymbol{M_c}$, where $c_j = 1$ if SNP j is causal and $c_j = 0$ otherwise. With p SNPs there are 2^p possible models. Let $d = ||\boldsymbol{c}||_1$, where $||\cdot||_1$ is the L^1 norm. Thus, d is the number of causal SNPs in model $\boldsymbol{M_c}$. For computational reasons, we restrict the model space to at most K causal SNPs. This is equivalent to those models $\boldsymbol{M_c}$ with $d \leq K$. For notational simplicity, we assume that, given $\boldsymbol{M_c}$, the SNPs have been reordered so that the d causal SNPs are labelled SNPs 1 to d (this allows us to use simpler subscripts). Let $\hat{\boldsymbol{\beta}}$ represent the column vector of length p of estimates of the effect sizes. Then, the posterior probability of model $\boldsymbol{M_c}$ is

$$P(\boldsymbol{M_c}|\hat{\boldsymbol{\beta}}) = \frac{P(\boldsymbol{M_c})\, f(\hat{\boldsymbol{\beta}}|\boldsymbol{M_c})}{\sum_{\boldsymbol{M_c}} P(\boldsymbol{M_c})\, f(\hat{\boldsymbol{\beta}}|\boldsymbol{M_c})}, \quad (1)$$

where $P(\boldsymbol{M_c})$ is the prior probability of the model and $f(\hat{\boldsymbol{\beta}}|\boldsymbol{M_c})$ is the model marginal likelihood. Typically, we would include an intercept in our model. Wakefield (2009) showed that this was not necessary in a univariate analysis. He showed that it was possible to reparameterise the model so that the joint distribution of the likelihood factorised into a term involving the intercept and a term involving the effect size parameter. As a consequence, the marginal likelihood for each model considered is multiplied by the same constant term (coming from integrating the intercept out of the joint distribution). As a result, these constant terms cancel in Equation (1). In the Appendix, we show that this factorisation property extends to the multivariate case. We

also show that the variance of $\hat{\boldsymbol{\beta}}$ in the model where we include an intercept is the same as the variance of $\hat{\boldsymbol{\beta}}$ in the reparameterised model. Essentially the Appendix shows that we can ignore the intercept in multivariate models when the interest is in the marginal likelihoods.

2.1 | Prior probability of the model

When analysing the simulated data we assume that each of the p SNPs is causal with probability q. The prior probability of model $\boldsymbol{M_c}$ with d causal SNPs is therefore $q^d\ (1-q)^{p-d}$. We also assume that the number of causal SNPs is K or fewer ($d \le K$). For the simulated data analyses, we assume that there is at least one causal SNP because regions are fine-mapped because of the presence of a statistically significant GWAS association. Thus, the prior probability for model $\boldsymbol{M_c}$ is given by

$$P(\boldsymbol{M_c}) = \frac{q^d\ (1-q)^{p-d}}{\sum_{k=0}^{K}\binom{p}{k} q^k\ (1-q)^{p-k}} \quad \text{for } d \le K, \tag{2}$$

where the denominator acts as a normalising constant. A good starting point for specifying q is to equate a guess at the prior expected number of causal SNPs with pq, the binomial expected value. We suggest that the user undertakes a sensitivity analysis if they have any doubts about an appropriate value of q or want to assess the sensitivity of the results to the value of q. Note that an alternative approach is to specify an arbitrary truncated discrete prior probability distribution on the number of causal SNPs and then calculate the prior probability of model $\boldsymbol{M_c}$ by assuming that each model with d causal SNPs is equally likely, for $d \le K$.

2.2 | The likelihood

Let $\boldsymbol{\beta_c}$ be a column vector of the effect sizes of the d causal SNPs in model $\boldsymbol{M_c}$. Let $\boldsymbol{\beta_N}$ be a column vector of $p-d$ zeroes representing the effect sizes of the noncausal SNPs. In fine-mapping we have large samples of case–control data, so we assume that the likelihood is distributed as

$$\begin{pmatrix} \hat{\boldsymbol{\beta}}_{\boldsymbol{c}}|\boldsymbol{\beta_c} \\ \hat{\boldsymbol{\beta}}_{\boldsymbol{N}}|\boldsymbol{\beta_N} \end{pmatrix} \sim N\left(\begin{pmatrix} \boldsymbol{\beta_c} \\ \mathbf{0} \end{pmatrix}, \Omega\right),$$

where Ω is the $p \times p$ covariance matrix of $\hat{\boldsymbol{\beta}}$. We let

$$\Omega^{-1} = \begin{pmatrix} \Sigma_c & \Sigma \\ \Sigma^T & \Sigma_N \end{pmatrix}, \tag{3}$$

where Σ_c is a $d \times d$ covariance matrix of the causal SNPs, Σ_N is a $(p-d)\times(p-d)$ covariance matrix of the noncausal SNPs and Σ is the $d \times (p-d)$ matrix of covariances between causal and noncausal SNPs. The likelihood probability density function (pdf) can be written as

$$f(\hat{\boldsymbol{\beta}}|\boldsymbol{\beta}) = ((2\pi)^p|\Omega|)^{-1/2} \exp\left(-\frac{1}{2}\left[\boldsymbol{\beta_c}^T\ \Sigma_c\ \boldsymbol{\beta_c} - 2\left(\hat{\boldsymbol{\beta}}_{\boldsymbol{c}}^T\ \Sigma_c\ +\ \hat{\boldsymbol{\beta}}_{\boldsymbol{N}}^T\ \Sigma^T\right)\boldsymbol{\beta_c} + \delta\right]\right), \tag{4}$$

where $\delta = \hat{\boldsymbol{\beta}}^T \Omega^{-1} \hat{\boldsymbol{\beta}}$.

