REHE: Fast Variance Components Estimation under Linear Mixed Models (Supplementary Material)

Kun Yue[1], Jing Ma[2,3], Timothy Thornton[1], and Ali Shojaie[1,*]

[1]Department of Biostatistics, University of Washington, Seattle, WA, USA
[2]Public Health Sciences Division, Fred Hutchinson Cancer Research Center, Seattle, Washington, USA
[3]Department of Statistics, Texas A&M University, College Station, Texas, USA
[*]e-mail:ashojaie@uw.edu

Supplementary Figure 1

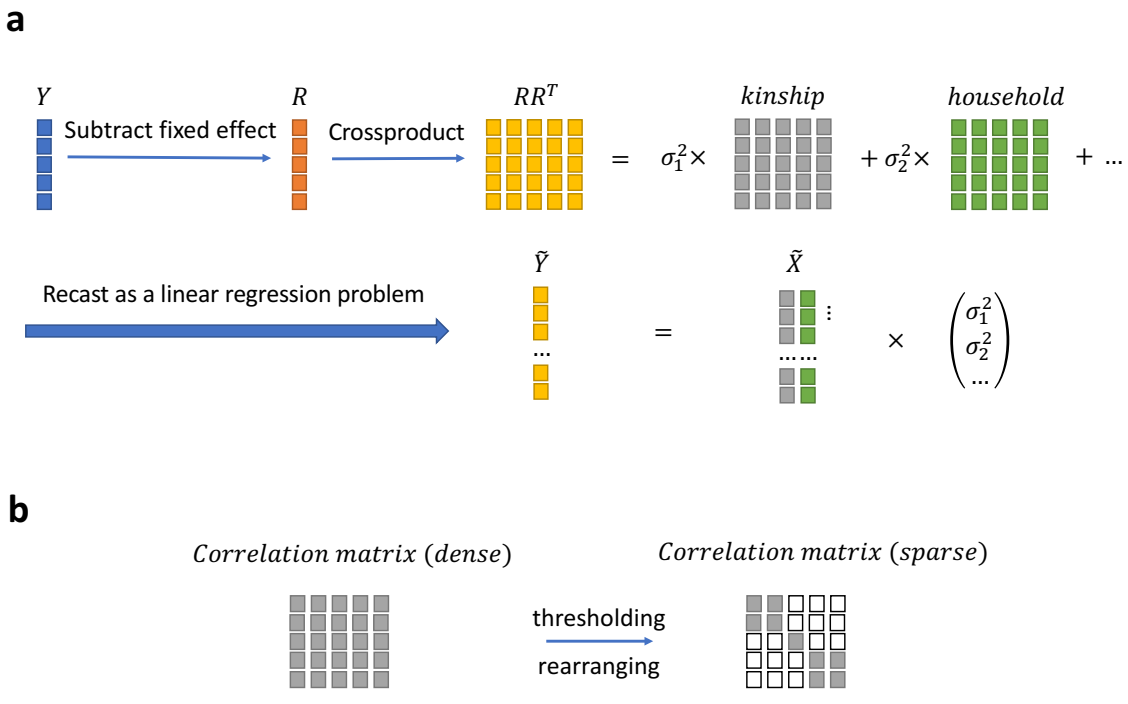

Figure S1: Illustration of HE/REHE workflow and correlation matrix sparsification. A: HE/REHE workflow. Y is the outcome vector and R is the residual vector after subtracting fixed effects from Y. Estimating equations are formed by decomposing the total variation $RR^\top$ into different sources (e.g. kinship, household relatedness) with corresponding variance components σ_k^2's. After vectorizing the total variation $RR^\top$ as $\widetilde{Y}$ and the relatedness matrices as $\widetilde{X}$, variance components are estimated using ordinary least squares. B: Illustration of correlation matrix sparsification.

Supplementary Figure 2

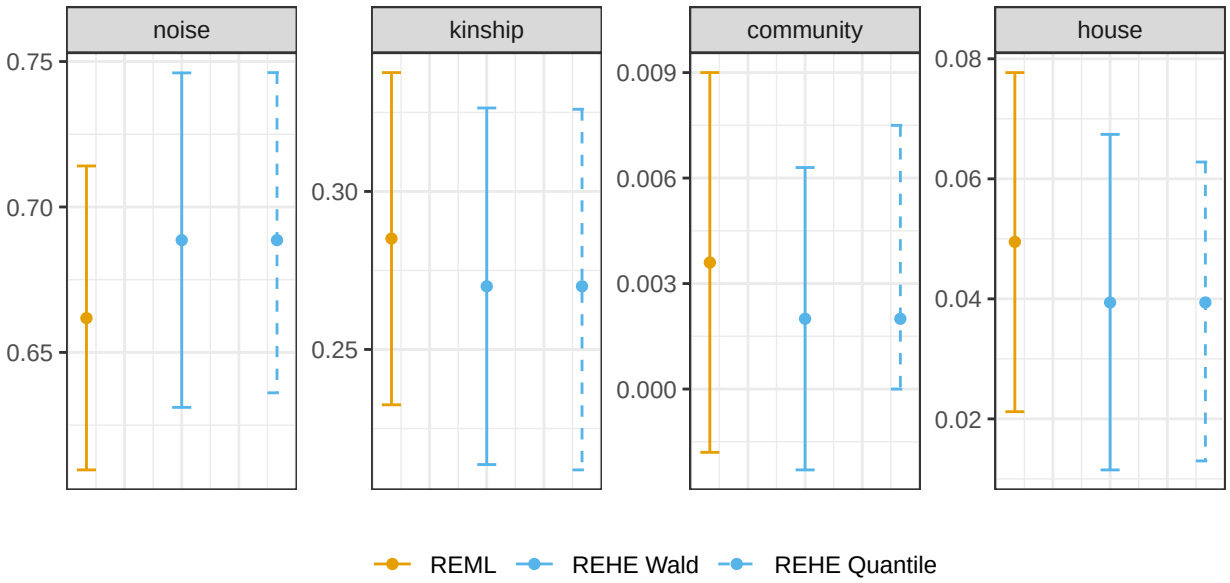

Figure S2: Confidence intervals (bar) and point estimates (dot) by REHE and REML for the null model in the HCHS/SOL data analysis, for proportion of variance attributed to noise, kinship (heritability), community block and household. Two types of REHE-based confidence intervals were presented: Wald type confidence intervals (REHE Wald), and quantile type confidence intervals (REHE Quantile).

