

**Supplementary material to “Efficient inference of  
parental origin effects using case-control mother-child  
genotype data” by**

Yuang Tian<sup>1</sup>, Hong Zhang<sup>2\*</sup>, Alexandre Bureau<sup>3</sup>, Hagit Hochner<sup>4</sup>,  
and Jinbo Chen<sup>5</sup>

<sup>1</sup>Shanghai Center for Mathematical Sciences, Fudan University, Shanghai, China

<sup>2</sup>Department of Statistics and Finance, School of Management, University of Science and Technology of China, Hefei, Anhui, China

<sup>3</sup>Department of social and preventive medicine, Université Laval, Québec, Canada

<sup>4</sup>Braun School of Public Health, The Hebrew University of Jerusalem, Jerusalem, Israel

<sup>5</sup>Department of Biostatistics and Epidemiology, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, U.S.A.

**Appendix A. Derivation of the log profile likelihood function (10)  
in the main text.**

We adopt the notations in the main text. The aim is to maximize the log empirical likelihood function

$$\log L(\boldsymbol{\Theta}, \boldsymbol{\pi}) = l_1(\boldsymbol{\Theta}) + \sum_{u=1}^n \log \pi_u$$

under the constraints

$$\sum_{u=1}^n L_u(\boldsymbol{\Theta}) \pi_u = f \tag{S1}$$

and

$$\sum_{u=1}^n \pi_u = 1. \tag{S2}$$

According to the Lagrange multiplier method, the above maximization problem is equivalent to maximizing

$$M(\boldsymbol{\Theta}, \boldsymbol{\pi}, \zeta, \lambda) = l_1(\boldsymbol{\Theta}) + \sum_{u=1}^n \log \pi_u - \zeta \left( \sum_{u=1}^n \pi_u - 1 \right) - n\lambda \left( \sum_{u=1}^n L_u(\boldsymbol{\Theta})\pi_u - f \right) \quad (\text{S3})$$

with respect to  $\boldsymbol{\Theta}$ ,  $\boldsymbol{\pi}$ ,  $\zeta$ , and  $\lambda$ . The maximizer of (S3) satisfies

$$\frac{\partial M}{\partial \pi_u}(\boldsymbol{\Theta}, \boldsymbol{\pi}, \zeta, \lambda) = 0, \quad (\text{S4})$$

or equivalently,

$$1 = \zeta \pi_u + n\lambda L_u(\boldsymbol{\Theta})\pi_u. \quad (\text{S5})$$

Summing the above equation over  $u = 1, \dots, n$ , we have

$$\zeta = n(1 - \lambda f). \quad (\text{S6})$$

Plugging this into (S5), we have

$$\pi_u = \frac{1}{n \{1 + \lambda (L_u(\boldsymbol{\Theta}) - f)\}}. \quad (\text{S7})$$

Plugging (S7) into (S1) and combining (S2), we obtain the following constraint on the multiplier  $\lambda$ :

$$\sum_{u=1}^n \frac{L_u(\boldsymbol{\Theta}) - f}{1 + \lambda (L_u(\boldsymbol{\Theta}) - f)} = 0. \quad (\text{S8})$$

Plugging (S7) and (S6) into (S3), we obtain the resulting log profile likelihood (10) in the main text, with  $\lambda$  satisfying the constraint (S8).

## Appendix B. Proof of Lemma 1 in the main text.

We need to prove

$$E \left[ \frac{\partial l_{\text{mp}}(\boldsymbol{\Theta})}{\partial \boldsymbol{\Theta}} \right] \Big|_{\boldsymbol{\Theta}=\boldsymbol{\Theta}_0} = 0, \quad (\text{S9})$$

under conditions (C2) in the main text.

Denote  $Z = (G^m, G^c, X)$ ,  $Z_u = (G_u^m, G_u^c, X_u)$ ,  $W = (g^m, X)$ ,  $W_u = (g_u^m, X_u)$ ,  $p_Z(Z; \boldsymbol{\Theta}) = P_{\boldsymbol{\Theta}}(Y = 1|Z)$ ,  $p_X(W; \boldsymbol{\Theta}) = P_{\boldsymbol{\Theta}}(Y = 1|W)$ . Let  $G$  denote a vector containing all genotypes except  $g^m$ , so that  $Z = (G, W)$ . For any measurable function  $t$ , let  $E\{t(Z)|Y = 1\}$

and  $E\{t(Z)|Y = 0\}$  be denoted by  $E_1\{t(Z)\}$  and  $E_0\{t(Z)\}$ , respectively. Furthermore, let  $E\{t(Z)\}$  denote the expectation of  $t(Z)$  in the population where cases and controls arise.