2.3 | Effect size priors

Given the model, the prior effect sizes are independent. Noncausal SNPs necessarily have effect sizes of zero. The joint prior of the d causal SNP effect sizes is given by $f(\boldsymbol{\beta_c}|M_{\mathbf{c}}) = \prod_{i=1}^{d} f(\beta_j)$, where $f(\beta_j)$ is either the pdf of the $N(0,w)$ distribution or the pdf of the Laplace distribution with rate parameter λ, denoted by La(λ). For the La(λ) prior this gives

$$f(\boldsymbol{\beta_c}|M_{\mathbf{c}}) = \left(\frac{\lambda}{2}\right)^d \exp(-\lambda \boldsymbol{\psi}^T \boldsymbol{\beta_c}), \tag{5}$$

where $\boldsymbol{\psi}$ is a length d column vector with jth element given by

$$\psi_j = \begin{cases} -1, & \text{if } \beta_j < 0, \\ 1, & \text{if } \beta_j \ge 0. \end{cases} \tag{6}$$

2.4 | Model marginal likelihood with a Laplace prior

The marginal likelihood is evaluated by integrating out $\boldsymbol{\beta_c}$. We have

$$f(\hat{\boldsymbol{\beta}}|M_{\mathbf{c}}) = \int_{\boldsymbol{\beta_c}} f(\boldsymbol{\beta_c}\ |M_{\mathbf{c}})\, f(\hat{\boldsymbol{\beta}}\ |\ \boldsymbol{\beta}) d\boldsymbol{\beta_c}. \tag{7}$$

Using the Laplace prior in Equation (5) and the likelihood in Equation (4), the integrand in Equation (7) is

$$\left(\frac{\lambda}{2}\right)^d ((2\pi)^p\ |\Omega|)^{-1/2} \exp\left(-\frac{1}{2}\left[\boldsymbol{\beta}_{\boldsymbol{c}}^T \Sigma_c \boldsymbol{\beta}_{\boldsymbol{c}} - 2\left(\hat{\boldsymbol{\beta}}_{\boldsymbol{c}}^T \Sigma_c + \hat{\boldsymbol{\beta}}_{\boldsymbol{N}}^T \Sigma^T - \lambda \boldsymbol{\psi}^T\right)\boldsymbol{\beta}_{\boldsymbol{c}} + \delta\right]\right) \tag{8}$$

$$= \left(\frac{\lambda}{2}\right)^d ((2\pi)^{p-d} |\Omega \Sigma_c|)^{-\frac{1}{2}} \exp\left(\frac{-\phi}{2}\right) \sqrt{(2\pi)^{-d}\ |\Sigma_c|} \exp\left(-\frac{1}{2}(\boldsymbol{\beta}_{\boldsymbol{c}} - \boldsymbol{\mu})^T\ \Sigma_c\ (\boldsymbol{\beta}_{\boldsymbol{c}} - \boldsymbol{\mu})\right), \tag{9}$$

where $\boldsymbol{\mu} = \Sigma_c^{-1}\boldsymbol{\alpha}^T$, $\phi = \delta - \boldsymbol{\alpha}\Sigma_c^{-1}\boldsymbol{\alpha}^T$ and $\boldsymbol{\alpha} = \hat{\boldsymbol{\beta}}_{\boldsymbol{c}}^T \Sigma_c + \hat{\boldsymbol{\beta}}_{\boldsymbol{N}}^T \Sigma^T - \lambda \boldsymbol{\psi}^T$. The integral in Equation (7) is over $\mathbb{R}^d$ but the integrand here is a function of $\boldsymbol{\psi}$, which depends on the signs of $\beta_1, ..., \beta_d$. As a consequence, we need to break the integral down over the 2^d combinations of $\boldsymbol{\psi}$ and use appropriate limits. For $\psi_j = 1$, the integration is over $\beta_j \in [0, \infty)$, while for $\psi_j = -1$, the integration is over $(-\infty, 0)$. Thus, the marginal likelihood is

$$f(\hat{\boldsymbol{\beta}}|\boldsymbol{M}_{\boldsymbol{c}}) = \left(\frac{\lambda}{2}\right)^d [(2\pi)^{p-d}|\Omega\Sigma_c|]^{-\frac{1}{2}} \sum_{\boldsymbol{\psi}\in\{-1,1\}^d} \exp\left(-\frac{\phi}{2}\right) \int_{\beta_1\in\theta_1} \cdots \int_{\beta_d\in\theta_d} \gamma d\beta_1 \cdots d\beta_d, \tag{10}$$

where γ is the pdf of a multivariate normal distribution with mean vector $\boldsymbol{\mu}$ and covariance matrix Σ_c^{-1} and

$$\theta_j = \begin{cases} (-\infty, 0), & \text{if } \psi_j = -1, \\ [0, \infty), & \text{if } \psi_j = 1. \end{cases}$$

There is no closed form expression for the integral in Equation (10) but it can be evaluated using the `pmvnorm` function in the `mvtnorm` package (Genz & Bretz, 2009) in R (R Core Team, 2021). Because some of our integral limits are not finite, we chose to use the Monte Carlo approach to evaluating the integrals as described in Genz and Bretz (2009). Mi et al. (2009) examined the accuracy of the probability approximations used in the `mvtnorm` package. For the cases they considered (where the truncated multidimensional integral values were known), the approximations were more than accurate enough for the number of decimal places of accuracy we considered necessary for the PIP calculations. The marginal PIP for SNP i is then the sum of the posterior probability of all models that include SNP i. The Gaussian prior leads to a tractable closed-form expression for the posterior model probability. The calculations are similar to the Laplace case and are presented in the Appendix.

2.5 | Choice of K

K limits the maximum number of causal SNPs allowed and the user needs to specify K in our approach. Keeping K small, while considering all reasonably plausible models, will keep computing time down. One approach to deciding on a value for K uses the marginal likelihood with K causal SNPs or fewer. The marginal likelihood is given by

$$f_K(\hat{\boldsymbol{\beta}}) = \sum_{\boldsymbol{M}_{\boldsymbol{c}}\in Q} f(\hat{\boldsymbol{\beta}}|\boldsymbol{M}_{\boldsymbol{c}}) P(\boldsymbol{M}_{\boldsymbol{c}}), \tag{11}$$

where $P(\boldsymbol{M}_{\boldsymbol{c}})$ and $f(\hat{\boldsymbol{\beta}}|\boldsymbol{M}_{\boldsymbol{c}})$ are given in Equations (2) and (10), respectively, and Q is the set of all models with no more than K causal SNPs, that is, $Q = \{\boldsymbol{M}_{\boldsymbol{c}} : ||\boldsymbol{c}||_1 \le K\}$. The method is to start with $K = 1$ and compare $f_2(\hat{\boldsymbol{\beta}})/f_1(\hat{\boldsymbol{\beta}})$. This ratio must be at least one because allowing larger numbers of causal SNPs must increase the marginal likelihood. If it is much larger than 1, then increment K by 1, calculate $f_3(\hat{\boldsymbol{\beta}})/f_2(\hat{\boldsymbol{\beta}})$, and make the same assessment. Stop incrementing K once the ratio is close enough to 1 (subjectively defined). This method stops increasing K once the multiplicative increase in the marginal likelihood is small enough. We give an example of implementing this for simulated data Scenario 1 to illustrate typical values of these ratios of marginal likelihoods.