Supplementary Note 1: Extensions of REHE

REHE with fixed effects covariates

In the main paper and Online Methods, REHE was presented for estimating variance components in linear mixed models without fixed effect covariates. The REHE estimation procedure can be easily modified to accommodate fixed effects. Let X denote the design matrix with p covariates and $\boldsymbol{\beta}$ denote the fixed effects. The full model becomes

$$Y = X\boldsymbol{\beta} + \sum_{k=0}^{K} \sigma_k \boldsymbol{\gamma_k}. \tag{S1}$$

Let $P_X^{\perp} = I_n - X(X^\top X)^{-1}X^\top$ denote the projection matrix onto the orthogonal complement of the column space of X. We project the outcome Y and the random effects $\boldsymbol{\gamma_k}$'s (including the noise term $\boldsymbol{\gamma_0}$) as

$$\begin{aligned} Y^\dagger &= P_X^{\perp} Y, \\ \boldsymbol{\gamma}_{\boldsymbol{k}}^\dagger &= P_X^{\perp} \boldsymbol{\gamma_k}. \end{aligned}$$

Recall that each random effect $\boldsymbol{\gamma_k}$ follows a normal distribution with mean 0 and covariance D_k. Writing $D_k^\dagger = P_X^{\perp} D_k P_X^{\perp}$, model (S1) becomes

$$Y^\dagger = \sum_{k=0}^{K} \sigma_k \boldsymbol{\gamma}_{\boldsymbol{k}}^\dagger, \quad \boldsymbol{\gamma}_{\boldsymbol{k}}^\dagger \sim N_n(0, D_k^\dagger), \quad k = 0, \ldots, K. \tag{S2}$$

With model (S2), REHE estimation of variance components directly follows the procedures introduced in the Online Methods. When the sample size n is large, computing the projected correlation matrices $D_k^\dagger$ is time-consuming. In practice, the number of fixed effect covariates p is typically very small compared to the sample size n. Thus the column space of X is negligible for dense and full rank correlation matrix D_k, and we expect $D_k^\dagger$ to be close to D_k. We observed little difference in the results when $D_k^\dagger$ was replaced by D_k in real data analysis and simulation studies. We thus suggest using D_k for computational efficiency when estimating model (S2) via REHE.

If the fixed effect coefficients β are of interest, they can be estimated using ordinary least squares as $\hat{\boldsymbol{\beta}} = (X^\top X)^{-1} X^\top Y$, or weighted least squares as $\hat{\boldsymbol{\beta}} = (X^\top \hat{\Sigma}^{-1} X)^{-1} X^\top \hat{\Sigma}^{-1} Y$, where $\hat{\Sigma} = \sum_{k=0}^{K} \hat{\sigma}_k^2 D_k$ is based on previously estimated variance components $\hat{\sigma}_k^2$'s. While the resulted estimates are consistent, one can iteratively update $\hat{\sigma}_k^2$'s and $\hat{\boldsymbol{\beta}}$ as illustrated in [1].

Analytical solutions to REHE

When there are only two variance components in the linear mixed model ($K = 1$), a closed form expression for REHE variance component estimates is available. Specifically, problem (4) in the Online Methods can be reformulated as:

$$(x, y) = \arg\min_{x \geq 0, y \geq 0} \; ax^2 + by^2 + cxy + dx + ey + f$$
$$\text{for some } a, b > 0, \; 4ab \neq c^2.$$

The condition $4ab \neq c^2$ guarantees that the matrix $\widetilde{X}^\top \widetilde{X}$ is invertible, where $\widetilde{X}$ is introduced in the Online Methods (equation (3)) as the HE regression design matrix. The closed form solution is:

$$(x, y) = \begin{cases} (\dfrac{ce - 2bd}{4ab - c^2}, \dfrac{cd - 2ae}{4ab - c^2}), & \text{if } \dfrac{ce - 2bd}{4ab - c^2} \geq 0, \; \dfrac{cd - 2ae}{4ab - c^2} \geq 0; \\ (0, -\dfrac{e}{2b}), & \text{if } \dfrac{ce - 2bd}{4ab - c^2} < 0, \; \dfrac{e}{2b} \leq 0; \\ (-\dfrac{d}{2a}, 0), & \text{if } \dfrac{cd - 2ae}{4ab - c^2} < 0, \; \dfrac{d}{2a} \leq 0; \\ (0, 0), & \text{if } \dfrac{d}{2a} > 0, \; \dfrac{e}{2b} > 0. \end{cases} \tag{S3}$$

REHE confidence interval construction with sparse correlation matrix

When the sample size n is large and the correlation matrices D_k's are dense, constructing REHE confidence intervals is computationally demanding. The main computational burden is sampling from $N_n(0, \Sigma)$ in step (a) of the bootstrap algorithm (see the Online Methods). A common practice to sample $\widetilde{Y}^* \sim N_n(0, \Sigma)$ in step (a) is through $\widetilde{Y}^* = \Sigma^{1/2} e$, where $\Sigma = \Sigma^{1/2}(\Sigma^{1/2})^\top$ and $e \sim N_n(0, I_n)$. $\Sigma^{1/2}$ is computed using Cholesky decomposition or singular value decomposition, which requires $O(n^3)$ operations. One way to speed up this procedure is to construct sparse block-diagonal correlation matrices D_k^*'s to approximate the dense D_k's [2]. In genetics studies, it is reasonable to assume high relatedness within small groups of subjects and low relatedness across groups, which holds for kinship and household membership relatedness. Such relatedness structure can be well-approximated by sparse block-diagonal correlation matrices, through thresholding the off-diagonal entries and rearranging the rows/columns of original dense correlation matrices [2]. The resulting sparse block-diagonal correlation matrices can substantially speed up the matrix decomposition needed in the bootstrap algorithm. Approximate confidence intervals obtained with sparse D_k^*'s are expected to have good coverage as long as each D_k is well approximated by D_k^*.

In simulation studies, we used the function *makeSparseMatrix* in the R package *GENESIS* (*v2.14.3*, [3]) to implement the kinship correlation matrix sparsification procedure. The sparsification threshold was set at $2^{-5.5}$ to make the kinship matrix sparse at the fifth degree relatedness [2].