We assume that the expectations appearing in the following context exist and are finite, as will be proved at the end of this section under condition (C2).

For any measurable function  $t$ , it can be immediately checked that the following equations hold true:

$$E_1\{t(Z)\} = E\left\{\frac{p_Z(Z; \Theta_0)}{f}t(Z)\right\} \text{ and } E_0\{t(Z)\} = E\left\{\frac{1 - p_Z(Z; \Theta_0)}{1 - f}t(Z)\right\}, \quad (\text{S10})$$

$$E\left\{\sum_{u=1}^n t(Z_u)\right\} = nE\left[\{1 + \lambda_0(p_Z(Z; \Theta_0) - f)\}t(Z)\right], \quad (\text{S11})$$

$$E\left\{\sum_{u=1}^n \frac{(2Y_u - 1)t(Z_u)}{1 - Y_u + (2Y_u - 1)p_Z(Z_u; \Theta_0)}\right\} = n\lambda_0 E\{t(Z)\}. \quad (\text{S12})$$

Note that

$$\begin{aligned} l_{\text{mp}}(\Theta) &= \sum_{u=1}^n \log P_{\Theta}(Y_u|Z_u) + \sum_{u=1}^n \log P_{\Theta}(Z_u|W_u) - \sum_{u=1}^n \log [n\{1 + \lambda_0(p_W(W_u; \Theta) - f)\}] \\ &:= \ell^{(1)}(\Theta) + \ell^{(2)}(\Theta) - \ell^{(3)}(\Theta). \end{aligned} \quad (\text{S13})$$

In view of (S12), we have that

$$E\left\{\frac{\partial}{\partial \Theta} \ell^{(1)}(\Theta_0)\right\} = E\left\{\sum_{u=1}^n \frac{(2Y_u - 1) \partial p_Z(Z_u; \Theta_0) / \partial \Theta}{1 - Y_u + (2Y_u - 1)p_Z(Z_u; \Theta_0)}\right\} = n\lambda_0 E\left\{\frac{\partial p_Z(Z; \Theta_0)}{\partial \Theta}\right\}. \quad (\text{S14})$$

In view of (S10), we have that

$$E\left\{\frac{\partial}{\partial \Theta} \ell^{(2)}(\Theta_0)\right\} = n_1 E\left\{\frac{p_Z(Z; \Theta_0)}{f} \frac{\partial}{\partial \Theta} \log P_{\Theta_0}(Z|W)\right\} \quad (\text{S15})$$

$$\begin{aligned} &+ n_0 E\left\{\frac{1 - p_Z(Z; \Theta_0)}{1 - f} \frac{\partial}{\partial \Theta} \log P_{\Theta_0}(Z|W)\right\} \\ &= \frac{n_0}{1 - f} E\left\{\frac{\partial}{\partial \Theta} \log P_{\Theta_0}(Z|W)\right\} + n\lambda_0 E\left\{p_Z(Z; \Theta_0) \frac{\partial}{\partial \Theta} \log P_{\Theta_0}(Z|W)\right\} \\ &= 0 + n\lambda_0 E\left\{P_{\Theta_0}(Y = 1|G, W) \frac{\partial}{\partial \Theta} \log P_{\Theta_0}(G|W)\right\} \\ &= n\lambda_0 E\left\{\sum_{g \in \mathcal{G}} P_{\Theta_0}(Y = 1|G = g, W) \frac{\partial}{\partial \Theta} P_{\Theta_0}(G = g|W)\right\}, \end{aligned} \quad (\text{S16})$$

where  $\mathcal{G}$  is the set of all possible values of  $G$ . In view of (S11)-(S16), we have that

$$\begin{aligned} E \left\{ \frac{\partial}{\partial \Theta} \ell^{(3)}(\Theta_0) \right\} &= \lambda_0 E \left[ \sum_{u=1}^n \frac{\partial p_W(W_u; \Theta_0) / \partial \Theta}{1 + \lambda_0 \{p_W(W_u; \Theta_0) - f\}} \right] \\ &= n \lambda_0 E \left[ \frac{1 + \lambda_0 \{p_Z(Z; \Theta_0) - f\}}{1 + \lambda_0 \{p_W(W; \Theta_0) - f\}} \frac{\partial p_W(W; \Theta_0)}{\partial \Theta} \right] \\ &= n \lambda_0 E \left[ \frac{\partial p_W(W; \Theta_0)}{\partial \Theta} \right] \end{aligned} \quad (\text{S17})$$