2.6 | Gaussian and Laplace priors and FINEMAP settings

We assume, a priori, that the effect size of causal SNPs follows a $N(0, 0.2^2)$ distribution, which is a common assumption in fine-mapping studies. To ensure a fair comparison between the Gaussian and Laplace prior, we equate the prior variances which leads to $\lambda = \sqrt{2/w}$, where $w = 0.2^2$ is the Gaussian prior effect size variance. This value of w gives $\lambda = 7.1$. For the prior distribution on the number of causal SNPs, we set $q = 0.04$ and $K = 2$, which gives prior probabilities of one and two causal SNPs of 0.37 and 0.63, respectively (assuming 82 SNPs, which is representative of the number of SNPs included in our simulations).

We compared the ranking performance of the shotgun search approach of FINEMAP (Benner et al., 2016) with those of the Laplace and Gaussian priors. To ensure identical effect size prior variances in all analyses, we specified a $N(0, 0.2^2)$ effect size prior in the FINEMAP analysis. The prior probabilities of the number of causal SNPs in the FINEMAP analysis were the same as for the

Laplace and Gaussian prior analysis. All other FINEMAP settings were set to their default value.

2.7 | Simulation details

We simulated haplotypes from the CASP8 region between base pairs 201,666,128 and 201,866,128 of the Hg19 build of chromosome 2 using Hapgen2 (Su et al., 2011). We used the European haplotypes of the August 2010 release of the 1000 Genomes Project as the reference data. We simulated haplotypes containing two causal SNPs with different minor allele frequencies (MAFs), effect sizes and levels of linkage disequilibrium (LD). For each simulation scenario we considered, we simulated 50 case–control data sets. Hapgen2 identified 412 SNPs in this region. We removed SNPs with MAF less than 0.01 and also retained just one SNP from each cluster of SNPs in high LD ($r^2 > 0.99$), making sure to retain the causal SNP if it was in a high LD cluster. Finally, we removed all SNPs with a univariate Bayes factor of less than 1 in all 50 data sets. This left between 77 and 87 SNPs in the simulated data sets. We considered six scenarios. In all scenarios, we fixed the odds ratio and MAF of the first (second) causal SNP to be 1.15 (1.25) and 0.3 (0.09), respectively. In all scenarios, we simulated equal numbers of case and control haplotypes. We considered total sample sizes of 8000, 14,000 and 18,000 individuals. These give univariate power to reject the null hypothesis of no association of approximately 8%, 46% and 73%, respectively, for each causal SNP. VanLiere and Rosenberg (2008) showed that if the MAFs of two SNPs satisfy $MAF_1 > MAF_2$, then the maximum r^2 value between two SNPs is given by

$$r^2_{\max} = \frac{(1 - MAF_1) \times MAF_2}{(1 - MAF_2) \times MAF_1}. \quad (12)$$

We report $r^2/r^2_{\max}$ along with the other simulation parameters in Table 1

2.8 | SNP prioritisation

We used ROC curves and area under the ROC curves (AUC) to compare the performance of the PIP as a ranking statistic. We used the ROCR package (Sing et al., 2005) to plot the curves and calculate the AUCs. Because we are really interested in the true-positive rates (TPRs) at low false-positive rates (FPRs), we report the partial AUC on FPRs <0.05 given as an integer percentage. The partial AUC for FPRs less than some maximum FPR value can be calculated using the ROCR package by setting the `fpr.stop` argument to this maximum FPR. For example, a reported partial AUC of 80 means that the AUC for FPR <0.05 was 0.04 (compared to a maximum possible partial AUC in this region of 0.05). We chose to use vertical averaging (Fawcett, 2006) to aggregate results over the multiple ROC curves. For a given FPR, vertical averaging averages TPRs across ROC curves and hence gives an average TPR at each FPR. For completeness, we report the AUCs on FPRs ≤1 in Supporting Information: File 1 and show the corresponding ROC curves.

As well as comparing the average ranking performance across data sets by scenario, we were also interested in comparing the tails of the PIP sampling distribution for the Laplace and Gaussian priors for the simulated data in Scenario 1 in Table 1. The question of interest is how the worst-case scenarios across simulations differ by type of prior. We report the actual PIPs and the PIP ranks for both causal SNPs for both priors, where a rank of one corresponds to the highest PIP. In each analysis, we set the prior probability that a SNP is causal at 0.04, regardless of K. We report the prior and posterior expected number of causal SNPs for each value of K for both priors.

Finally, we illustrate how to choose the user-specified value K for the Gaussian prior in simulated scenario 1.

TABLE 1 Simulation scenarios used to compare the Gaussian and Laplace priors and FINEMAP

Scenario number	Causal SNP1 MAF	Causal SNP2 MAF	Total sample size (1000s)	$r^2/r^2_{\max}$
1	0.33	0.08	8	0.03
2	0.33	0.08	14	0.03
3	0.33	0.08	18	0.03
4	0.28	0.10	8	0.92
5	0.28	0.10	14	0.92
6	0.28	0.10	18	0.92

Abbreviations: MAF, minor allele frequenc; SNP, single-nulceotide polymorphism.

2.9 | iCOGS analysis

We compared the PPIs using the Gaussian and Laplace priors in fine-mapping the CASP8 region between base positions 201,500,074 and 202,569,992 on Chromosome 2. The data comes from a custom-built iCOGS array developed by the Collaborative Oncological Gene-Environment Study (COGS) Consortium (Michailidou et al., 2013). Of the 1733 variants, 501 were genotyped and 1232 were imputed using Impute2 (Marchini & Howie, 2010); they were genotyped in 42,600 controls and 46,450 breast cancer cases. We report the top 10 PIP-ranked SNPs for $K = 1, 2$ and 3 for both the Gaussian and Laplace priors. We applied a similar filtering to that used in the analysis of the simulated data. We removed all iCOGS SNPs that had a MAF less than 0.05 and all SNPs that had a Wakefield Bayes factor less than 2. We choose a more stringent Bayes factor threshold in the iCOGS analysis than in the simulated data analysis simply to reduce the number of eligible SNPs. This left 90 SNPs in the CASP8 region. We provide the SNP number from 1 to 1733 (rather than 1 to 90) for the top 10 ranked SNPs. This is for ease of comparison with Walters et al. (2021). The prior probabilities of the number of causal SNPs used in this analysis are provided in Table 2.

