The American Journal of Human Genetics, Volume *107*

Supplemental Data

Detecting Allele-Specific Alternative Splicing from Population-Scale RNA-Seq Data

Levon Demirdjian, Yungang Xu, Emad Bahrami-Samani, Yang Pan, Shayna Stein, Zhijie Xie, Eddie Park, Ying Nian Wu, and Yi Xing

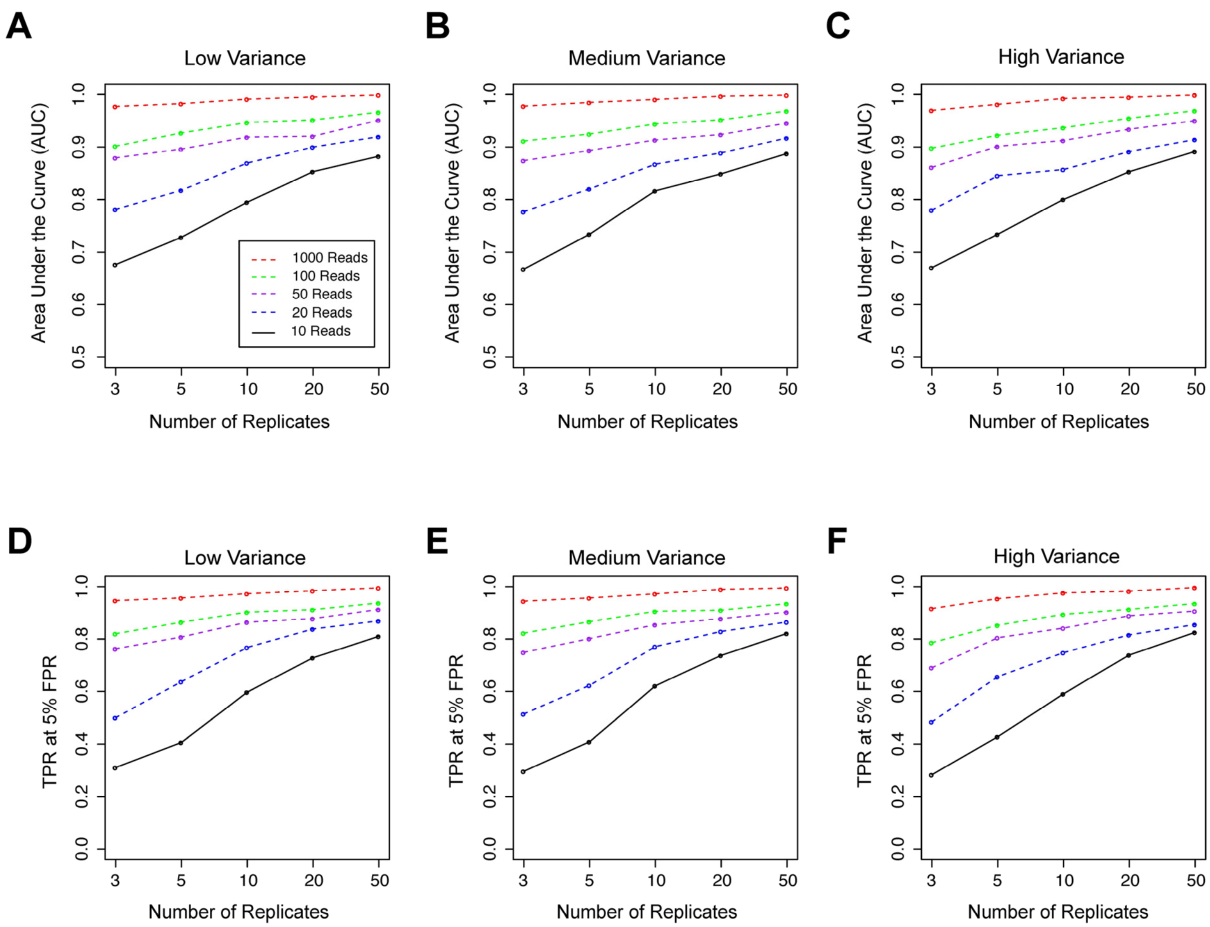

Figure S1. Simulation studies to evaluate the power of PAIRADISE with different RNA-seq read counts and different numbers of replicates. **(A-C)** The area under curve (AUC) of PAIRADISE with numbers of replicates equal to 3, 5, 10, 20, and 50; numbers of reads equal to 10, 20, 50, 100, and 1000. We used three different variance settings corresponding to low, medium, and high variability, with variance terms sampled from the 1st, 2nd, and 3rd quartiles of the empirical variance estimated from the Geuvadis CEU dataset. **(D-F)** The true positive rate (TPR) at 5% false positive rate (FPR) for various simulation settings.