$$\begin{aligned} &= n \lambda_0 E \left\{ \frac{\partial}{\partial \Theta} \sum_{g \in \mathcal{G}} P_{\Theta_0}(Y = 1 | G = g, W) P_{\Theta_0}(G = g | W) \right\} \\ &= n \lambda_0 E \left\{ \sum_{g \in \mathcal{G}} \frac{\partial}{\partial \Theta} P_{\Theta_0}(Y = 1 | G = g, W) P_{\Theta_0}(G = g | W) \right. \\ &\quad \left. + \sum_{g \in \mathcal{G}} P_{\Theta_0}(Y = 1 | G = g, W) \frac{\partial}{\partial \Theta} P_{\Theta_0}(G = g | W) \right\} \\ &= E \left\{ \frac{\partial}{\partial \Theta} \ell^{(1)}(\Theta_0) \right\} + E \left\{ \frac{\partial}{\partial \Theta} \ell^{(2)}(\Theta_0) \right\}, \end{aligned} \quad (\text{S18})$$

which is equivalent to (S9) according to (S13).

Now we show that the above expectations exist and are finite. First,

$$\begin{aligned} E \left\{ \left| \sum_{u=1}^n t(Z_u) \right| \right\} &\leq n E \left[ \{1 + \lambda_0 (p_Z(Z; \Theta_0) - f)\} |t(Z)| \right], \\ E \left\{ \left| \sum_{u=1}^n \frac{(2Y_u - 1) t(Z_u)}{1 - Y_u + (2Y_u - 1) p_Z(Z_u; \Theta_0)} \right| \right\} &\leq \left( \frac{n_1}{f} + \frac{n_0}{1 - f} \right) E \{|t(Z)|\}. \end{aligned}$$

Then, following the process we went through before, we have

$$\begin{aligned} E \left\{ \left| \frac{\partial}{\partial \Theta} \ell^{(1)}(\Theta_0) \right| \right\} &\leq \left( \frac{n_1}{f} + \frac{n_0}{1 - f} \right) E \left\{ \left| \frac{\partial p_Z(Z; \Theta_0)}{\partial \Theta} \right| \right\}, \\ E \left\{ \left| \frac{\partial}{\partial \Theta} \ell^{(2)}(\Theta_0) \right| \right\} &\leq n \lambda_0 E \left\{ \left| \frac{\partial}{\partial \Theta} \log P_{\Theta_0}(Z | W) \right| \right\}, \end{aligned}$$

and

$$E \left\{ \left| \frac{\partial}{\partial \Theta} \ell^{(3)}(\Theta_0) \right| \right\} \leq n \lambda_0 \left( E \left\{ \left| \frac{\partial p_Z(Z; \Theta_0)}{\partial \Theta} \right| \right\} + E \left\{ \left| \frac{\partial}{\partial \Theta} \log P_{\Theta_0}(Z | W) \right| \right\} \right).$$

Under condition (C2), the above expectations are finite. Consequently,

$$E \left[ \left| \frac{\partial \ell_{\text{mp}}(\Theta_0)}{\partial \Theta} \right| \right] \leq E \left\{ \left| \frac{\partial}{\partial \Theta} \ell^{(1)}(\Theta_0) \right| \right\} + E \left\{ \left| \frac{\partial}{\partial \Theta} \ell^{(2)}(\Theta_0) \right| \right\} + E \left\{ \left| \frac{\partial}{\partial \Theta} \ell^{(3)}(\Theta_0) \right| \right\} < \infty.$$

According to the Dominated Convergence Theorem, all relevant expectations shown above exist and are finite under condition (C2).

## Appendix C. Proof of Theorem 1 in the main text.

First we prove (i), i.e.,  $l_{\text{mp}}(\Theta)$  takes the value of negative infinity on the boundary of the parameter space. When  $\|\beta\| \rightarrow \infty$ , with probability tend to 1, there exists a family such that each compatible haplotype combination  $h_{iu}^m h_{ij}^m, h_{wu}^c, h_{lu}^c$  satisfies  $P(Y_u | h_{iu}^m h_{ij}^m, h_{wu}^c, h_{lu}^c, X_u) \rightarrow 0$ , which is equivalent to  $l_{\text{mp}}(\Theta) \rightarrow -\infty$ . Then we examine the performance of  $l_{\text{mp}}(\Theta)$  when  $\mu_s \rightarrow 0$  for some  $s$ . Note that the true value of  $\mu_s$  is not equal to 0, so with probability tend to 1 we can find a family that both mother and child have diplotype  $h_s h_s$ . For that family we have  $H_u^m = \{h_{su}^m h_{su}^m\}$ ,  $H_u^c = \{(h_{su}^c, h_{su}^c)\}$  and  $P(h_{su}^m h_{su}^m, h_{su}^c, h_{su}^c) \rightarrow 0$ , which means  $l_{\text{mp}}(\Theta) \rightarrow -\infty$ . Under condition (C1), the maximizer of  $l_{\text{mp}}(\Theta)$  is an interior point of the parameter space, which ensures that there exists a solution  $\hat{\Theta}_{\text{mp}}$  in the interior of the parameter space.