Supplementary Note 2: Statistical Properties of REHE Estimator

REHE variance component estimator is consistent and asymptotically normal under mild conditions. For expositional clarity, we assume the simple model introduced in the Online Methods:

$$Y = \sigma_0 \boldsymbol{\gamma_0} + \sigma_1 \boldsymbol{\gamma_1},$$

where $\boldsymbol{\gamma_k} \sim N(0, D_k)$, $k = 0,\ 1$. We additionally assume D_1 is sparse and block-diagonal:

$$D_1 = \begin{bmatrix} D_1^{(1)} & 0 & 0 & \dots \\ 0 & D_1^{(2)} & 0 & \dots \\ \vdots & & \ddots & \\ 0 & 0 & 0 & D_1^{(M)} \end{bmatrix},$$

where $D_1^{(m)}$ $(m = 1, \dots, M)$ are square blocks along the diagonal of D_1, and M is the total number of blocks. The following arguments can be easily generalized to models with fixed effect covariates or with more than two variance components. To facilitate the discussion, let s_m denote the number of rows/columns in $D_1^{(m)}$, and $Y^{(m)}$ denote the $s_m \times 1$ subvector of Y corresponding to the m^{th} block. For example, $Y^{(1)}$ is the subvector by taking the first s_1 elements from Y. By construction, $Y^{(m)}$'s are independent and normally distributed with zero mean and covariance $\sigma_0^2 I_{s_m} + \sigma_1^2 D_1^{(m)}$, for $m = 1, \dots, M$.

We first establish consistency and asymptotic normality of the HE estimator. We have shown in the Online Methods that the HE estimate is the ordinary least squares solution to a linear regression problem. For sparse block-diagonal D_1, we can simplify $\widetilde{Y}$ and $\widetilde{X}$ (respectively the outcome and design matrix in the HE regression; see Online Methods) by discarding elements corresponding to zero entries of D_1:

$$\widetilde{Y} = \begin{pmatrix} \widetilde{Y}^{(1)} \\ \widetilde{Y}^{(2)} \\ \vdots \\ \widetilde{Y}^{(M)} \end{pmatrix}, \quad \widetilde{X} = \begin{pmatrix} \widetilde{X}^{(1)} \\ \widetilde{X}^{(2)} \\ \vdots \\ \widetilde{X}^{(M)} \end{pmatrix},$$

where we denote $\widetilde{Y}^{(m)} = \mathrm{vec}(Y^{(m)} Y^{(m)\top})$ and $\widetilde{X}^{(m)} = (\mathrm{vec}(I_n^{(m)}), \mathrm{vec}(D_1^{(m)}))$, for $m = 1, \dots, M$. The independence among $Y^{(m)}$'s implies independence among $\widetilde{Y}^{(m)}$'s. Thus, the HE estimator falls within the framework of generalized estimating equations with I_n being the 'working correlation matrix'. The consistency and asymptotic normality properties of such an estimator are well-established (see [4] Example 2.1 and 5.1). Specifically, as the

number of blocks $M \to \infty$, regardless of whether the maximum block size $s = \max_{1 \le m \le M}(s_m)$ goes to infinity or not, example 2.1 in [4] shows that the HE estimates $\hat{\boldsymbol{\sigma}}^2_{sM} = \left(\hat{\sigma}^2_{0,sM}, \hat{\sigma}^2_{1,sM}\right)$ converge in probability to the true variance components $\boldsymbol{\sigma}^2 = (\sigma^2_0, \sigma^2_1)$ at a rate of at least $\sqrt{M}$. When the maximum block size s does not go to infinity too fast and $M \to \infty$, the HE estimates are asymptotically normal, with a rate of at least $\sqrt{M}$ [4]:

$$W_{sM}^{-1/2} H_{sM}(\hat{\boldsymbol{\sigma}}^2_{sM} - \boldsymbol{\sigma}^2) \xrightarrow{d} N_2(0, I_2),$$

where

$$W_{sM} = \sum_{m=1}^{M} (\widetilde{X}^{(m)})^\top \bar{R}^{(m)} \widetilde{X}^{(m)}, \quad \bar{R}^{(m)} = \mathrm{Cor}(\widetilde{Y}^{(m)}), \quad H_{sM} = \sum_{m=1}^{M} (\widetilde{X}^{(m)})^\top \widetilde{X}^{(m)}.$$

In genetic studies, it is typical that subjects belong to small groups (blocks) where s is bounded and the number of groups (blocks) increases with increasing sample sizes, which satisfy the conditions for HE estimator's consistency and asymptotic normality.

We are now ready to establish the consistency and asymptotic normality of the REHE estimator, $\widetilde{\boldsymbol{\sigma}}^2_{sM} = \left(\widetilde{\sigma}^2_{0,sM}, \widetilde{\sigma}^2_{1,sM}\right)$. As illustrated in (S3), the REHE estimates will only be different from the HE estimates $\hat{\boldsymbol{\sigma}}^2_{sM}$ when HE gives negative estimates for some variance components. Let $\lambda_{sM} = \mathrm{P}(\hat{\sigma}^2_{0,sM} \ge 0,\ \hat{\sigma}^2_{1,sM} \ge 0)$ denote the probability of the HE estimates being nonnegative, which also equals $\mathrm{P}(\widetilde{\boldsymbol{\sigma}}^2_{sM} = \hat{\boldsymbol{\sigma}}^2_{sM})$. We can view the distribution of the REHE estimator as a mixture of two distributions:

$$\begin{aligned}\mathrm{P}(\widetilde{\sigma}^2_{0,sM} < t_0,\ \widetilde{\sigma}^2_{1,sM} < t_1) =& \lambda_{sM} \mathrm{P}(\widetilde{\sigma}^2_{0,sM} < t_0,\ \widetilde{\sigma}^2_{1,sM} < t_1 \mid \widetilde{\boldsymbol{\sigma}}^2_{sM} = \hat{\boldsymbol{\sigma}}^2_{sM}) \\ &+ (1 - \lambda_{sM}) \mathrm{P}(\widetilde{\sigma}^2_{0,sM} < t_0, \widetilde{\sigma}^2_{1,sM} < t_1 \mid \widetilde{\boldsymbol{\sigma}}^2_{sM} \ne \hat{\boldsymbol{\sigma}}^2_{sM}).\end{aligned}$$

As $M \to \infty$, the HE estimates converges in probability to strictly positive $\boldsymbol{\sigma}^2$. Thus the probability of the HE estimates being negative goes to zero, i.e. $1 - \lambda_{sM} \to 0$. This implies that $\mathrm{P}(\widetilde{\sigma}^2_{0,sM} < t_0,\ \widetilde{\sigma}^2_{1,sM} < t_1) \to \mathrm{P}(\hat{\sigma}^2_{0,sM} < t_0,\ \hat{\sigma}^2_{1,sM} < t_1)$, or equivalently the distribution of the REHE estimator converges to the distribution of the HE estimator. Therefore, the REHE estimator $\widetilde{\boldsymbol{\sigma}}^2_{sM}$ shares the asymptotic properties of the HE estimator $\hat{\boldsymbol{\sigma}}^2_{sM}$ as $M \to \infty$. In particular, the REHE estimator are consistent and asymptotically normally distributed under the mild conditions for the HE estimator.