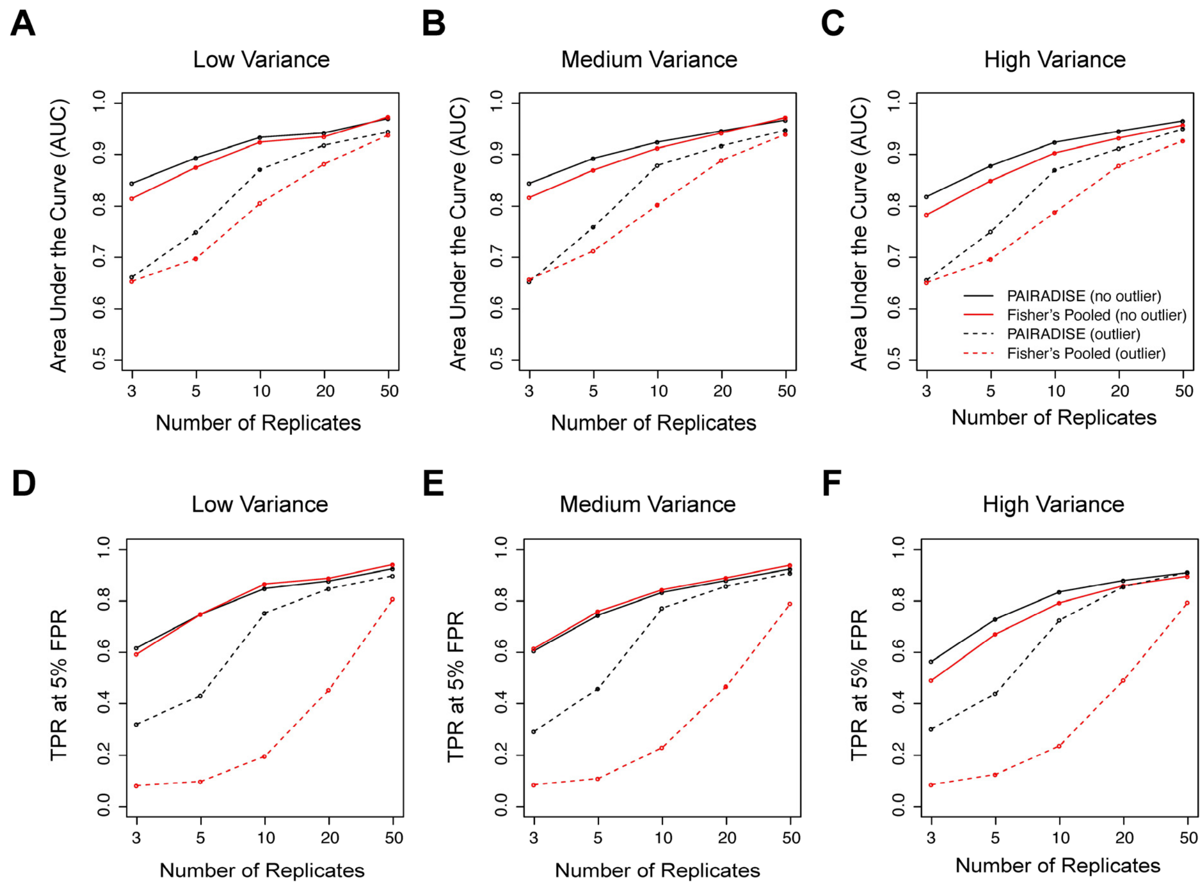

Figure S2. Simulation studies to compare the performance of PAIRADISE to Fisher's exact test using reads pooled from all replicates of the two alleles ("Fisher's pooled"), in the absence or presence of an outlier. **(A–C)** The area under curve (AUC) of both methods in simulation settings with numbers of replicates equal to 3, 5, 10, 20, and 50, and three settings of variability (low, medium, and high) sampled from the first, second, and third quartiles of the empirical variance estimated from the Geuvadis CEU dataset. **(D-F)** The true positive rate (TPR) at 5% false positive rate (FPR) of both methods in various simulation settings. Solid lines and dashed lines denote simulation settings without or with outlier, respectively.