3 | RESULTS

3.1 | SNP prioritisation

Figure 1 compares the performance of the Laplace prior, the Gaussian prior and FINEMAP as the sample size varies for the scenarios described in Table 1. The partial AUCs are also provided in the figure legends. As might be expected, the performance of the Laplace and Gaussian priors equalises as the sample size increases. In all six scenarios, the Laplace prior gave the highest partial AUC, generally followed by the Gaussian prior. FINEMAP was not the best performing in any of the scenarios we considered. FINEMAP performed particularly badly in the high LD case. This might reflect the fact that FINEMAP is selecting SNPs in high LD with the causal SNPs in its shotgun search strategy, rather than the causal SNPs themselves. If we redefined the target SNPs as the causal SNPs and those in very high LD with them, we would expect the performance of FINEMAP to improve considerably. These ROC curves show that the Laplace prior can give higher PIP ranks than both the Gaussian prior and FINEMAP across a range of simulation scenarios.

Table 3 also shows the posterior expected number of causal SNPs by K, along with the prior expected number of causal SNPs. We see that as the prior expected number increases, so does the posterior number of causal SNPs, although the posterior is always higher than the prior. There is not much difference in the posterior number between the Laplace and Gaussian priors. Table 3 also shows values from the tails of the PIP sampling distribution across simulations (and also from the sampling distribution of ranks). Here, we are interested in comparing low percentiles of the PIP distribution that correspond to poor performance. For each causal SNP, the 10th percentile of the 50 PIP values is never less for the Laplace prior than for the Gaussian prior (although the differences are relatively modest), except for the case $K = 1$ for the rare causal SNP (causal SNP 2). This exception is not identifiable given the number of decimal places presented in the table but can be seen in the 90th percentile of the ranks. The most stark difference between the two priors occurs in the worst rank seen across the 50 simulations. These can be quite marked, particularly for higher values of K, but even at $K = 2$, the worst rank for the common causal SNP is 8 for the Laplace prior and 10 for the Gaussian. These results show that not only can the Laplace prior perform better, on average, than the Gaussian prior in terms of causal SNP ranks, but also gives the causal SNPs higher ranks in the event of an extreme set of sampled genotypes.

Figure 2 shows the performance of the Laplace prior by maximum number of causal SNPs allowed in the model (K) for the Scenario 1 in Table 1. As might be expected, there is a marked improvement in performance as K increases from one to two, with an increase in partial AUC of nearly 50%. As K increases further the improvement in AUC becomes increasingly modest. Table 3 shows the median PIPs for the two causal SNPs. We see a similar levelling-off of the median PIP beyond $k = 2$ for the common causal SNP but this effect is less apparent for the rarer causal SNP, where there are still improvements up to and including $K = 4$.

Table 4 presents the prior probabilities of the number of causal SNPs and multiplicative increases in the

TABLE 2 Prior probability distributions of the number of causal SNPs implemented in the iCOGS analysis.

Maximum number of causal SNPs allowed	Prior probability of k causal SNPs			
	$k = 0$	$k = 1$	$k = 2$	$k = 3$
1	0.50	0.50	0	0
2	0.45	0.45	0.10	0
3	0.35	0.35	0.20	0.10

Abbreviation: SNP, single-nuceotide polymorphism.

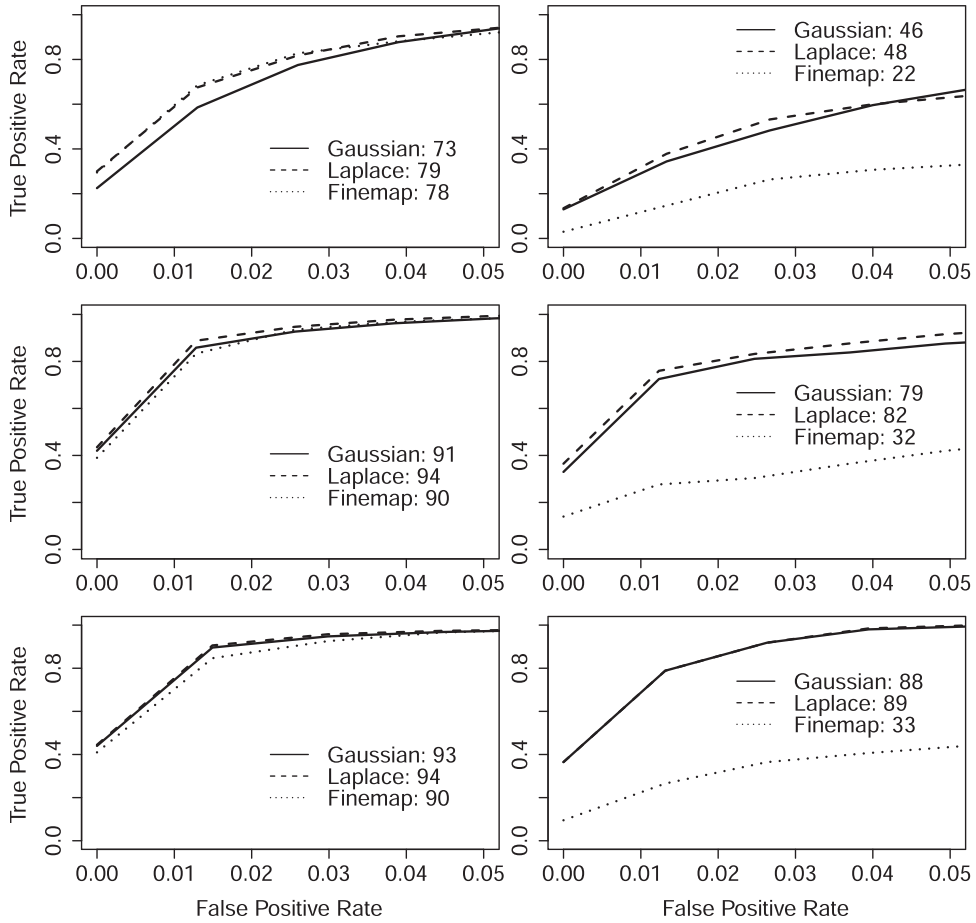

FIGURE 1 Receiver-operating characteristic (ROC) curves comparing the ranking performance of the posterior inclusion probability for the Laplace and Gaussian prior and FINEMAP. Each ROC curve shows the average true-positive rate (TPR) for a given false-positive rate (FPR), averaged over 50 data sets, simulated from Hapgen2. The simulation parameters are given in Table 1. We allowed models with a maximum of two causal SNPs. The plots on the left are for Scenarios 1–3 (top left is Scenario 1) and the plots on the right are for Scenarios 4–6 (top right is Scenario 4). Partial area under the ROC curves are given in the figure legends.