Supplementary Note 3: Additional Simulation Results (Main Paper Simulation Study)

As we have discussed in Supplementary Note 1, correlation matrix sparsification can speed up confidence interval construction with REHE. REML can also benefit computationally from sparsification [2]. We thus applied sparsification to both REHE confidence interval construction and REML, to examine their performances in terms of computation time, mean squared error (MSE), confidence interval coverage and width. *sREML* refers to REML applied with sparsified kinship correlation matrix. *sREHE Wald* and *sREHE Quantile* refers to, respectively, REHE Wald and Quantile confidence intervals based on sparsified kinship matrix. We did not apply sparsification to REHE point estimation. This is because REHE point estimation is very fast even with dense correlation matrices, and we have observed that REHE point estimates are much more accurate than those obtained with sparsification.

Both REHE confidence interval construction and REML based on sparse kinship matrix were substantially faster than their counterparts based on the original kinship matrix (Figure S3). The sparsification procedure itself was computationally demanding, but it eventually showed an advantage over REML when the sample size is large. At $n = 12000$, sparsification with one correlation matrix (one random effect) required only half of the time needed by REML. On the other hand, for a fixed sample size, when there are multiple random effects in the model, time spent on sparsifying multiple correlation matrices will add up, thereby compromising the speed gain by sparsification.

The MSE of point estimates by sREML was always larger than REML, and sometimes could be much larger than MSE of REHE estimates (Figure S4). The coverage of CIs by methods that use sparsification was not as stable: sREHE based CIs showed conservative coverage; sREML based CIs had severe under-coverage at sample size 3000, dropping below 0.75 (Figure S5). This gives a warning sign as sparsification might speed up computation but affect inference accuracy. For the same method, confidence interval width was generally wider with sparsification than without (Figure S6). We found sREML confidence interval width comparable to REHE based intervals.

Supplementary Note 4: Additional Simulation Studies

Simulation setup

We conducted additional simulation studies to compare REHE with REML and HE using artificial kinship matrices D_1 generated under three settings. In the first two settings, we generated block-diagonal kinship matrix D_1 with 3×3 blocks D_b, such that D_1 had the form

$$D_1 = \begin{pmatrix} D_b & 0 & 0 & \dots \\ 0 & D_b & 0 & \dots \\ 0 & 0 & D_b & \\ \vdots & & & \ddots \end{pmatrix}.$$

Setting 1 represents a population with high correlation within groups, where

$$D_b = \begin{pmatrix} 1 & 0.8 & 0.2 \\ 0.8 & 1 & 0.4 \\ 0.2 & 0.4 & 1 \end{pmatrix}.$$

Setting 2 represents a population where subjects are remotely related even within a group, where

$$D_b = \begin{pmatrix} 1 & 0.05 & 0.05 \\ 0.05 & 1 & 0.1 \\ 0.05 & 0.1 & 1 \end{pmatrix}.$$

For setting 3, we generated an approximate block-diagonal kinship matrix: in addition to setting

$$D_b = \begin{pmatrix} 1 & 0.1 & 0.05 \\ 0.1 & 1 & 0.3 \\ 0.05 & 0.3 & 1 \end{pmatrix},$$

we generated random values from $\text{Unif}[-0.001, 0.001]$ to fill the off-diagonal zero entries in D_1. Under setting 3, we tried to mimic the dense kinship matrix in real data applications, which might make estimation and inference more difficult. For the above three different kinship matrices, we ran 200 replicates for each combination of $(\sigma_0^2,\ \sigma_1^2) \in \{(0.1,\ 0.1),\ (0.04,\ 0.1),\ (0.1, 0.04),\ (0.01,\ 0.1),\ (0.1,\ 0.01)\}$ and $n \in \{3000,\ 6000,\ 9000,\ 12000\}$. Kinship matrix sparsification was applied for confidence interval construction with REHE and for REML.

Simulation results

Computation time under different scenarios was very similar to those presented in the main paper and Supplementary Note 3 Figure S3. It is worth pointing out that even when the kinship matrix was already sparse and block-diagonal under setting 1 and 2, the sparsification procedure did not immediately recognize that. It actually required similar computation time compared to sparsifying dense kinship matrices of the same size. This indicates room of improvement for the sparsification procedure.

The MSEs for different methods are presented in Figure S7. We noticed the MSE of sREML was always indistinguishable from REML. REHE estimates were close to REML estimates, and their MSEs often had negligible difference (e.g. Figure S7-b, S7-e, S7-h). When the underlying variance component values were different, improvement of REHE over HE was more pronounced than those illustrated in the main paper and Supplementary Note 3 (Figure S7-e, S7-f). Such improvement in MSE happened when REHE corrected negative values of HE estimates. We counted the frequency of HE giving negative estimates (before being truncated to zero) under each scenario, and observed as high as 45% of the replicates giving negative HE estimates ($n = 3000$, D_1 under setting 2, $(\sigma_0^2,\ \sigma_1^2) \in \{(0.01, 0.1),\ (0.1, 0.01)\}$). Even at the sample size $n = 12000$, HE still gave negative estimates 30% of the time. Large decrease in the MSE of REHE compared to HE was subsequently observed (Figure S7-e, S7-f). With less extreme variance component values $(\sigma_0^2,\ \sigma_1^2) \in \{(0.04, 0.1),\ (0.1, 0.04)\}$, we still observed 15% of HE estimates being negative at the sample size $n = 3000$.

Confidence intervals had close to nominal coverage for all methods under $(\sigma_0^2, \sigma_1^2) = (0.1, 0.1)$, but showed more volatility when the true variance component values were different (Figure S8). Taking into consideration the Monte Carlo error of 0.03 from 200 replicates, REHE Quantile confidence interval always had acceptable coverage around 0.95. The coverage of REHE Wald type confidence intervals was generally close to 0.95, but could be slightly off under some scenarios (Figure S8-c, S8-f, S8-i). This suggests REHE Quantile interval is more robust than Wald type interval. REHE confidence intervals with sparsification had similar patterns of coverage. It has been observed in the main paper and Supplementary Note 3 that REML and sREML could have poor coverage under some scenarios. While the coverage improved quickly with increasing sample size in those cases, the under-coverage could persist even under large sample sizes (Figure S8-c, S8-e, S8-h). With $(\sigma_0^2, \sigma_1^2) = (0.01, 0.1)$ and kinship matrix structure under setting 2 or 3, coverage of REML/sREML was lower than 0.7 even with $n = 12000$ samples (Figure S8-e, S8-h). In addition, it is worth noting that REML may fail to produce a confidence interval due to computational issues, which happens when one variance component estimate reaches the zero boundary during the iterative updates. With unequal true variance component values and small sample sizes, this issue occurred more than 45% of the time. Under setting 2 with $(\sigma_0^2, \sigma_1^2) = (0.01, 0.1)$, the proportion of replicates where REML was unable to compute a confidence interval was over 28% even at a sample size of $n = 12000$. REHE confidence intervals do not suffer from this issue and can always be obtained.