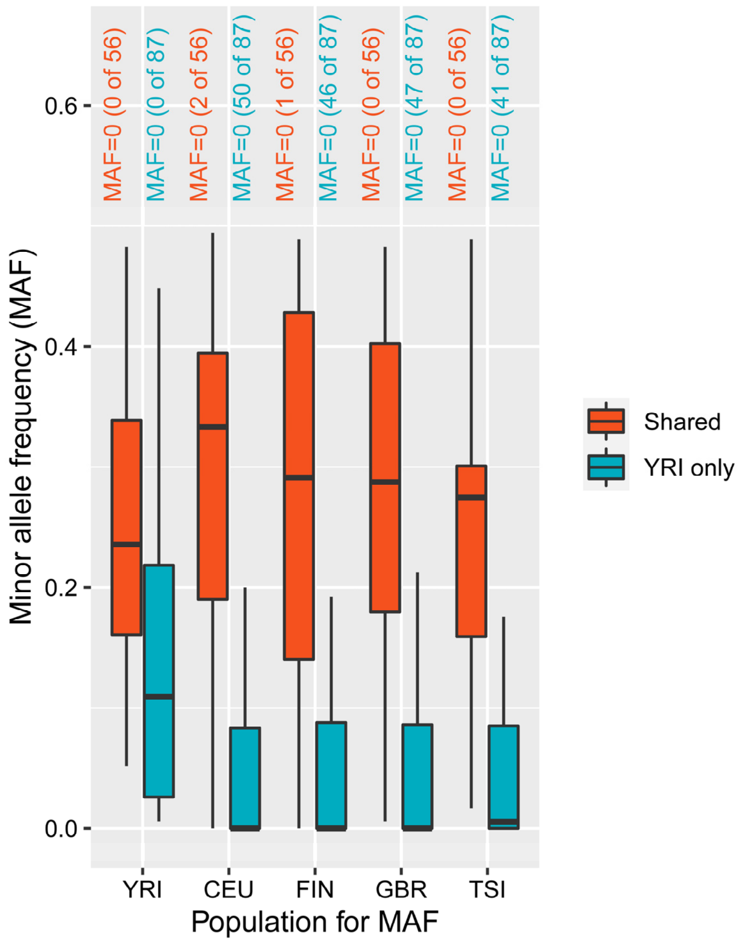

Figure S3. The minor allele frequency (MAF) of YRI ASAS SNPs in African and European populations. The boxplots show the MAFs of the YRI ASAS SNPs in each of the five populations, with the orange and blue boxplots representing ASAS SNPs of YRI ASAS events detected also in European populations ("Shared") or detected only in YRI ("YRI only") respectively. The middle line of the boxplot represents the median value. The low and high ends of the box represent the 25% and 75% quantile respectively. The two whiskers extend to the minimum and maximum value respectively. The text above each boxplot shows the number of YRI ASAS SNPs with MAF=0 in each population.

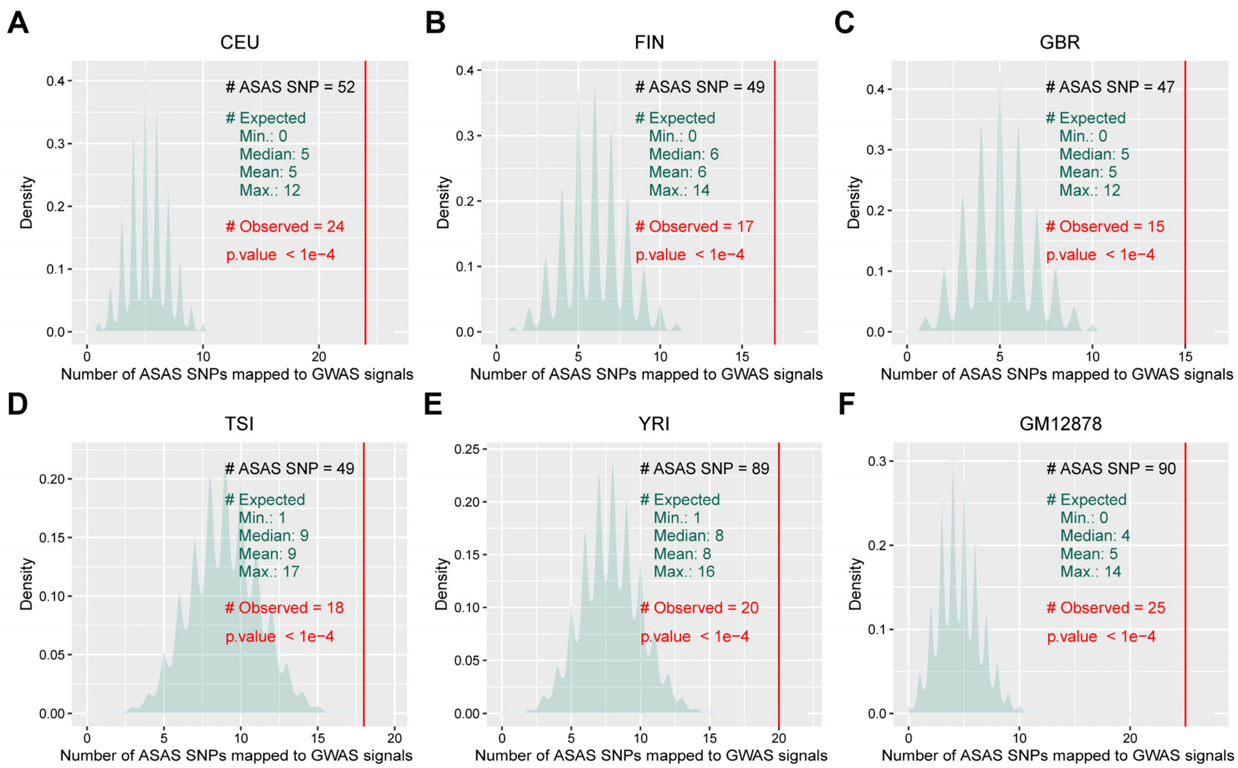

Figure S4. Enrichment of ASAS SNPs for GWAS signals. In each dataset, the red vertical line indicates the observed number of ASAS SNPs in high LD with GWAS SNPs. The density plot shows the distribution of the expected number of control non-ASAP SNPs in high LD with GWAS SNPs, based on 10,000 times of random sampling of control non-ASAP SNPs. Results from the five Geuvadis populations **(A-E)** and GM12878 **(F)** are shown.

Detecting allele-specific alternative splicing from population-scale RNA-seq data

Demirdjian et al.