marginal likelihoods for different K as a function of the prior probability that each SNP is causal (q) for the first simulated data set of Scenario 1 (see Table 1) with a $N(0, 0.04)$ effect size prior. As q increases, we observe greater probability of 3, 4 and 5 causal SNPs. The sensitivity of the marginal likelihoods ratios to q can also be observed. For $q = 0.04$ we can see that there are large multiplicative increases in the marginal likelihood as K increases. For $q = 0.04$ the marginal likelihood with a maximum of five causal SNPs is 1.31 times the marginal likelihood with a maximum of four causal SNPs. It is difficult to interpret these numbers directly, but by using Figure 2 we can see how the various marginal likelihood ratios in Table 4 relate to the ROC curves for simulated Scenario 1 data (where $q = 0.04$). Increasing K from 1 to 2 increases the partial AUC considerably (the marginal likelihood ratio for this change is 4.93). Changing from $K = 3$ to 4 yields a small increase in partial AUC (the marginal likelihood ratio for this change is 1.64). So, it seems that multiplicative increases in the marginal likelihood of around 1.5 or less might not warrant increasing K. Note that with 80 SNPs, the code took less than a minute to run for $K = 3$ but 100 min for $K = 5$.

3.2 | iCOGS analysis

Table 5 compares the top PIP-ranked iCOGS SNPs for the Gaussian and Laplace prior by maximum number of causal SNPs allowed. The table shows that the results for the two priors are very similar. It is not until the sixth ranked SNP that there is any disagreement between the two analyses. This strong similarity is not unexpected because the sample size analysed here is very large (a total sample size of nearly 90,000) compared to those used in the simulations. There is a marked difference in the top two ranked SNPs as the maximum allowed number of causal SNPs increases from 1 to 2 for both the

TABLE 3 Percentiles of the distributions of the PIP and the PIP ranks for the two causal SNPs in Scenario 1 in Table 1.

		$K = 1$		$K = 2$		$K = 3$		$K = 4$	
		L	G	L	G	L	G	L	G
	Prior expected number of causal SNPs	1.00	1.00	1.61	1.61	2.15	2.15	2.58	2.58
	Posterior expected number of causal SNPs	1.00	1.00	1.77	1.76	2.39	2.42	2.86	2.95
Causal SNP 1	PIP 50th percentile	0.36	0.36	0.50	0.47	0.52	0.48	0.53	0.49
	PIP 10th percentile	0.07	0.06	0.12	0.11	0.13	0.11	0.14	0.11
	PIP rank 50th percentile	1.0	1.5	1.0	2.0	1.0	1.5	1.5	1.5
	PIP rank 90th percentile	3.0	3.0	3.1	4.0	4.0	4.0	4.0	5.0
	PIP rank maximum	6	6	8	10	10	12	13	19
Causal SNP 2	PIP 50th percentile	0.00	0.00	0.26	0.21	0.37	0.31	0.42	0.38
	PIP 10th percentile	0.00	0.00	0.06	0.05	0.09	0.07	0.11	0.09
	PIP rank 50th percentile	5	5	2	3	2	2	2	2
	PIP rank 90th percentile	13.0	10.1	5.0	5.0	4.0	5.0	4.0	5.0
	PIP rank maximum	25	25	27	38	26	39	26	37

Note: Results are presented according to the maximum number of causal SNPs allowed in the model (K) and prior type (L is Laplace, G is Gaussian). Prior and posterior expected numbers of causal SNPs are also presented.

Abbreviations: PIP, posterior inclusion probability; SNP, single-nuceotide polymorphism.

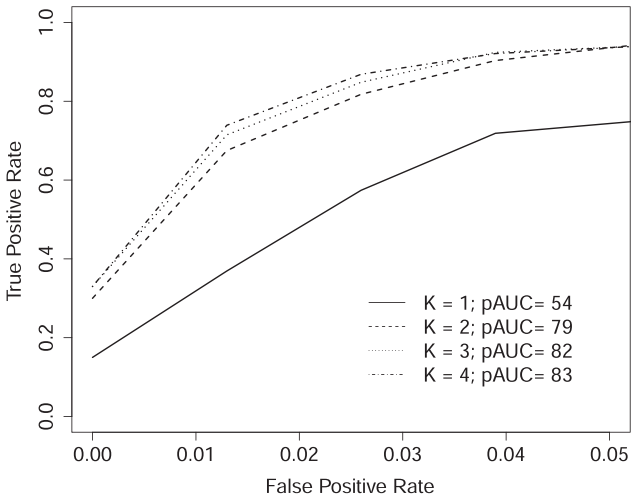

FIGURE 2 Receiver-operating characteristic (ROC) curves comparing the ranking performance of the posterior inclusion probability for the Laplace prior for Scenario 1 in Table 1 by maximum number of causal SNPs allowed in the model (K). Each ROC curve shows the average true-positive rate (TPR) for a given false-positive rate (FPR), averaged over 50 data sets, simulated from Hapgen2. Partial area under the ROC curves are given in the figure legends.

Gaussian and Laplace priors. SNP 1639 is ranked 18 for $K = 1$ for both priors with PIPs less than 5×10^{-7}. For $K = 2$, SNP 1639 is ranked 2 with PIPs greater than 0.09 for both priors. For $K = 3$, the PIPs for SNP 1639 jump to 0.094 and 0.167 for the Laplace and Gaussian priors, respectively. SNP 31 is clearly the top ranked SNP in all six analyses in Table 5 but there are differences in the PIPs of this SNP across the six analyses. For the Gaussian prior the PIPs vary between 0.650 and 0.733, whereas for the Laplace prior, they vary between 0.696 and 0.715 across the three values of K. The Laplace prior therefore yields PIPs for the top ranking SNP with a smaller range than for the Gaussian prior. Note that SNP 31 was also the top ranked SNP in the univariate iCOGs analysis undertaken in Walters et al. (2021) which used two different values of λ in the Laplace prior.

4 | DISCUSSION

The results in this paper concur somewhat with the results of the univariate Laplace analysis undertaken in Walters et al. (2021): utilising a Laplace prior in a Bayesian fine-mapping multi-SNP analysis can lead to better causal SNP prioritisation than using a Gaussian prior, although it depends on the actual causal SNP effect sizes and allele frequencies. The results demonstrate that, as expected, differences in causal SNP ranks using different priors decrease with sample size.

A clear drawback of the Laplace prior approach is that it leads to an increase in computing time and realistically limits the number of causal SNPs considered to three for most analyses, as is the case for most of the methods that enumerate over all possible models. There

TABLE 4 Prior probabilities of the number of causal SNPs and multiplicative increase in the marginal likelihood as K increments by 1 by the prior probability that each SNP is causal (q) for the first simulated data set of Scenario 1 in Table 1 with a $N(0, 0.04)$ effect size prior.