Then we prove (2), i.e.,  $\hat{\Theta}_{\text{mp}}$  is consistent for  $\Theta_0$ . According to Lemma 1 and conditions (C4)-(C5), for any sufficiently small  $a > 0$ , we have

$$E[l_{\text{mp}}(\Theta_0)] > \sup\{E[l_{\text{mp}}(\Theta)] : \Theta \in \partial B_a\},$$

where  $B_a = \{\Theta : \|\Theta - \Theta_0\| \leq a\}$  and  $\partial B_a$  is the boundary of  $B_a$ . It follows that

$$P\left(\liminf_{n \rightarrow \infty} \{l_{\text{mp}}(\Theta_0) - l_{\text{mp}}(\Theta) : \Theta \in \partial B_a\} > 0\right) = 1$$

according to the Strong Law of Large Number, which means there exists a solution to the score function in  $B_a$  with a large  $n$ . Using the standard theory for Z-statistic, we can find  $\hat{\Theta}_{\text{mp}}$ , a series of solutions to the score equation, which strongly consistent for  $\Theta_0$ .

Next we prove (3), i.e.,  $\hat{\Theta}_{\text{mp}}$  is asymptotically normal. Under condition (C6) and the consistence of  $\hat{\Theta}_{\text{mp}}$ , Taylor's series expansion gives that

$$\frac{\partial l_{\text{mp}}(\hat{\Theta}_{\text{mp}})}{\partial \Theta} = \frac{\partial l_{\text{mp}}(\Theta_0)}{\partial \Theta} + \frac{\partial^2 l_{\text{mp}}(\Theta_0)}{\partial \Theta \partial \Theta^\tau} (\hat{\Theta}_{\text{mp}} - \Theta_0) + (\hat{\Theta}_{\text{mp}} - \Theta_0) o_p(n). \quad (\text{S19})$$

We have  $\partial l_{\text{mp}}(\hat{\Theta}_{\text{mp}})/\partial \Theta = 0$  because  $\hat{\Theta}_{\text{mp}}$  is an interior point solution of  $\partial l_{\text{mp}}(\Theta)/\partial \Theta = 0$ . Under conditions (C3) and (C5), we have  $\partial^2 l_{\text{mp}}(\Theta_0)/\partial \Theta \partial \Theta^\tau = n A_{\text{mp}}(\Theta_0) + o_p(n)$ . Then (S19) can be rewritten as

$$\sqrt{n} (A_{\text{mp}}(\Theta_0) + o_p(1)) (\hat{\Theta}_{\text{mp}} - \Theta_0) = -\frac{1}{\sqrt{n}} \frac{\partial l_{\text{mp}}(\Theta_0)}{\partial \Theta}. \quad (\text{S20})$$

Under conditions (C3) and (C7), Levy's Central Limit Theorem gives

$$\frac{1}{\sqrt{n}} \frac{\partial l_{\text{mp}}(\Theta_0)}{\partial \Theta} \xrightarrow{d} N\{0, \Sigma_{\text{mp}}(\Theta_0)\}. \quad (\text{S21})$$

Then, combining (S20) and (S21), we have

$$\sqrt{n}(\hat{\Theta}_{\text{mp}} - \Theta_0) \xrightarrow{d} N\{0, A_{\text{mp}}^{-1}(\Theta_0) \Sigma_{\text{mp}}(\Theta_0) A_{\text{mp}}^{-1}(\Theta_0)\}.$$

Finally we show (4), i.e.,  $\hat{\Theta}_{\text{mp}}$  is asymptotically efficient. It suffices to show that the profile MLE  $\hat{\Theta}_p$  and the modified profile MLE  $\hat{\Theta}_{\text{mp}}$  are asymptotic equivalent since the former is asymptotically efficient. Let the corresponding score functions be denoted by

$$U_p(\Theta) = \frac{\partial l_p(\Theta)}{\partial \Theta} = \frac{\partial l_0(\Theta, \lambda)}{\partial \Theta} \Big|_{\lambda=\lambda(\Theta)} \text{ and } U_{\text{mp}}(\Theta) = \frac{\partial l_0(\Theta, \lambda_0)}{\partial \Theta},$$

and denote

$$S_{\text{mp}}(\Theta) = \frac{\partial^2 l_0(\Theta, \lambda_0)}{\partial \Theta \partial \Theta^\tau}.$$