1 Supplementary Methods

1.1 PAIRADISE statistical model

Let I_{ijk} be the number of RNA-seq reads corresponding to the exon inclusion isoform for exon $i, i = 1, ..., n$, replicate $k, k = 1, ..., M$ and allele/group $j, j = 1, 2$. Similarly, define S_{ijk} to be the number of RNA-seq reads corresponding to the exon skipping isoform for replicate k and allele j. PAIRADISE uses a binomial distribution to model the estimation uncertainty in individual replicates, i.e.

$$I_{i1k}|\psi_{i1k} \sim \text{Bin}\left(n_{i1k} = I_{i1k} + S_{i1k}, p_{i1k} = \frac{\ell_{iI}\psi_{i1k}}{\ell_{iI}\psi_{i1k} + \ell_{iS}(1-\psi_{i1k})}\right) \tag{1}$$

$$I_{i2k}|\psi_{i2k} \sim \text{Bin}\left(n_{i2k} = I_{i2k} + S_{i2k}, p_{i2k} = \frac{\ell_{iI}\psi_{i2k}}{\ell_{iI}\psi_{i2k} + \ell_{iS}(1-\psi_{i2k})}\right).$$

ℓ_{iI} and ℓ_{iS} are the effective lengths (the number of unique isoform-specific read positions) of the inclusion isoform and skipping isoform respectively. ψ_{ijk} is the exon inclusion level for replicate k in group j. To model the variability among replicates, let

$$\text{logit}(\psi_{i1k}) = \alpha_{ik} + \epsilon_{i1k}, \qquad \text{logit}(\psi_{i2k}) = \alpha_{ik} + \delta_i + \epsilon_{i2k}, \tag{2}$$

where the subject effect for exon i, α_{ik}, is assumed to follow a normal distribution

$$\alpha_{ik} \overset{iid}{\sim} N(\mu_i, \sigma_i^2), \quad k = 1, ..., M; \tag{3}$$

in other words, the α_{ik} all follow the same normal distribution with mean μ_i and variance σ_i^2. In expression (2), we are assuming that

$$\epsilon_{i1k} \overset{iid}{\sim} N(0, \sigma_{i1}^2), \qquad \epsilon_{i2k} \overset{iid}{\sim} N(0, \sigma_{i2}^2), \quad k = 1, ..., M, \tag{4}$$

and that ϵ_{i1k} and ϵ_{i2k} are independent of each other. The variable δ_i in (2) measures the expected difference between $\text{logit}(\psi_{i2k})$ and $\text{logit}(\psi_{i1k})$, i.e.

$$\delta_i = E\left[\text{logit}(\psi_{i2k}) - \text{logit}(\psi_{i1k})|\alpha_{ik}\right].$$

Equations (2), (3), and (4) imply the following conditional distributions:

$$\text{logit}(\psi_{i1k})\big|\alpha_{ik} \sim N(\alpha_{ik}, \sigma_{i1}^2)$$

$$\text{logit}(\psi_{i2k})\big|\alpha_{ik}, \delta_i \sim N(\alpha_{ik} + \delta_i, \sigma_{i2}^2).$$

The above setup yields the following joint distribution of $\text{logit}(\psi_{i1k})$ and $\text{logit}(\psi_{i2k})$:

$$\left(\begin{bmatrix} \text{logit}(\psi_{i1k}) \\ \text{logit}(\psi_{i2k}) \end{bmatrix} \Big| \alpha_{ik}, \sigma_{i1}, \sigma_{i2}, \delta_i \right) \sim N\left(\begin{pmatrix} \alpha_{ik} \\ \alpha_{ik} + \delta_i \end{pmatrix}, \begin{pmatrix} \sigma_{i1}^2 & 0 \\ 0 & \sigma_{i2}^2 \end{pmatrix} \right). \tag{5}$$

1.2 Derivation of the likelihood function

PAIRADISE aims to test whether there is a significant difference in the means of the distributions of $\text{logit}(\psi_{i1k})$ and $\text{logit}(\psi_{i2k})$. Adopting the notation of hypothesis testing, PAIRADISE tests the null hypothesis $H_0 : \delta_i = 0$ (no difference) against the alternative hypothesis $H_a : \delta_i \neq 0$ (there is a difference) using a likelihood ratio test. In the PAIRADISE framework, I_{i1k} and I_{i2k} are the observed data and ψ_{i1k}, ψ_{i2k}, and α_{ik} are all regarded as latent, unobserved variables. In order to make inference about the parameter δ_i, we must first derive an expression for the observed data likelihood. For a given exon i, the complete data likelihood function (the likelihood of the observed and latent variables) is given by

$$\prod_{k=1}^{M} f(I_{i1k}, I_{i2k}, \text{logit}(\psi_{i1k}), \text{logit}(\psi_{i2k}), \alpha_{ik} | \theta_i), \tag{6}$$

where we have set $\theta_i = (\delta_i, \sigma_{i1}, \sigma_{i2}, \sigma_i, \mu_i)$ for notational simplicity. To derive an expression for the observed data likelihood, we can integrate the latent variables out of the complete data likelihood in (6) to obtain

$$\begin{aligned} &\prod_{k=1}^{M} f(I_{i1k}, I_{i2k} | \theta_i) \\ &= \prod_{k=1}^{M} \int f(I_{i1k}, I_{i2k}, \text{logit}(\psi_{i1k}), \text{logit}(\psi_{i2k}), \alpha_{ik} | \theta_i) d\,\text{logit}(\psi_{i1k}) \cdot d\,\text{logit}(\psi_{i2k}) \cdot d\alpha_{ik}. \end{aligned} \tag{7}$$