	Prior probability of d causal SNPs								
q	$d = 1$	$d = 2$	$d = 3$	$d = 4$	$d = 5$	$f_2(\hat{\beta})/f_1(\hat{\beta})$	$f_3(\hat{\beta})/f_2(\hat{\beta})$	$f_4(\hat{\beta})/f_3(\hat{\beta})$	$f_5(\hat{\beta})/f_4(\hat{\beta})$
0.04	0.17	0.26	0.26	0.20	0.11	4.93	2.40	1.64	1.31
0.03	0.27	0.31	0.23	0.13	0.06	3.92	1.97	1.40	1.17
0.02	0.43	0.33	0.16	0.06	0.02	2.92	1.57	1.20	1.06
0.01	0.67	0.25	0.06	0.01	0.01	1.95	1.21	1.05	1.01

Abbreviation: SNP, single-nuceotide polymorphism.

TABLE 5 SNP number (from 1 to 1733) of the top 10 ranked iCOGS SNPs using Gaussian and Laplace priors by maximum number of causal SNPs allowed (K)

	Gaussian prior			Laplace prior		
Rank	$K = 1$	$K = 2$	$K = 3$	$K = 1$	$K = 2$	$K = 3$
1	31	31	31	31	31	31
2	2	1639	1639	2	1639	1639
3	1	2	602	1	2	602
4	3	1	2	3	1	2
5	16	3	1	16	3	1
6	24	602	1056	24	602	3
7	7	16	3	7	16	681
8	29	811	1038	29	681	1671
9	27	816	1043	27	811	1056
10	602	812	681	602	816	1656

Abbreviation: SNP, single-nuceotide polymorphism.

are several ways that the computing time could be reduced. The first approach would be to implement parallel computing where the calculations for the different models are partitioned over different PC cores. This would approximately reduce the computing time pro rata to the number of cores. A second approach is to attempt to restrict the number of regions of $\boldsymbol{R}^k$ that need to be integrated over. There are 2^K different regions, where K is the maximum number of causal SNPs allowed in the model. Some of these will contribute very little to the model marginal likelihood. It seems that it should be possible to use information from the eigenvalues/vectors of the $\hat{\boldsymbol{\beta}}$ covariance matrix, along with the estimated SNP effect sizes, to inform which regions could be be ignored. A third approach is to only consider sets of $k + 1$ SNPs for SNPs, which appear in at least one k-SNP model with high posterior probability. This has the potential for large computational savings but the probability threshold would need to be chosen carefully, and this approach risks removing the causal SNP from further inclusion. A final approach is to avoid using Monte Carlo methods to perform the low-dimensional integration. Because the problem is low-dimensional, it is possible to accurately use numerical approximations, for example, quadrature methods or a Laplace approximation.

Our analysis assumes that the covariance matrix of $\hat{\boldsymbol{\beta}}$ is available. This can be obtained by running a logistic regression model on the genotype/phenotype data if it is available. As in other methods (Benner et al., 2016), it is possible to approximate this matrix if the standard errors of each SNP effect size estimate are available because the square of these values give the leading diagonals of the covariance matrix. These standard errors are available from a univariate SNP analysis and are frequently reported, or can be inferred if the effect size estimate and the p value are reported. The other covariances in this matrix (off-diagonals) can be approximated using the SNP correlation estimate from a reference panel, and using the relationship between the correlation and covariance.

A key question for the method proposed here is the choice of K. We suggested selecting K based on comparing marginal likelihoods but there are other simple strategies that could be employed. For simulated data, it is possible to compare the partial AUCs but this is clearly only possible when the causal SNPs are known. For real data, like the iCOGS data we analysed, there are a few approaches the user could take. The simplest approach is to increase K and assess the change in the SNP PIPs or ranks. This could be done across all SNPs but is probably more usefully done only for the top ranked SNPs because these are the ones of interest. For example, if the intention is to further investigate 10 SNPs via functional studies, then K could be incrementally increased and the effect on the top 10 ranked SNPs assessed. If the set of the 10 top-ranked SNPs remains

stable as K increases from say two to three, then there seems little to be gained increasing K any further.

The results of fine-mapping are likely to be sensitive to the choice of the hyperparameter λ in all but the largest of fine-mapping studies. There are several ways that λ could be selected. One is to use expert elicitation. This is a common approach to subjective prior formation in Bayesian analysis that involves eliciting quantiles from experts and using these to infer priors (Morris et al., 2014). An alternative approach is to use the effect sizes of disease-specific GWAS top hits (Walters et al., 2019). In this paper, the authors derived the maximum-likelihood estimator of λ based on the estimated effect sizes of GWAS top hits and an estimate of the number of yet-to-be-discovered top hits.

ACKNOWLEDGEMENTS

Some of this study was carried out as part of a PhD funded by the Ministry of Higher Education Malaysia and Universiti Malaya.

DATA AVAILABILITY STATEMENT

The simulated data are available at https://github.com/Kevin-walters/Code_and_data_for_multiSNP_Laplace_prior_paper. Because of the number and size of the Hapgen files produced, we only make the processed (not raw) data available. The iCOGS data are available at https://bcac.ccge.medschl.cam.ac.uk/bcacdata/. The R code needed to produce the results in this paper is available at https://github.com/Kevin-walters/Code_and_data_for_multiSNP_Laplace_prior_paper. This code makes use of some packages available for download at https://github.com/Kevin-walters/.

ORCID

Kevin Walters http://orcid.org/0000-0002-5718-5734
Hannuun Yaacob http://orcid.org/0000-0002-1056-792X

REFERENCES

Alenazi, A. A., Cox, A., Juarez, M., Lin, W.-Y., & Walters, K. (2019). Bayesian variable selection using partially observed categorical prior information in fine mapping association studies. *Genetic Epidemiology*, *43*(6), 690–703.

Benner, C., Spencer, C. C., Havulinna, A. S., Salomaa, V., Ripatti, S., & Pirinen, M. (2016). Finemap: Efficient variable selection using summary data from genome-wide association studies. *Bioinformatics*, *32*(10), 1493–1501.

Boggis, E., Milo, M., & Walters, K. (2016). equips: eqtl analysis using informed partitioning of SNPs—A fully Bayesian approach. *Genet Epidemiol*, *40*(4), 273–283.

Bottolo, L R. S. (2010). Evolutionary stochastic search for Bayesian model exploration. *Bayesian Analysis*, *5*(3), 583–618.