Note that the maximizer of the profile likelihood  $(\hat{\Theta}_p, \lambda(\hat{\Theta}_p))$  is consistent for  $(\Theta_0, \lambda_0)$ , which can be proved similar to Appendix B. Consequently, we have

$$n^{1/2} (\hat{\Theta}_p - \Theta_0) = -\{S_{\text{mp}}(\Theta_0)/n\}^{-1} n^{-1/2} U_p(\Theta_0) \{1 + o_p(1)\} \quad (\text{S22})$$

and

$$n^{1/2} (\hat{\Theta}_{\text{mp}} - \Theta_0) = -\{S_{\text{mp}}(\Theta_0)/n\}^{-1} n^{-1/2} U_{\text{mp}}(\Theta_0) \{1 + o_p(1)\}. \quad (\text{S23})$$

Using Taylor's series expansion, we have that

$$n^{-1/2} \{U_p(\Theta_0) - U_{\text{mp}}(\Theta_0)\} = \frac{1}{n} \frac{\partial^2 l_0(\Theta_0, \lambda_0)}{\partial \Theta \partial \lambda^\tau} n^{1/2} \{\lambda(\Theta_0) - \lambda_0\} (1 + o_p(1)). \quad (\text{S24})$$

Condition (C8) and the Weak Law of Large Numbers give that

$$\frac{1}{n} \frac{\partial^2 l_0(\Theta_0, \lambda_0)}{\partial \Theta \partial \lambda^\tau} \rightarrow 0 \text{ in probability.} \quad (\text{S25})$$

It follows from (S24) and (S25) that

$$n^{-1/2} \{U_p(\Theta_0) - U_{\text{mp}}(\Theta_0)\} = o_p(1). \quad (\text{S26})$$

Now we establish the asymptotic equivalence between  $\hat{\Theta}_{\text{mp}}$  and  $\hat{\Theta}_p$ :

$$n^{-1/2} \{\hat{\Theta}_{\text{mp}} - \hat{\Theta}_p\} = o_p(1) \quad (\text{S27})$$

in view of (S22), (S23), and (S26). According to the asymptotically efficiency of the standard profile MLE  $\hat{\Theta}_p$  [2], we have the asymptotically efficiency of  $\hat{\Theta}_{\text{mp}}$ .

## Appendix D. A simulation study for checking the regularity condition (C8) in the main text

We conducted a simulation study to check the rationality of regularity condition (C8) in the main text, that is,

$$E \{ \partial^2 l_0(\Theta_0, \lambda_0) / \partial \Theta \partial \lambda^\tau \} = 0.$$

We used the sample mean  $n^{-1} \partial^2 l_0(\Theta_0, \lambda_0) / \partial \Theta \partial \lambda^\tau$  to estimate the expectation. We generated data in the way described in Section 3.1 in the main text, with  $\beta_{g^m} = \log(1.8)$ ,  $\beta_{g^c} = \log(1.5)$ ,  $\beta_X = \log(1.8)$ ,  $\beta_{im} = 0$  or  $\beta_{im} = \log(1.5)$ , and  $\eta = 0$ . The simulation results with different sample size are summarized in Table S1, which were based on 10,000 replicates of simulations. For various sample sizes, the estimates are very close to 0, which indicates that the assumption that  $E \{ \partial^2 l_0(\Theta_0, \lambda_0) / \partial \Theta \partial \lambda^\tau \} = 0$  approximately holds true in the simulation situation. We also tried some other parameter combinations and reached the same conclusion (results not shown).

## Appendix E. Proof of Theorem 2 in the main text

The idea of proof follows from [4] and [3]. Let  $g_u^m$  be the target SNP genotype extracted from the joint genotype  $G_u^m$ , then we have

$$P_{\Theta}(Y_u, G_u^c, G_u^m, h_u | g_u^m, X_u) = P_{\Theta}(h_u | Y_u, G_u^c, G_u^m, X_u) P_{\Theta}(Y_u, G_u^c, G_u^m | g_u^m, X_u)$$

for any haplotype  $h_u$ . Consequently, we can decompose  $l_1(\Theta)$  as follows:

$$\begin{aligned} \sum_{u=1}^n \log P_{\Theta}(Y_u, G_u^c, G_u^m | g_u^m, X_u) &= \sum_{u=1}^n \log P_{\Theta}(Y_u, G_u^c, G_u^m, h_u | g_u^m, X_u) \\ &\quad - \sum_{u=1}^n \log P_{\Theta}(h_u | Y_u, G_u^c, G_u^m, X_u) \end{aligned}$$