Since there is no closed-form expression for the integral in (7), we proceed by using Laplace's method to obtain an approximation of this integral. Let $f_1 = \log(f)$, and let $\widehat{\alpha_{ik}}, \text{logit}(\widehat{\psi_{i1k}})$, and $\text{logit}(\widehat{\psi_{i2k}})$ be the MLE's of $\alpha_{ik}, \text{logit}(\psi_{i1k})$, and $\text{logit}(\psi_{i2k})$. Then for $k = 1, ..., M$,

$$\begin{aligned} &\int f(I_{i1k}, I_{i2k}, \text{logit}(\psi_{i1k}), \text{logit}(\psi_{i2k}), \alpha_{ik} | \theta_i) d\,\text{logit}(\psi_{i1k}) \cdot d\,\text{logit}(\psi_{i2k}) \cdot d\alpha_{ik} \\ &= \int \exp(f_1(I_{i1k}, I_{i2k}, \text{logit}(\psi_{i1k}), \text{logit}(\psi_{i2k}), \alpha_{ik} | \theta_i)) d\,\text{logit}(\psi_{i1k}) \cdot d\,\text{logit}(\psi_{i2k}) \cdot d\alpha_{ik} \\ &= \int \exp\{f_1(I_{i1k}, I_{i2k}, \text{logit}(\widehat{\psi_{i1k}}), \text{logit}(\widehat{\psi_{i2k}}), \widehat{\alpha_{ik}} | \theta_i) \end{aligned}$$

$$+ \frac{1}{2} \begin{bmatrix} \text{logit}(\psi_{i1k}) - \text{logit}(\widehat{\psi_{i1k}}) \\ \text{logit}(\psi_{i2k}) - \text{logit}(\widehat{\psi_{i2k}}) \\ \alpha_{ik} - \widehat{\alpha_{ik}} \end{bmatrix}' \Sigma_{ik} \begin{bmatrix} \text{logit}(\psi_{i1k}) - \text{logit}(\widehat{\psi_{i1k}}) \\ \text{logit}(\psi_{i2k}) - \text{logit}(\widehat{\psi_{i2k}}) \\ \alpha_{ik} - \widehat{\alpha_{ik}} \end{bmatrix}$$

$$+ o((\text{logit}(\psi_{i1k}) - \text{logit}(\widehat{\psi_{i1k}}))^2) + o((\text{logit}(\psi_{i2k}) - \text{logit}(\widehat{\psi_{i2k}}))^2)$$

$$+ o((\alpha_{ik} - \widehat{\alpha_{ik}})^2)\} \quad d\,\text{logit}(\psi_{i1k}) \cdot d\,\text{logit}(\psi_{i2k}) \cdot d\alpha_{ik}$$

$$\approx (2\pi)^{3/2}(-|\Sigma_{ik}|)^{-1/2} \exp\{f_1(I_{i1k}, I_{i2k}, \text{logit}(\widehat{\psi_{i1k}}), \text{logit}(\widehat{\psi_{i2k}}), \widehat{\alpha_{ik}}|\theta_i)\}, \tag{8}$$

where

$$\Sigma_{ik} = \begin{bmatrix} \dfrac{\partial^2 f_1(z_{ik})}{\partial\,\text{logit}^2(\psi_{i1k})} & \dfrac{\partial^2 f_1(z_{ik})}{\partial\,\text{logit}(\psi_{i1k})\partial\,\text{logit}(\psi_{i2k})} & \dfrac{\partial^2 f_1(z_{ik})}{\partial\,\text{logit}(\psi_{i1k})\partial\alpha_{ik}} \\ \dfrac{\partial^2 f_1(z_{ik})}{\partial\,\text{logit}(\psi_{i1k})\partial\,\text{logit}(\psi_{i2k})} & \dfrac{\partial^2 f_1(z_{ik})}{\partial\,\text{logit}^2(\psi_{i2k})} & \dfrac{\partial^2 f_1(z_{ik})}{\partial\,\text{logit}(\psi_{i2k})\partial\alpha_{ik}} \\ \dfrac{\partial^2 f_1(z_{ik})}{\partial\alpha_{ik}\partial\,\text{logit}(\psi_{i1k})} & \dfrac{\partial^2 f_1(z_{ik})}{\partial\alpha_{ik}\partial\,\text{logit}(\psi_{i2k})} & \dfrac{\partial^2 f_1(z_{ik})}{\partial\alpha_{ik}^2} \end{bmatrix}, \tag{9}$$

and where we have used the shorthand $f_1(z_{ik}) := f_1(\text{logit}(\widehat{\psi_{i1k}}), \text{logit}(\widehat{\psi_{i2k}}), \widehat{\alpha_{ik}}|I_{i1k}, I_{i2k}, \theta_i)$. Note that the determinant of the above Hessian matrix is always negative (shown at the end of the appendix). Combining (7) and (8) yields an expression for the observed data likelihood:

$$\prod_{k=1}^{M} f(I_{i1k}, I_{i2k}|\theta_i) \approx C_1 \prod_{k=1}^{M} (-|\Sigma_{ik}|)^{-1/2} f(I_{i1k}, I_{i2k}, \text{logit}(\widehat{\psi_{i1k}}), \text{logit}(\widehat{\psi_{i2k}}), \widehat{\alpha_{ik}}|\theta_i)$$

or

$$\sum_{k=1}^{M} \log f(I_{i1k}, I_{i2k}|\delta_i, \sigma_{i1}, \sigma_{i2}, \mu_i, \sigma_i) \approx$$
$$\sum_{k=1}^{M} \left\{ f_1(I_{i1k}, I_{i2k}, \text{logit}(\widehat{\psi_{i1k}}), \text{logit}(\widehat{\psi_{i2k}}), \widehat{\alpha_{ik}}|\delta_i, \sigma_{i1}, \sigma_{i2}, \mu_i, \sigma_i) - \frac{1}{2}\log(-|\Sigma_{ik}|) \right\} + C_2 \tag{10}$$

for some constants C_1 and C_2.