Chen, W., Larrabee, B. R., Ovsyannikova, I. G., Kennedy, R. B., Haralambieva, I. H., Poland, G. A., & Schaid, D. J. (2015). Fine mapping causal variants with an approximate Bayesian method using marginal test statistics. *Genetics*, *200*(3), 719–736.

Fawcett, T. (2006). An introduction to ROC analysis. *Pattern Recognition Letters*, *27*(8), 861–874.

Genz, A., & Bretz, F. (2009). *Computation of multivariate normal and* t *probabilities. Lecture notes in statistics*. Springer.

Hoggart, C. J., Whittaker, J. C., De Iorio, M., & Balding, D. J. (2008). Simultaneous analysis of all SNPs in genome-wide and re-sequencing association studies. *PLoS genetics*, *4*(7), e1000130.

Hutchinson, A., Asimit, J., & Wallace, C. (2020). Fine-mapping genetic associations. *Human Molecular Genetics*, *29*(R1), R81–R88.

Lu, T.-T., & Shiou, S.-H. (2002). Inverses of 2 × 2 block matrices. *Computers and Mathematics with Applications*, *43*(1), 119–129.

Marchini, J., & Howie, B. (2010). Genotype imputation for genome-wide association studies. *Nature Reviews Genetics*, *11*(7), 499–511.

Mi, X., Miwa, T., & Hothorn, T. (2009). mvtnorm: New numerical algorithm for multivariate normal probabilities. *The R Journal*. *1*(1), 1.

Michailidou, K., Hall, P., Gonzalez-Neira, A., Ghoussaini, M., Dennis, J., Milne, R. L., Schmidt, M. K., Chang-Claude, J., Bojesen, S. E., Bolla, M. K., Wang, Q., Ducks, E., Lee, A., Turnbull, C., & Rahman, N., Breast and Ovarian Cancer Research Group Netherlands. (2013). Large-scale genotyping identifies 41 new loci associated with breast cancer risk. *Nat Genet*, *45*(4), 353–361.

Morris, D. E., Oakley, J. E., & Crowe, J. A. (2014). A web-based tool for eliciting probability distributions from experts. *Environmental Modelling and Software*, *52*, 1–4.

R Core Team. (2021). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing.

Sing, T., Sander, O., Beerenwinkel, N., & Lengauer, T. (2005). ROCR: Visualizing classifier performance in r. *Bioinformatics*, *21*(20), 7881.

Spencer, A. V., Cox, A., Lin, W.-Y., Easton, D. F., Michailidou, K., & Walters, K. (2015). Novel Bayes factors that capture expert uncertainty in prior density specification in genetic association studies. *Genetic Epidemiology*, *39*(4), 239–248.

Spencer, A. V., Cox, A., Lin, W.-Y., Easton, D. F., Michailidou, K., & Walters, K. (2016). Incorporating functional genomic information in genetic association studies using an empirical Bayes approach. *Genetic Epidemiology*, *40*(3), 176–187.

Su, Z., Marchini, J., & Donnelly, P. (2011). Hapgen2: Simulation of multiple disease snps. *Bioinformatics*, *27*(16), 2304–2305.

VanLiere, J. M., & Rosenberg, N. A. (2008). Mathematical properties of the r^2 measure of linkage disequilibrium. *Theoretical Population Biology*, *74*(1), 130–137.

Wakefield, J. (2009). Bayes factors for genome-wide association studies: comparison with *p*-values. *Genetic Epidemiology*, *33*(1), 79–86.

Walters, K., Cox, A., & Yaacob, H. (2019). Using GWAS top hits to inform priors in Bayesian fine-mapping association studies. *Genetic Epidemiology*, *43*(6), 675–689.

Walters, K., Cox, A., & Yaacob, H. (2021). The utility of the laplace effect size prior distribution in Bayesian fine-mapping studies. *Genetic Epidemiology*, *45*(4), 386–401.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Walters, K., & Yaacob, H. (2023). Bayesian multivariant fine mapping using the laplace prior. *Genet Epidemiol*, 1–12. https://doi.org/10.1002/gepi.22517

APPENDIX A: JUSTIFICATION FOR NOT INCLUDING THE INTERCEPT IN THE MODEL

Let ρ and β represent the intercept and the SNP effect size vector, and let $\hat{\rho}$ and $\hat{\boldsymbol{\beta}}$ represent their estimators. Further, let V represent $\text{var}\begin{pmatrix}\hat{\rho}\\ \hat{\boldsymbol{\beta}}\end{pmatrix}$ and let

$$V^{-1} = \begin{pmatrix} A & \boldsymbol{D} \\ \boldsymbol{D}^T & \boldsymbol{C} \end{pmatrix}, \tag{A1}$$

where A is a scalar, $\boldsymbol{D}$ is a $1 \times p$ matrix and $\boldsymbol{C}$ is a $p \times p$ matrix. The variance matrix V can be found using a standard result (see, e.g., Lu & Shiou, 2002) about inverses of 2×2 block matrices. This gives

$$V = \begin{pmatrix} * & * \\ * & \left(\boldsymbol{C} - \frac{\boldsymbol{D}^T\boldsymbol{D}}{A}\right)^{-1} \end{pmatrix}, \tag{A2}$$

where * represent some matrices not of interest. It follows that the variance of $\hat{\boldsymbol{\beta}}$ is

$$\text{Cov}(\hat{\boldsymbol{\beta}}) = \left(\boldsymbol{C} - \frac{\boldsymbol{D}^T\boldsymbol{D}}{A}\right)^{-1}. \tag{A3}$$

We now consider reparameterisng the model. Ignoring additive constant terms, the log-likelihood (l) including the intercept is

$$\begin{aligned} l &= -\tfrac{1}{2}\begin{pmatrix}\rho - \hat{\rho}\\ \boldsymbol{\beta} - \hat{\boldsymbol{\beta}}\end{pmatrix}^T V^{-1} \begin{pmatrix}\rho - \hat{\rho}\\ \boldsymbol{\beta} - \hat{\boldsymbol{\beta}}\end{pmatrix} \\ &= -\tfrac{1}{2}[(\rho - \hat{\rho})^2 A + 2(\boldsymbol{\beta} - \hat{\boldsymbol{\beta}})^T\boldsymbol{D}^T(\rho - \hat{\rho}) \\ &\quad + (\boldsymbol{\beta} - \hat{\boldsymbol{\beta}})^T\boldsymbol{C}(\boldsymbol{\beta} - \hat{\boldsymbol{\beta}})]. \end{aligned}$$