with  $h_u$  being any haplotype, which implies that

$$\begin{aligned} &\sum_{u=1}^n \log P_{\Theta}(Y_u, G_u^c, G_u^m | g_u^m, X_u) \\ &= E \left[ \sum_{u=1}^n \log P_{\Theta}(Y_u, G_u^c, G_u^m, H_u | g_u^m, X_u) \middle| \Theta', D \right] - E \left[ \sum_{u=1}^n \log P_{\Theta}(H_u | Y_u, G_u^c, G_u^m, X_u) \middle| \Theta', D \right] \end{aligned}$$

for any given  $\Theta'$ , where  $D = \{(Y_u, G_u^c, G_u^m, X_u), u = 1, \dots, n\}$ . As a result,

$$\begin{aligned} l_{\text{mp}}(\Theta) &= \sum_{u=1}^n \log P_{\Theta}(Y_u, G_u^c, G_u^m | g_u^m, X_u) - l_2(\Theta, \lambda_0) \\ &= \left( E \left[ \sum_{u=1}^n \log P_{\Theta}(Y_u, G_u^c, G_u^m, H_u | g_u^m, X_u) \middle| \Theta', D \right] - l_2(\Theta, \lambda_0) \right) \\ &\quad - E \left[ \sum_{u=1}^n \log P_{\Theta}(H_u | Y_u, G_u^c, G_u^m, X_u) \middle| \Theta', D \right] \\ &:= Q(\Theta, \Theta') - H(\Theta, \Theta'). \end{aligned} \tag{S28}$$

Here  $Q(\Theta, \Theta')$  is the objective function in the M-step of the  $r$ th iteration when  $\Theta' = \Theta^{(r)}$ .

Let  $M(\Theta) = \underset{\Theta' \in \Phi}{\operatorname{argmax}} Q(\Theta', \Theta)$  denote a point-to-set map on the parameter space  $\Phi$ , and the parameter sequence  $\{\Theta^{(r)}\}$  can be generated by  $\Theta^{(r+1)} \in M(\Theta^{(r)})$ .

Let  $\Gamma$  denote the set containing all stationary points of  $l_{\text{mp}}$ . In what follows, we prove that all limit points of  $\{\Theta^{(r)}\}$  belong to  $\Gamma$ . We have that

$$\Theta^{(r)} \text{ belongs to a compact set } \Phi_P \subset \Phi \text{ for } r = 1, 2, \dots \tag{S29}$$

because  $l_{\text{mp}}(\Theta)$  tends to minus infinity as  $\|\Theta\| \rightarrow \infty$ . We can prove  $Q(\Theta^{(r+1)}, \Theta^{(r)}) \geq Q(\Theta^{(r)}, \Theta^{(r)})$  and  $H(\Theta^{(r+1)}, \Theta^{(r)}) \leq H(\Theta^{(r)}, \Theta^{(r)})$  using Jensen's inequality. It immediately follows that



$$l_{\text{mp}}(\Theta^{(r+1)}) \geq l_{\text{mp}}(\Theta^{(r)}) \quad (\text{S30})$$

by the definition of  $l_{\text{mp}}(\Theta)$  given in (S28). Furthermore, for  $\Theta^{(r)} \notin \Gamma$ , in what follows we prove

$$l_{\text{mp}}(\Theta^{(r+1)}) > l_{\text{mp}}(\Theta^{(r)}). \quad (\text{S31})$$

In fact, if  $l_{\text{mp}}(\Theta^{(r+1)}) = l_{\text{mp}}(\Theta^{(r)})$ , then  $Q(\Theta^{(r+1)}, \Theta^{(r)}) = Q(\Theta^{(r)}, \Theta^{(r)})$  and  $H(\Theta^{(r+1)}, \Theta^{(r)}) = H(\Theta^{(r)}, \Theta^{(r)})$ . As a result,

$$\frac{\partial l_{\text{mp}}(\Theta^{(r)})}{\partial \Theta} = \left[ \frac{\partial Q(\Theta, \Theta^{(r)})}{\partial \Theta} \right] \Big|_{\Theta=\Theta^{(r)}} - \left[ \frac{\partial H(\Theta, \Theta^{(r)})}{\partial \Theta} \right] \Big|_{\Theta=\Theta^{(r)}} = 0,$$

which establishes a contradiction. The inequality (S31) is thus proved. By (S29), the parameter sequence  $\{\Theta^{(r)}\}$  must have a convergent subsequence, which means that there exists a subset  $\mathbb{N}^1 \subset \mathbb{N}$  such that

$$\Theta^{(r)} \rightarrow \Theta^{(\infty)} \quad \text{for } r \in \mathbb{N}^1, \quad (\text{S32})$$