1.3 Optimization

Next, we outline an iterative procedure that will produce estimates $(\hat{\delta_i}, \hat{\sigma_{i1}}, \hat{\sigma_{i2}}, \hat{\mu_i}, \hat{\sigma_i})$ based on the observed data log likelihood in (10). The optimization procedure consists of the following steps:

1.3.1 Initializing the latent variables

For $k = 1, ..., M$, initialize $\text{logit}(\psi_{i1k})$ and $\text{logit}(\psi_{i2k})$ from the individual binomial distribution of each replicate:

$$\text{logit}(\widehat{\psi_{i1k}^{(0)}}) = \text{logit}\left(\frac{I_{i1k}\ell_{iS}}{I_{i1k}\ell_{iS} + S_{i1k}\ell_{iI}}\right), \quad \text{logit}(\widehat{\psi_{i2k}^{(0)}}) = \text{logit}\left(\frac{I_{i2k}\ell_{iS}}{I_{i2k}\ell_{iS} + S_{i2k}\ell_{iI}}\right).$$

Since $\text{logit}(\psi_{i1k}) = \alpha_{ik} + \epsilon_{i1k}$, one can set $\widehat{\alpha_{ik}}^{(0)} = \text{logit}(\widehat{\psi_{i1k}^{(0)}})$.

1.3.2 Estimating the parameters

Next, let $t \leftarrow 1$ and proceed to the following step.

Step 1: Estimate the MLEs of the observed data likelihood based on the estimated values of $\text{logit}(\widehat{\psi_{i1k}^{(t-1)}}), \text{logit}(\widehat{\psi_{i2k}^{(t-1)}})$, and $\widehat{\alpha_{ik}}^{(t-1)}$. That is, maximize expression (10):

$$(\hat{\delta_i}^{(t)}, \hat{\sigma_{i1}}^{(t)}, \hat{\sigma_{i2}}^{(t)}, \hat{\mu_i}^{(t)}, \hat{\sigma_i}^{(t)}) =$$

$$\underset{\delta_i,\sigma_{i1},\sigma_{i2},\mu_i,\sigma_i}{\text{argmax}} \sum_{k=1}^{M} \left(f_1(I_{i1k}, I_{i2k}, \text{logit}(\widehat{\psi_{i1k}^{(t-1)}}), \text{logit}(\widehat{\psi_{i2k}^{(t-1)}}), \widehat{\alpha_{ik}}^{(t-1)} | \delta_i, \sigma_{i1}, \sigma_{i2}, \mu_i, \sigma_i) - \frac{1}{2}\log(-|\Sigma_{ik}^{(t-1)}|) \right).$$

$\Sigma_{ik}^{(t-1)}$ is the Hessian matrix given in (9) where the partial derivatives are evaluated using the estimates $\widehat{\alpha_{ik}}^{(t-1)}, \text{logit}(\widehat{\psi_{i1k}^{(t-1)}})$, and $\text{logit}(\widehat{\psi_{i2k}^{(t-1)}})$.

Step 2: For $k = 1, ..., M$, update the estimates $\widehat{\alpha_{ik}}^{(t)}, \text{logit}(\widehat{\psi_{i1k}^{(t)}})$, and $\text{logit}(\widehat{\psi_{i2k}^{(t)}})$ based on the complete data likelihood (6) and the latest MLEs of $\hat{\delta_i}^{(t)}, \hat{\sigma_{i1}}^{(t)}, \hat{\sigma_{i2}}^{(t)}, \hat{\mu_i}^{(t)}, \hat{\sigma_i}^{(t)}$:

$$\begin{aligned}
&(\text{logit}(\widehat{\psi_{i1k}^{(t)}}), \text{logit}(\widehat{\psi_{i2k}^{(t)}}), \widehat{\alpha_{ik}}^{(t)}) = \\
&\underset{\text{logit}(\psi_{i1k}),\text{logit}(\psi_{i2k}),\alpha_{ik}}{\text{argmax}} \Bigg\{ -\frac{(\text{logit}(\psi_{i1k}) - \alpha_{ik})^2}{2\hat{\sigma_{i1}}^{2(t)}} - \frac{(\text{logit}(\psi_{i2k}) - \alpha_{ik} - \hat{\delta_i}^{(t)})^2}{2\hat{\sigma_{i2}}^{2(t)}} \\
&+ I_{i1k}\log\left(\frac{\ell_{iI}\psi_{i1k}}{\ell_{iI}\psi_{i1k} + \ell_{iS}(1-\psi_{i1k})}\right) + S_{i1k}\log\left(\frac{\ell_{iS}(1-\psi_{i1k})}{\ell_{iI}\psi_{i1k} + \ell_{iS}(1-\psi_{i1k})}\right) \\
&+ I_{i2k}\log\left(\frac{\ell_{iI}\psi_{i2k}}{\ell_{iI}\psi_{i2k} + \ell_{iS}(1-\psi_{i2k})}\right) + S_{i2k}\log\left(\frac{\ell_{iS}(1-\psi_{i2k})}{\ell_{iI}\psi_{i2k} + \ell_{iS}(1-\psi_{i2k})}\right) \\
&- \frac{1}{2\hat{\sigma}_i^{2(t)}}(\alpha_{ik} - \hat{\mu_i}^{(t)})^2 \Bigg\}.
\end{aligned}$$