We reparameterise via the transformation $\boldsymbol{\gamma} = \boldsymbol{\beta}$ and $\alpha = \rho + \frac{\boldsymbol{D}}{A}\boldsymbol{\beta}$. The likelihood under the new parameterisation (l^*) is

$$\begin{aligned} l^* = -\frac{1}{2}\Bigg[&\left(\alpha - \frac{\boldsymbol{D}}{A}\boldsymbol{\gamma} - \hat{\alpha}\right)^2 A + 2(\boldsymbol{\gamma} - \hat{\boldsymbol{\gamma}})^T\boldsymbol{D}^T \\ &\left(\alpha - \frac{\boldsymbol{D}}{A}\boldsymbol{\gamma} - \hat{\alpha}\right) + (\boldsymbol{\gamma} - \hat{\boldsymbol{\gamma}})^T\boldsymbol{C}(\boldsymbol{\gamma} - \hat{\boldsymbol{\gamma}})\Bigg]. \end{aligned}$$

Three of the second derivatives are

$$\begin{aligned} &\frac{\partial^2 l^*}{\partial\alpha^2} = -A, \\ &\frac{\partial^2 l^*}{\partial\boldsymbol{\gamma}\partial\alpha} = \frac{\partial^2 l^*}{\partial\alpha\partial\boldsymbol{\gamma}} = -\frac{1}{2}\left[2A\left(-\frac{\boldsymbol{D}}{A}\mathbf{1}_p\right) + 2\,\mathbf{1}_p^T\boldsymbol{D}^T\right] = \mathbf{0}, \end{aligned}$$

where $\mathbf{1}_p$ is $p \times 1$ vector of 1s. To find the second derivative of the log-likelihood with respect to $\boldsymbol{\gamma}$, we first rewrite l^* by separating second-order terms in $\boldsymbol{\gamma}$ and other terms. Because $\boldsymbol{D}\boldsymbol{\gamma}$ is a scalar we have $\boldsymbol{D}\boldsymbol{\gamma} = \boldsymbol{\gamma}^T\boldsymbol{D}^T$. Therefore,

$$\begin{aligned} l^* = &-\frac{1}{2}\left[\frac{(\boldsymbol{D}\boldsymbol{\gamma})^2}{A} - \frac{2\boldsymbol{\gamma}^T\boldsymbol{D}^T\boldsymbol{D}\boldsymbol{\gamma}}{A} + \boldsymbol{\gamma}^T\boldsymbol{C}\boldsymbol{\gamma}\right] + \\ &\text{other terms} \\ = &-\frac{1}{2}\left[\boldsymbol{\gamma}^T\left(C - \frac{\boldsymbol{D}^T\boldsymbol{D}}{A}\right)\boldsymbol{\gamma}\right] + \text{other terms}. \end{aligned} \tag{A4}$$

This gives

$$\frac{\partial^2 l^*}{\partial\boldsymbol{\gamma}^2} = \frac{\boldsymbol{D}^T\boldsymbol{D}}{A} - \boldsymbol{C}. \tag{A5}$$

The expected information matrix under the new parameterisation is, therefore,

$$I_{\text{new}} = \begin{pmatrix} A & \mathbf{0}_p \\ \mathbf{0}_p^T & \boldsymbol{C} - \frac{\boldsymbol{D}^T\boldsymbol{D}}{A} \end{pmatrix}. \tag{A6}$$

Inverting this matrix gives the covariance matrix

$$V_{\text{new}} = \begin{bmatrix} A^{-1} & \mathbf{0}_p \\ \mathbf{0}_p^T & \left(C - \frac{\boldsymbol{D}^T\boldsymbol{D}}{A}\right)^{-1} \end{bmatrix}. \tag{A7}$$

It is clear that the variance of the SNP effect size estimators is the same in the original model and the reparameterised model. The result demonstrates that we do not need to include the intercept in the likelihood (because the likelihood will factorise into an intercept term and other terms) and we can use the variance matrix of the estimated SNP effect sizes calculated when fitting a logistic regression model that includes an intercept.

Marginal likelihood for the Gaussian prior

The marginal likelihood is

$$f(\hat{\boldsymbol{\beta}}|\boldsymbol{M_c}) = \int_{\boldsymbol{\beta_c}} f(\boldsymbol{\beta_c}\ |M_{\mathbf{c}})\, f(\hat{\boldsymbol{\beta}}\ |\ \boldsymbol{\beta}) d\boldsymbol{\beta_c}. \tag{A8}$$

If the effect size prior is given by

$$\boldsymbol{\beta_c}|\boldsymbol{M_c} \sim N_d(\mathbf{0}, W),$$

where $W = \text{diag}(w_1, w_2, ..., w_d)$, then the pdf is

$$f(\boldsymbol{\beta_c}|\boldsymbol{M_c}) = [(2\pi)^d|W|]^{-1/2} \exp\left(-\frac{1}{2}\boldsymbol{\beta_c}^T\ W^{-1}\ \boldsymbol{\beta_c}\right).$$

Let

$$\begin{aligned}
\Gamma &= W^{-1} + \Sigma_c, \\
\boldsymbol{\nu} &= \hat{\boldsymbol{\beta}}_{\boldsymbol{c}}^T \Sigma_c + \hat{\boldsymbol{\beta}}_{\boldsymbol{N}}^T \Sigma^T, \\
\boldsymbol{\eta} &= \Gamma^{-1}\boldsymbol{\nu}^T, \\
\chi &= \delta - \boldsymbol{\nu}\Gamma^{-1}\boldsymbol{\nu}^T.
\end{aligned}$$

Using these definitions and the definitions in the main text, the integrand in Equation (A8) can be written as

$$\begin{aligned}
&((2\pi)^{p+d}\ |\Sigma W|)^{-1/2} \\
&\quad \exp\left(-\frac{1}{2}\left[\boldsymbol{\beta_c}^T\ \Gamma\ \boldsymbol{\beta_c} - 2\left(\hat{\boldsymbol{\beta}}_{\boldsymbol{c}}^T \Sigma_c + \hat{\boldsymbol{\beta}}_{\boldsymbol{N}}^T \Sigma^T\right)\boldsymbol{\beta_c} + \delta\right]\right) \\
&= ((2\pi)^{p+d}|\Sigma W|)^{-1/2} \exp\left(-\frac{1}{2}[(\boldsymbol{\beta_c} - \boldsymbol{\eta})^T \right. \\
&\quad \left. \Gamma\ (\boldsymbol{\beta_c} - \boldsymbol{\eta}) + \chi]\right).
\end{aligned}$$

It follows that

$$f(\hat{\boldsymbol{\beta}}|\boldsymbol{M_c}) = ((2\pi)^p\ |\Gamma W \Omega|)^{-1/2} \exp\left(-\frac{\chi}{2}\right).$$