where  $\Theta^{(\infty)}$  denotes the limit of this subsequence. Consequently, we have

$$\lim_{r \rightarrow +\infty} l_{\text{mp}}(\Theta^{(r)}) = l_{\text{mp}}(\Theta^{(\infty)})$$

in view of the inequality (S30) and the continuity of  $l_{\text{mp}}$ . Now consider  $\{\Theta^{(r+1)}\}_{r \in \mathbb{N}^1}$ . Again, it follows from (S29) that there exists a subset  $\mathbb{N}^2 \subset \mathbb{N}^1$  such that

$$\Theta^{(r+1)} \rightarrow \Theta^{(\infty+1)} \quad \text{for } r \in \mathbb{N}^2, \quad (\text{S33})$$

where  $\Theta^{(\infty+1)}$  denotes the limit of  $\{\Theta^{(r+1)}\}_{r \in \mathbb{N}^2}$ . Similarly, we have

$$l_{\text{mp}}(\Theta^{(\infty+1)}) = \lim_{r \rightarrow +\infty} l_{\text{mp}}(\Theta^{(r+1)}) = \lim_{r \rightarrow +\infty} l_{\text{mp}}(\Theta^{(r)}) = l_{\text{mp}}(\Theta^{(\infty)}). \quad (\text{S34})$$

To prove  $\Theta^{(\infty)} \in \Gamma$ , we first assume  $\Theta^{(\infty)} \notin \Gamma$  then establish a contradiction in what follows. Note that  $Q(\Theta', \Theta)$  is continuous with respect to both  $\Theta'$  and  $\Theta$ , which indicates that  $M$  is a closed operator. According to (S32) and (S33), we have

$$\Theta^{(\infty+1)} \in M(\Theta^{(\infty)}).$$

As a result,  $\Theta^{(\infty)} \notin \Gamma$  implies

$$l_{\text{mp}}(\Theta^{(\infty+1)}) > l_{\text{mp}}(\Theta^{(\infty)})$$

in view of (S31), which is contradicted with (S34). This proves  $\Theta^{(\infty)} \in \Gamma$ .

According to the assumption of Theorem 2,  $l_{\text{mp}}$  takes different values at different points in  $\Gamma$ , which ensures that  $\{\Theta^{(r)}\}$  has a unique limit point due to (S30). In summary,  $\{\Theta^{(r)}\}$  converge to a stationary point of  $l_{\text{mp}}$ . Furthermore,  $\{\Theta^{(r)}\}$  would converge to the maximizer of  $l_{\text{mp}}$  if an appropriate initial value is chosen.

## References

- [1] Dongyu Lin, Clarice R Weinberg, Rui Feng, Hagit Hochner, and Jinbo Chen. A multi-locus likelihood method for assessing parent-of-origin effects using case-control mother-child pairs. *Genetic Epidemiology*, 37(2):152–162, 2013.
- [2] Jin Qin and Jerry Lawless. Empirical likelihood and general estimating equations. *The Annals of Statistics*, 22(1):300–325, 1994.
- [3] CF Jeff Wu. On the convergence properties of the em algorithm. *The Annals of Statistics*, 11(1):95–103, 1983.
- [4] Willard I Zangwill. *Nonlinear programming: a unified approach*, volume 52. Prentice-hall Englewood Cliffs, NJ, 1969.