Figure S9 presents distributions of confidence interval half width. REHE was comparable to REML: REML tends to have narrower confidence intervals when true values of variance components were not too different (Figure S9-a, S9-d, S9-g); otherwise REHE intervals were narrower (Figure S9-c, S9-f, S9-i). As we have seen scenarios with poor coverage for REML and sREML confidence intervals (Figure S8-e, S8-h), corresponding interval width had suspiciously skewed distributions (Figure S9-e, S9-h).

References

[1] Sofer, T. Confidence intervals for heritability via Haseman-Elston regression. *Statistical Applications in Genetics and Molecular Biology* **16**, 259–273 (2017).

[2] Gogarten, S. M. *et al.* Genetic association testing using the genesis r/bioconductor package. *Bioinformatics* **35**, 5346–5348 (2019).

[3] Conomos, M. P. *et al. GENESIS: GENetic EStimation and Inference in Structured samples (GENESIS): Statistical methods for analyzing genetic data from samples with population structure and/or relatedness* (2019). URL `https://github.com/UW-GAC/GENESIS`. R package version 2.14.3.

[4] Xie, M. & Yang, Y. Asymptotics for generalized estimating equations with large cluster sizes. *The Annals of Statistics* **31**, 310–347 (2003).

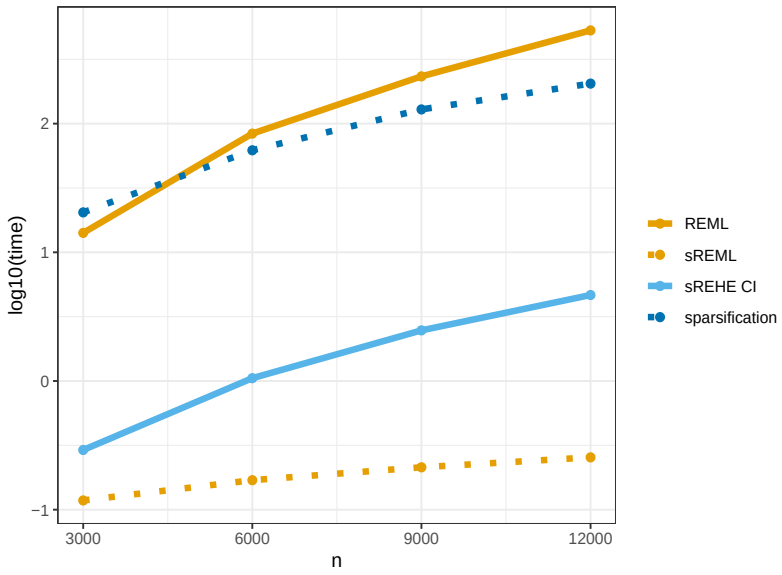

Figure S3: Additional results for the simulation study described in the main paper and Online Methods: Computation time for REML and REHE based on matrix sparsification, benchmarked with computation time for REML. We separately illustrate time for sparsifying the kinship matrix (sparsification), time for fitting the model with REML after obtaining the sparsified kinship matrix (sREML), and time for REHE confidence interval construction based on sparsified kinship matrix (sREHE CI).

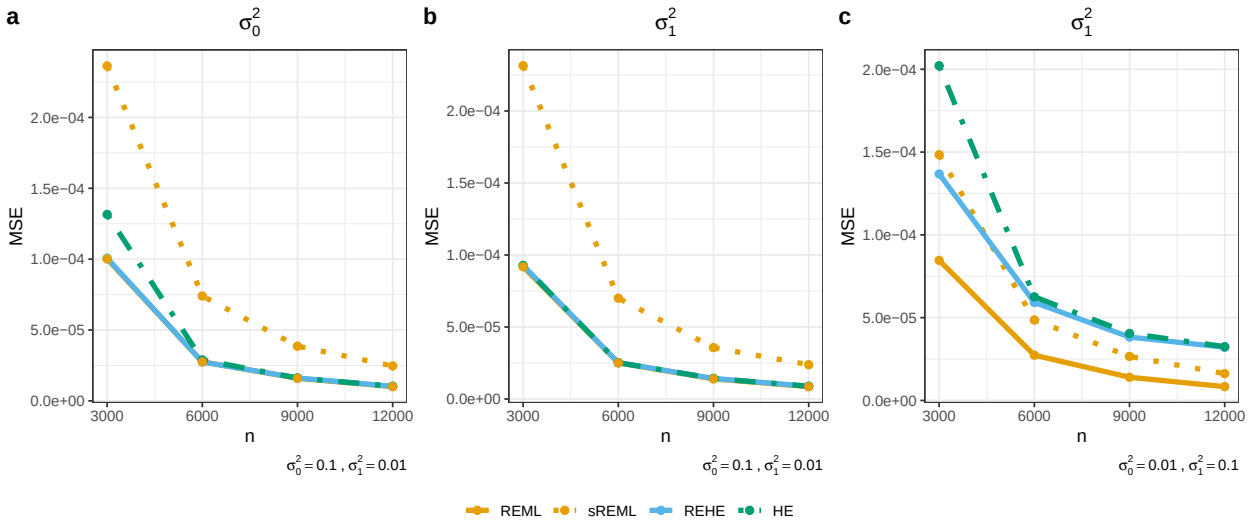

Figure S4: Additional results for the simulation study described in the main paper and Online Methods: MSE of point estimates by REML, REML with matrix sparsification (sREML), REHE and HE. True values of variance components under each scenario were presented on the bottom-right corner of each sub-plot.

