Supplementary material to "Asymptotic theory of dependent Bayesian multiple testing procedures under possible model misspecification"

Noirrit Kiran Chandra* Sourabh Bhattacharya†

S-1 Assumptions of Shalizi (2009)

(S1) Consider the following likelihood ratio:

$$R_n(\boldsymbol{\xi}) = \frac{f_{\boldsymbol{\xi}}(\boldsymbol{X}_n)}{p(\boldsymbol{X}_n)}. \tag{S-1}$$

Assume that $R_n(\boldsymbol{\xi})$ is $\sigma(\boldsymbol{X}_n) \times \mathcal{T}$-measurable for all $n > 0$.

(S2) For each $\boldsymbol{\xi} \in \Xi$, the generalized or relative asymptotic equipartition property holds, and so, almost surely,

$$\lim_{n\to\infty} \frac{1}{n} \log R_n(\boldsymbol{\xi}) = -h(\boldsymbol{\xi}),$$

where $h(\boldsymbol{\xi})$ is given in (S3) below.

(S3) For every $\boldsymbol{\xi} \in \Xi$, the KL-divergence rate

$$h(\boldsymbol{\xi}) = \lim_{n\to\infty} \frac{1}{n} E\left(\log \frac{p(\boldsymbol{X}_n)}{f_{\boldsymbol{\xi}}(\boldsymbol{X}_n)}\right). \tag{S-2}$$

exists (possibly being infinite) and is $\mathcal{T}$-measurable.

(S4) Let $I = \{\boldsymbol{\xi} : h(\boldsymbol{\xi}) = \infty\}$. The prior π satisfies $\pi(I) < 1$.

(S5) There exists a sequence of sets $\mathcal{G}_n \to \Xi$ as $n \to \infty$ such that:

(1)
$$\pi(\mathcal{G}_n) \geq 1 - \alpha \exp(-\varsigma n), \text{ for some } \alpha > 0,\ \varsigma > 2h(\Xi); \tag{S-3}$$

(2) The convergence in (S3) is uniform in θ over $\mathcal{G}_n \setminus I$.

(3) $h(\mathcal{G}_n) \to h(\Xi)$, as $n \to \infty$.

For each measurable $A \subseteq \Xi$, for every $\delta > 0$, there exists a random natural number $\tau(A, \delta)$ such that

$$n^{-1} \log \int_A R_n(\boldsymbol{\xi})\pi(\boldsymbol{\xi})d\boldsymbol{\xi} \leq \delta + \limsup_{n\to\infty} n^{-1} \log \int_A R_n(\boldsymbol{\xi})\pi(\boldsymbol{\xi})d\boldsymbol{\xi}, \tag{S-4}$$

for all $n > \tau(A, \delta)$, provided $\limsup_{n\to\infty} n^{-1} \log \pi(\mathbb{I}_A R_n) < \infty$. Regarding this, the following assumption has been made by Shalizi:

*Department of Statistics and Data Science, University of Texas at Austin, 2317 Speedway D9800, Austin, TX 78712-1823, USA

†Interdisciplinary Statistical Research Unit, Indian Statistical Institute, 203 B. T. Road, Kolkata, WB, 700108, India

(S6) The sets $\mathcal{G}_n$ of (S5) can be chosen such that for every $\delta > 0$, the inequality $n > \tau(\mathcal{G}_n, \delta)$ holds almost surely for all sufficiently large n.

(S7) The sets $\mathcal{G}_n$ of (S5) and (S6) can be chosen such that for any set A with $\pi(A) > 0$,

$$h\left(\mathcal{G}_n \cap A\right) \to h\left(A\right) \text{ as } n \to \infty \tag{S-5}$$

S-2 Comparisons of versions of FNR

With respect to the new notions of errors in (2) and (3), $\text{FNR}_{\boldsymbol{X}_n}$ can be modified as

$$\begin{aligned} modified\ \text{FNR}_{\boldsymbol{X}_n} &= E_{\boldsymbol{\xi}|\boldsymbol{X}_n}\left[\sum_{\boldsymbol{d}\in\mathbb{D}} \frac{\sum_{i=1}^m (1-d_i) r_i z_i}{\sum_{i=1}^m (1-d_i) \vee 1} \delta_{\mathcal{M}}\left(\boldsymbol{d}|\boldsymbol{X}_n\right)\right] \\ &= \sum_{\boldsymbol{d}\in\mathbb{D}} \frac{\sum_{i=1}^m (1-d_i) w_{in}(\boldsymbol{d})}{\sum_{i=1}^m (1-d_i) \vee 1} \delta_{\mathcal{M}}(\boldsymbol{d}|\boldsymbol{X}_n). \end{aligned}$$