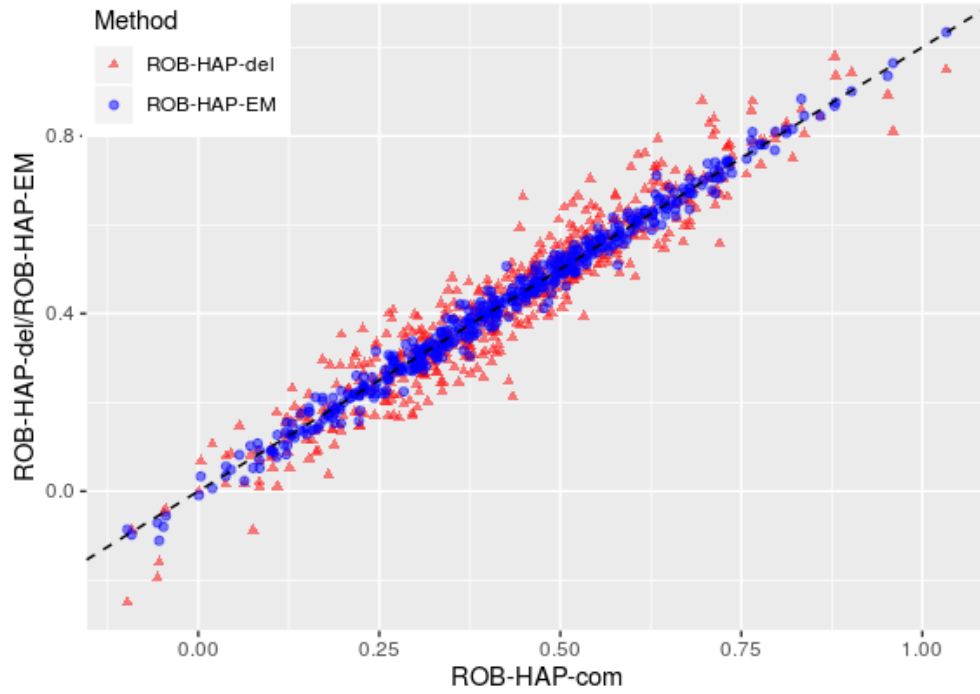


Figure S1: Estimation results for POEs ( $\beta_{im}$ ) by three versions of ROB-HAP mentioned in Section 3.7 in the main text. The true value  $\beta_{im}$  was 0.405. ROB-HAP-com, a version of ROB-HAP utilizing complete genotypes; ROB-HAP-del, a version of ROB-HAP deleting families with incomplete genotypes; ROB-HAP-EM, a version of ROB-HAP handling missing genotypes through EM algorithm.

Table S1: Simulation results for checking regularity condition (C8) in the main text<sup>a</sup>

$\beta_{im}$	$n$	$n^{-1}\partial^2 l_0 / \partial \beta_{gm} \partial \lambda$		$n^{-1}\partial^2 l_0 / \partial \beta_{gc} \partial \lambda$		$n^{-1}\partial^2 l_0 / \partial \beta_{im} \partial \lambda$		$n^{-1}\partial^2 l_0 / \partial \beta_x \partial \lambda$	
		EST <sup>b</sup>	SD <sup>c</sup>	EST <sup>b</sup>	SD <sup>c</sup>	EST <sup>b</sup>	SD <sup>c</sup>	EST <sup>b</sup>	SD <sup>c</sup>
0	200	4.19e-03	3.46e-04	4.56e-03	1.96e-04	2.57e-04	1.97e-04	2.50e-03	5.42e-04
	400	4.18e-03	2.45e-04	4.56e-03	1.39e-04	2.54e-04	1.39e-04	2.49e-03	3.89e-04
	800	4.19e-03	1.74e-04	4.56e-03	9.86e-05	2.55e-04	9.87e-05	2.50e-03	2.74e-04
log(1.5)	200	4.35e-03	3.34e-04	4.32e-03	2.14e-04	1.28e-03	2.11e-04	2.39e-03	5.37e-04
	400	4.35e-03	2.35e-04	4.32e-03	1.51e-04	1.28e-03	1.48e-04	2.38e-03	3.79e-04
	800	4.35e-03	1.66e-04	4.32e-03	1.06e-04	1.28e-03	1.04e-04	2.39e-03	2.73e-04

<sup>a</sup>Simulation setup is described in Appendix E; <sup>b</sup>Mean of sample mean based on 10,000 replicates of simulations; <sup>c</sup>Standard error of sample mean based on 10,000 replicates of simulations.

Table S2: Haplotype configurations and frequencies used in simulation studies<sup>a</sup>

Haplotype	SNP1 <sup>b</sup>	SNP2 <sup>b</sup>	SNP3 <sup>b</sup>	SNP4 <sup>b</sup>	SNP5 <sup>b</sup>	Frequency <sup>c</sup>
$H_1$	0	0	0	0	1	0.298
$H_2$	0	0	0	0	0	0.267
$H_3$	0	0	1	0	1	0.152
$H_4$	1	1	0	1	0	0.117
$H_5$	0	0	1	0	0	0.099
$H_6$	1	0	0	1	0	0.034
$H_7$	1	1	1	0	0	0.032
MAF <sup>d</sup>	0.183	0.149	0.283	0.151	0.450	

<sup>a</sup>SNP1-SNP5 are located in the gene GPX1 [1]. <sup>b</sup>0 and 1 refer to the common and rare allele, respectively; <sup>c</sup>Haplotype frequency; <sup>d</sup>Minor allele frequency.

Table S3: Sensitivity analysis of ROB-HAP w.r.t. Hardy-Weinberg equilibrium<sup>a</sup>

Fixation index $F$	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
Type-I error rate	0.05	0.052	0.042	0.046	0.064	0.056	0.038	0.03	0.056	0.056

<sup>a</sup>The target SNP was SNP1 in gene GPX1 shown in Table S2.