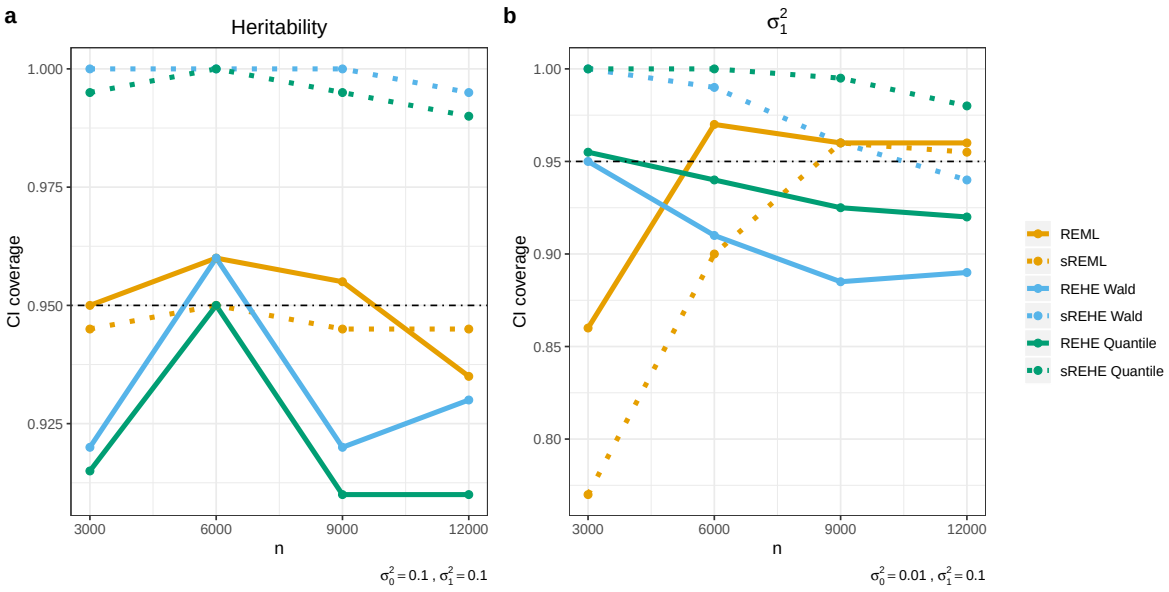

Figure S5: Additional results for the simulation study described in the main paper and Online Methods: confidence interval coverage with 200 replicates. Monte Carlo error of 0.03 is expected for 200 simulation replicates. We present results for confidence intervals constructed based on REML, REML with matrix sparsification (sREML), Wald type intervals with REHE (REHE Wald), Wald type intervals with REHE and matrix sparsification (sREHE Wald), quantile type intervals with REHE (REHE Quantile), and quantile type intervals with REHE and matrix sparsification (sREHE Quantile). True values of variance components under each scenario were presented on the bottom-right corner of each sub-plot.

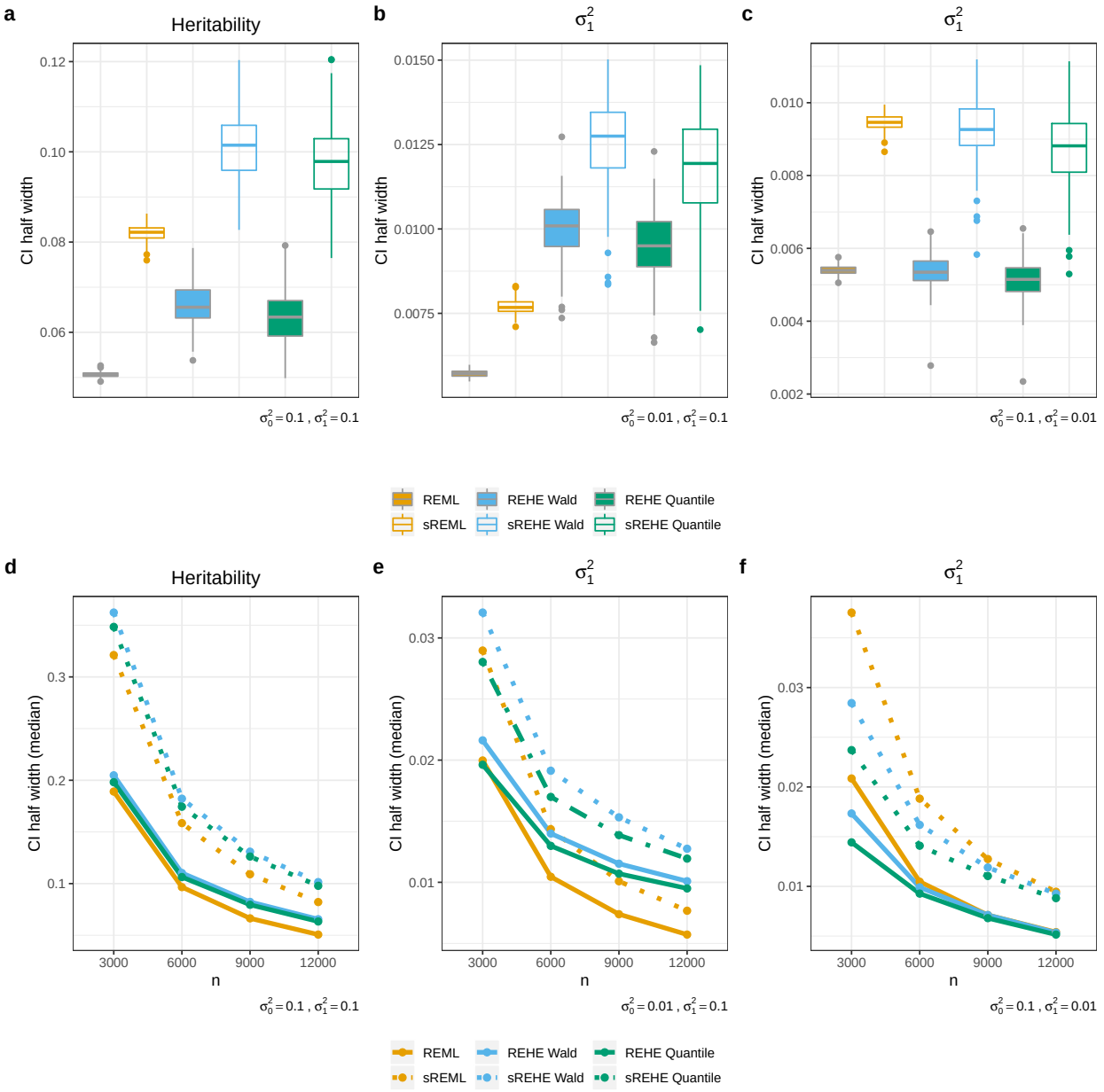

Figure S6: Additional results for the simulation study described in the main paper and Online Methods: boxplots of confidence interval half width, and line charts of median confidence interval half width against sample sizes. The boxplots follow standard Tukey representations, where center line represents median, box limits represent lower and upper quartiles, whiskers represent 1.5x interquartile range, and points represent outliers. We present results for confidence intervals constructed based on REML, REML with matrix sparsification (sREML), Wald type intervals with REHE (REHE Wald), Wald type intervals with REHE and matrix sparsification (sREHE Wald), quantile type intervals with REHE (REHE Quantile), and quantile type intervals with REHE and matrix sparsification (sREHE Quantile). True values of variance components under each scenario were presented on the bottom-right corner of each sub-plot.