Step 3: Let $t \leftarrow t+1$ and go to step 1. Iterate between steps 1 and 2 until the difference in log likelihoods between 2 consecutive iterations is smaller than some threshold ϵ, say $\epsilon = 10^{-2}$. Use an optimization algorithm (e.g. L-BFGS-B or BOBYQA) to optimize the likelihood function with the parameters $\sigma_{i1}, \sigma_{i2}, \sigma_i$ constrained within $(0, \infty)$, and $\alpha_{ik}, \mu_i, \delta_i, \text{logit}(\psi_{i1k}), \text{logit}(\psi_{i2k})$ unconstrained.

The above optimization procedure is performed for two cases: the unconstrained model, and the model constrained under the null hypothesis (i.e. the model with $\delta_i = 0$). The likelihood-ratio test statistic then asymptotically (in M) follows a χ^2 distribution with 1 degree of freedom:

$$-2(\log L_{\delta_i=0} - \log L) \sim \chi^2,$$

where $L_{\delta_i=0}$ is the likelihood function under the null hypothesis and L is the likelihood function under the alternative hypothesis.

1.4 Computing the Hessian Σ_{ik}

The expressions for the partial derivatives in the Hessian matrix Σ_{ik} given in (9), evaluated at $\widehat{\alpha_{ik}}, \text{logit}(\widehat{\psi_{i1k}}), \text{logit}(\widehat{\psi_{i2k}})$, are given by

$$\frac{\partial^2 f_1(z_{ik})}{\partial \, \text{logit}^2\,(\psi_{i1k})} = \frac{\ell_{iI}\ell_{iS}\hat{\psi}_{i1k}(\hat{\psi}_{i1k}-1)(I_{i1k}+S_{i1k})}{[\ell_{iI}\hat{\psi}_{i1k}+\ell_{iS}(1-\hat{\psi}_{i1k})]^2} - \frac{1}{\sigma_{i1}^2} \tag{11}$$

$$\frac{\partial^2 f_1(z_{ik})}{\partial \, \text{logit}^2\,(\psi_{i2k})} = \frac{\ell_{iI}\ell_{iS}\hat{\psi}_{i2k}(\hat{\psi}_{i2k}-1)(I_{i2k}+S_{i2k})}{[\ell_{iI}\hat{\psi}_{i2k}+\ell_{iS}(1-\hat{\psi}_{i2k})]^2} - \frac{1}{\sigma_{i2}^2} \tag{12}$$

$$\frac{\partial^2 f_1(z_{ik})}{\partial \alpha_{ik}^2} = -\left[\frac{1}{\sigma_{i1}^2} + \frac{1}{\sigma_{i2}^2} + \frac{1}{\sigma_i^2}\right] \tag{13}$$

$$\frac{\partial^2 f_1(z_{ik})}{\partial \alpha_{ik} \partial \, \text{logit}(\psi_{i1k})} = \frac{\partial^2 f_1(z_{ik})}{\partial \, \text{logit}(\psi_{i1k}) \partial \alpha_{ik}} = \frac{1}{\sigma_{i1}^2}$$

$$\frac{\partial^2 f_1(z_{ik})}{\partial \alpha_{ik} \partial \, \text{logit}(\psi_{i2k})} = \frac{\partial^2 f_1(z_{ik})}{\partial \, \text{logit}(\psi_{i2k}) \partial \alpha_{ik}} = \frac{1}{\sigma_{i2}^2}$$

$$\frac{\partial^2 f_1(z_{ik})}{\partial \, \text{logit}\,(\psi_{i1k}) \partial \, \text{logit}\,(\psi_{i2k})} = \frac{\partial^2 f_1(z_{ik})}{\partial \, \text{logit}\,(\psi_{i2k}) \partial \, \text{logit}\,(\psi_{i1k})} = 0.$$

The determinant of $|\Sigma_{ik}^{(t)}|$ is therefore given by the following expression:

$$\left|\Sigma_{ik}^{(t)}\right| = \frac{1}{(\sigma_{i1}^2)^2}\left(\frac{1}{\sigma_{i2}^2} - \frac{\ell_{iI}\ell_{iS}\hat{\psi}_{i2k}^{(t)}(\hat{\psi}_{i2k}^{(t)}-1)(I_{i2k}+S_{i2k})}{[\ell_{iI}\hat{\psi}_{i2k}^{(t)}+\ell_{iS}(1-\hat{\psi}_{i2k}^{(t)})]^2}\right) + \left(\frac{1}{\sigma_{i1}^2} - \frac{\ell_{iI}\ell_{iS}\hat{\psi}_{i1k}^{(t)}(\hat{\psi}_{i1k}^{(t)}-1)(I_{i1k}+S_{i1k})}{[\ell_{iI}\hat{\psi}_{i1k}^{(t)}+\ell_{iS}(1-\hat{\psi}_{i1k}^{(t)})]^2}\right)$$

$$\cdot \left[\left(\frac{\ell_{iI}\ell_{iS}\hat{\psi}_{i2k}^{(t)}(\hat{\psi}_{i2k}^{(t)}-1)(I_{i2k}+S_{i2k})}{[\ell_{iI}\hat{\psi}_{i2k}^{(t)}+\ell_{iS}(1-\hat{\psi}_{i2k}^{(t)})]^2} - \frac{1}{\sigma_{i2}^2}\right)\left(\frac{1}{\sigma_{i1}^2}+\frac{1}{\sigma_{i2}^2}+\frac{1}{\sigma_i^2}\right) + \frac{1}{(\sigma_{i2}^2)^2}\right].$$