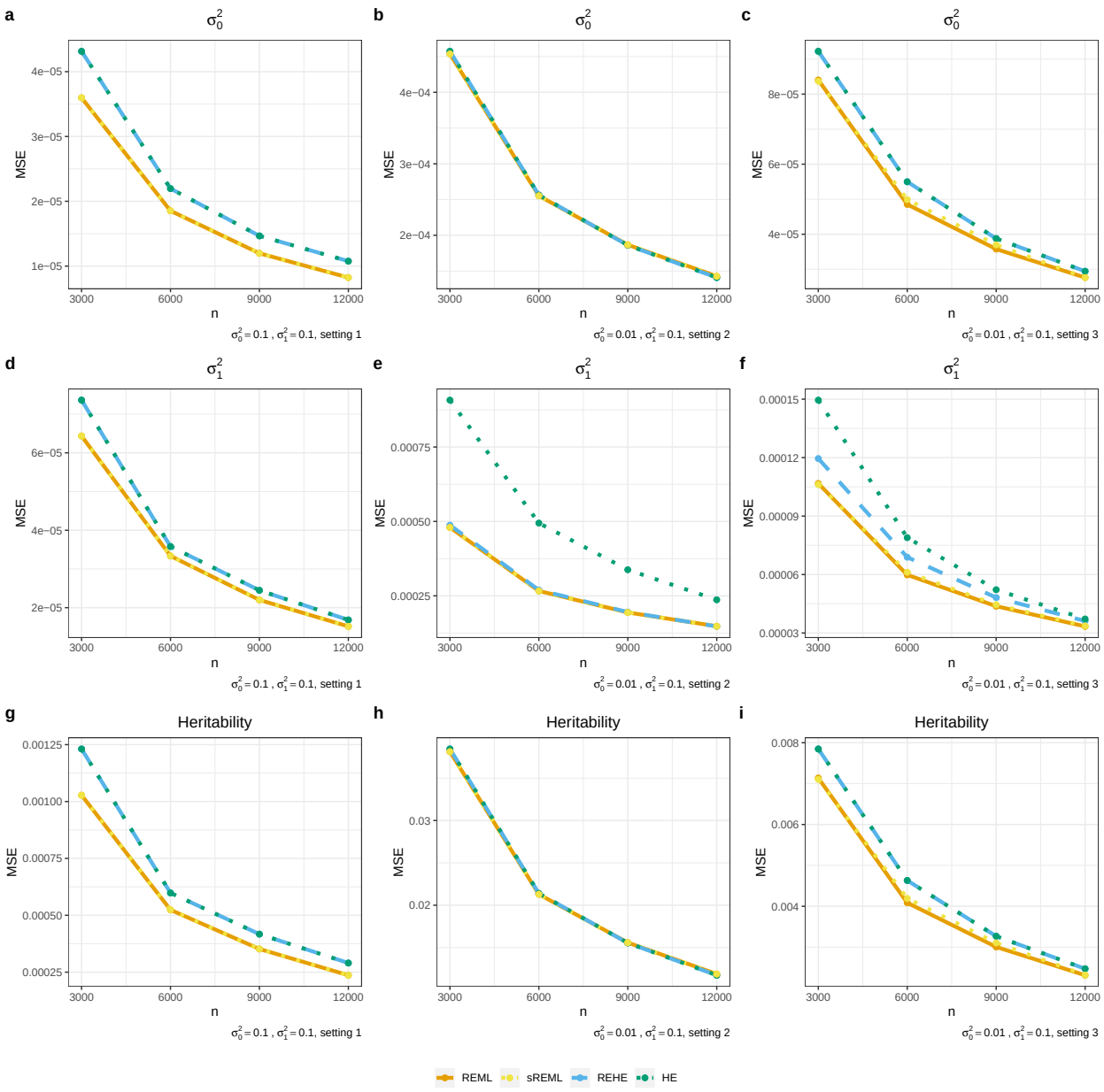

Figure S7: MSE of estimates for σ_0^2, σ_1^2 and heritability, based on REML, REML with matrix sparsification (sREML), REHE and HE. Parameter true values for σ_0^2 and σ_1^2, and kinship matrix structure setting under each scenario are indicated on the bottom-left corner of each subplot.