1.5 Proof that the determinant of Σ_{ik} is negative

To ease notation, rewrite the Hessian in (9) as

$$\Sigma_{ik} = \begin{bmatrix} x_1 & 0 & \frac{1}{\sigma_{i1}^2} \\ 0 & x_2 & \frac{1}{\sigma_{i2}^2} \\ \frac{1}{\sigma_{i1}^2} & \frac{1}{\sigma_{i2}^2} & x_3 \end{bmatrix}$$

where x_1, x_2 and x_3 are defined as in (11), (12) and (13) (we ignore the indices i and k for additional clarity). Next, let

$$a_1 = \frac{\ell_{iI}\ell_{iS}\hat{\psi}_{i1k}(\hat{\psi}_{i1k}-1)(I_{i1k}+S_{i1k})}{[\ell_{iI}\hat{\psi}_{i1k}+\ell_{iS}(1-\hat{\psi}_{i1k})]^2}$$

and

$$a_2 = \frac{\ell_{iI}\ell_{iS}\hat{\psi}_{i2k}(\hat{\psi}_{i2k}-1)(I_{i2k}+S_{i2k})}{[\ell_{iI}\hat{\psi}_{i2k}+\ell_{iS}(1-\hat{\psi}_{i2k})]^2}$$

so that

$$x_1 = a_1 - \frac{1}{\sigma_{i1}^2}$$

and

$$x_2 = a_2 - \frac{1}{\sigma_{i2}^2}.$$

It follows that

$$\det(\Sigma_{ik}) = (x_3 - 1)x_1x_2 + \det\begin{bmatrix} x_1 - \frac{1}{\sigma_{i1}^4} & -\frac{1}{\sigma_{i1}^2\sigma_{i2}^2} \\ -\frac{1}{\sigma_{i1}^2\sigma_{i2}^2} & x_2 - \frac{1}{\sigma_{i2}^4} \end{bmatrix}$$

$$
\begin{aligned}
&= x_1 x_2 x_3 - \left[\frac{x_1}{\sigma_{i2}^4} + \frac{x_2}{\sigma_{i1}^4}\right] \\
&= -\left[\frac{1}{\sigma_i^2} + \frac{1}{\sigma_{i1}^2} + \frac{1}{\sigma_{i2}^2}\right]\left[(a_1 - \frac{1}{\sigma_{i1}^2})(a_2 - \frac{1}{\sigma_{i2}^2})\right] - \left[\frac{a_1}{\sigma_{i2}^4} - \frac{1}{\sigma_{i1}^2\sigma_{i2}^4} + \frac{a_2}{\sigma_{i1}^4} - \frac{1}{\sigma_{i2}^2\sigma_{i1}^4}\right] \\
&= -\frac{a_1 a_2}{\sigma_{i1}^2} + \frac{a_1}{\sigma_{i1}^2\sigma_{i2}^2} + \frac{a_2}{\sigma_{i1}^4} - \frac{1}{\sigma_{i1}^4\sigma_{i2}^2} - \frac{a_1 a_2}{\sigma_{i2}^2} + \frac{a_1}{\sigma_{i2}^4} + \frac{a_2}{\sigma_{i1}^2\sigma_{i2}^2} - \frac{1}{\sigma_{i1}^2\sigma_{i2}^4} \\
&- \frac{a_1 a_2}{\sigma_i^2} + \frac{a_1}{\sigma_i^2\sigma_{i2}^2} + \frac{a_2}{\sigma_i^2\sigma_{i1}^2} - \frac{1}{\sigma_i^2\sigma_{i1}^2\sigma_{i2}^2} - \left[\frac{a_1}{\sigma_{i2}^4} + \frac{a_2}{\sigma_{i1}^4} - \frac{1}{\sigma_{i1}^2\sigma_{i2}^4} - \frac{1}{\sigma_{i2}^2\sigma_{i1}^4}\right] \\
&= -a_1 a_2\Big(\frac{1}{\sigma_i^2} + \frac{1}{\sigma_{i1}^2} + \frac{1}{\sigma_{i2}^2}\Big) + a_1\Big(\frac{1}{\sigma_{i1}^2\sigma_{i2}^2} + \frac{1}{\sigma_i^2\sigma_{i2}^2}\Big) + a_2\Big(\frac{1}{\sigma_{i1}^2\sigma_{i2}^2} + \frac{1}{\sigma_i^2\sigma_{i1}^2}\Big) - \frac{1}{\sigma_i^2\sigma_{i1}^2\sigma_{i2}^2} < 0,
\end{aligned}
$$

which follows since $a_1, a_2 < 0$.