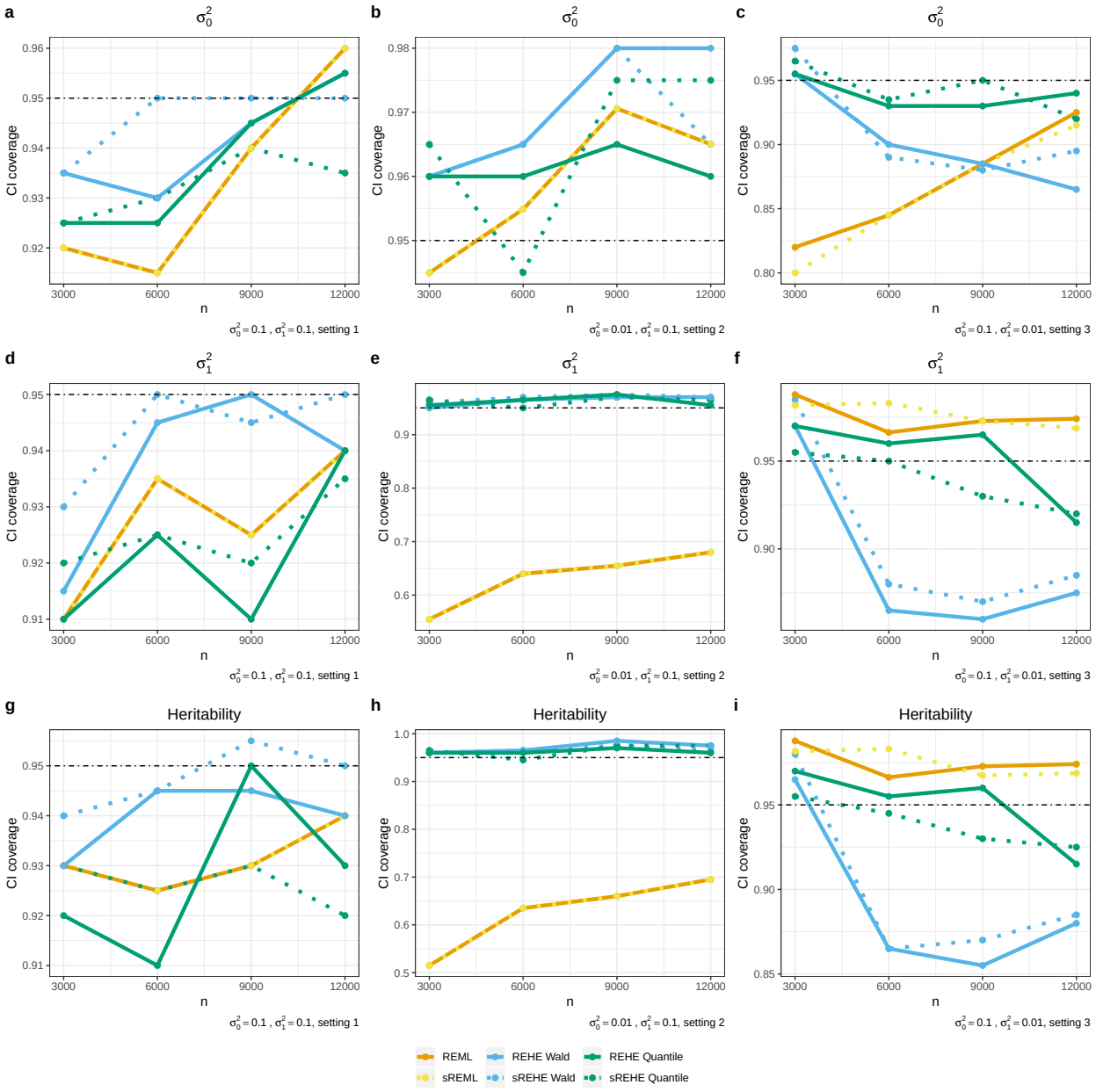

Figure S8: Confidence interval coverage for σ_0^2, σ_1^2 and heritability with 200 replicates. Monte Carlo error of 0.03 is expected for 200 simulation replicates. We present results for confidence intervals constructed based on REML, REML with matrix sparsification (sREML), Wald type intervals with REHE (REHE Wald), Wald type intervals with REHE and matrix sparsification (sREHE Wald), quantile type intervals with REHE (REHE Quantile), and quantile type intervals with REHE and matrix sparsification (sREHE Quantile). Parameter true values for σ_0^2 and σ_1^2, and kinship matrix structure setting under each scenario are indicated on the bottom-left corner of each subplot.

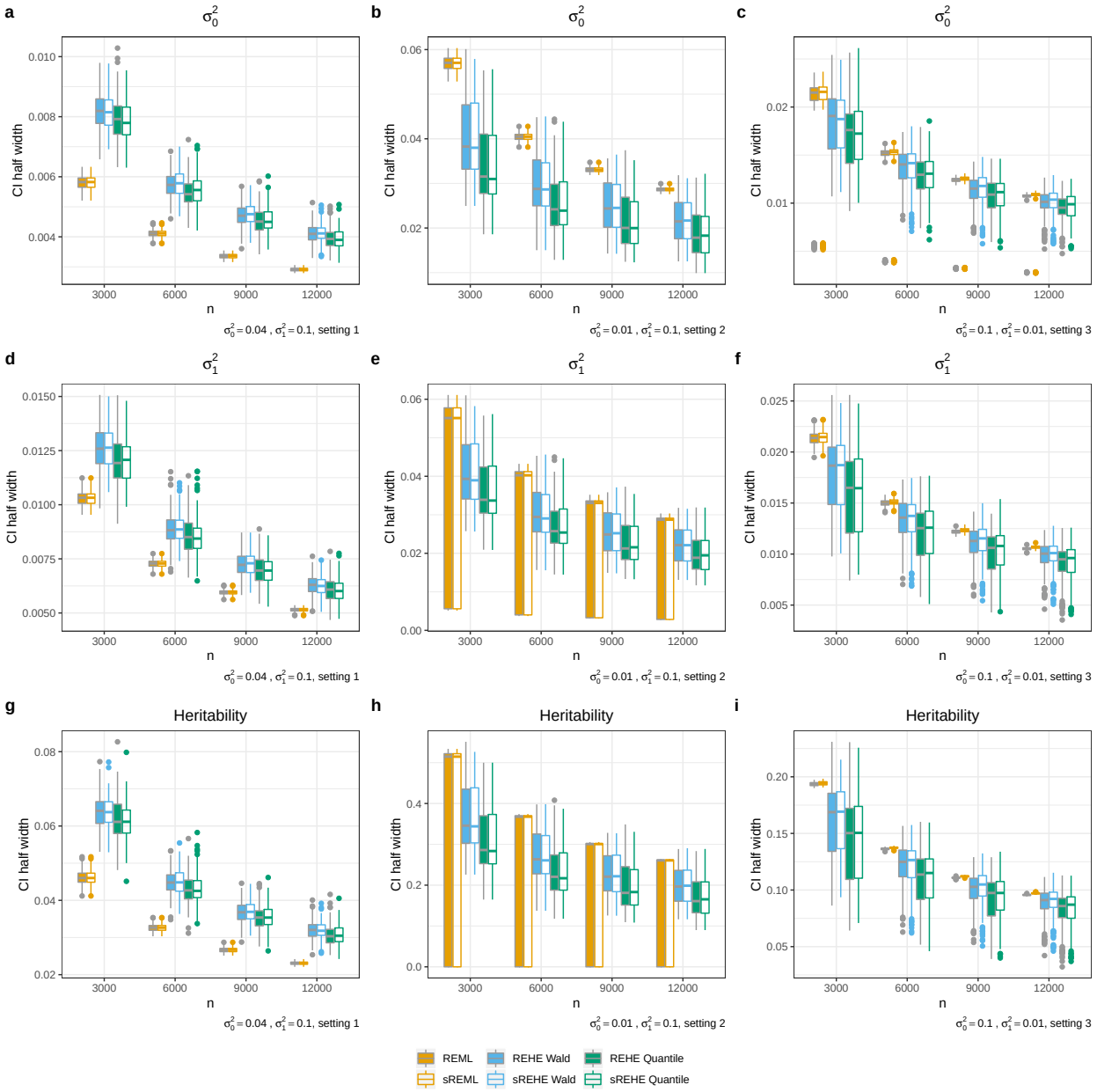

Figure S9: Boxplots of confidence interval half width for σ_0^2, σ_1^2 and heritability. The boxplots follow standard Tukey representations, where center line represents median, box limits represent lower and upper quartiles, whiskers represent 1.5x interquartile range, and points represent outliers. We present results for confidence intervals constructed based on REML, REML with matrix sparsification (sREML), Wald type intervals with REHE (REHE Wald), Wald type intervals with REHE and matrix sparsification (sREHE Wald), quantile type intervals with REHE (REHE Quantile), and quantile type intervals with REHE and matrix sparsification (sREHE Quantile). Parameter true values for σ_0^2 and σ_1^2, and kinship matrix structure setting under each scenario are indicated on the bottom-left corner of each subplot.