We denote $modified\ \text{FNR}_{\boldsymbol{X}_n}$ as $\text{mFNR}_{\boldsymbol{X}_n}$. Now, from Theorem 1, $d_i^t = 0$ implies

$$\exp\left[-n\left(J\left(\Xi_{\boldsymbol{d}^t,i}\right) + \epsilon\right)\right] < w_{in}(\boldsymbol{d}^t) < \exp\left[-n\left(J\left(\Xi_{\boldsymbol{d}^t,i}\right) - \epsilon\right)\right].$$

Similar to Theorem 4, using the above bounds, we can obtain the asymptotic convergence rate of $\text{mFNR}_{\boldsymbol{X}_n}$, formalized in the following theorem:

Theorem S-2.1 *Assume conditions (A1) and (A2). Let* $\tilde{J}_{\min} = \min_{i:d_i^t=0} J(\Xi_{\boldsymbol{d}^t,i})$. *Then for the non-marginal multiple testing procedure*

$$\lim_{n\to\infty} \frac{1}{n} \log \text{mFNR}_{\boldsymbol{X}_n} = -\tilde{J}_{\min}. \tag{S-6}$$

Proof. Following the proof of Lemma S-4.1, we have

$$\exp(-n\epsilon) \times \frac{\sum_{i=1}^m (1-d_i^t) e^{-nJ(\Xi_{i\boldsymbol{d}^t})}}{\sum_{i=1}^m (1-d_i^t)} \leq \text{mFNR}_{\boldsymbol{X}_n} \leq \exp(n\epsilon) \times \frac{\sum_{i=1}^m (1-d_i^t) e^{-nJ(\Xi_{i\boldsymbol{d}^t})}}{\sum_{i=1}^m (1-d_i^t)},$$

from which the proof follows. ■

If $J\left(\Xi_{\boldsymbol{d}^t,i}\right) = J\left(H_{1i}\right)$ for $i = 1, \ldots, m$, it would follow that $w_{in}(\boldsymbol{d}^t)$ and v_{in} have the same lower and upper bounds. Lemma S-2.2 shows that indeed $J\left(\Xi_{\boldsymbol{d}^t,i}\right) = J\left(H_{1i}\right)$ for $i = 1, \ldots, m$, under a very mild assumption given by the following.

(A3) For any decision configuration $\boldsymbol{d}$, define $S(\boldsymbol{d}) = \{i : d_i = d_i^t\}$. Then for two decision configurations $\boldsymbol{d}$ and $\tilde{\boldsymbol{d}}$, if $S(\boldsymbol{d}) \subset S(\tilde{\boldsymbol{d}})$, then $J(\Xi(\boldsymbol{d})) > J(\Xi(\tilde{\boldsymbol{d}}))$.

Notably in (A3), $S(\boldsymbol{d})$ is the set of correct decisions. Note that $S(\boldsymbol{d}) \subset S(\tilde{\boldsymbol{d}})$ implies that number of correct decisions is more in $\tilde{\boldsymbol{d}}$ compared to $\boldsymbol{d}$. Hence, the model directed by $\boldsymbol{d}$ should procure greater divergence. This assumption is easily seen to hold in independent cases, and also in dependent models such as multivariate normal.

Lemma S-2.2 *Under (A3),* $J\left(\Xi_{\boldsymbol{d}^t,i}\right) = J\left(H_{1i}\right)$, *for all i such that* $d_i^t = 0$.

Proof. For all i such that $d_i^t = 0$, define $\boldsymbol{d}^{(i)}$, where $d_j^{(i)} = d_j^t$ for all $j \neq i$, and $d_j^{(i)} = 1$ and $S_i = \{\boldsymbol{d} : d_i = 1\}$. Then

$$P_{\boldsymbol{\xi}|\boldsymbol{X}_n}\left(H_{1i}\right) = \sum_{\boldsymbol{d}\in S_i} P_{\boldsymbol{\xi}|\boldsymbol{X}_n}\left(H_{1i} \cap \left\{\cap_{j\neq i} H_{d_j,j}\right\}\right), \tag{S-7}$$

so that dividing both sides of (S-7) by $P_{\boldsymbol{\xi}|\boldsymbol{X}_n}\left(\Xi_{\boldsymbol{d}^t,i}\right)$ yields

$$\frac{P_{\boldsymbol{\xi}|\boldsymbol{X}_n}\left(H_{1i}\right)}{P_{\boldsymbol{\xi}|\boldsymbol{X}_n}\left(\Xi_{\boldsymbol{d}^t,i}\right)} = 1 + \sum_{\boldsymbol{d}\in S_i\setminus\{\boldsymbol{d}^{(i)}\}} \frac{P_{\boldsymbol{\xi}|\boldsymbol{X}_n}\left(H_{1i}\cap\left\{\cap_{j\neq i}H_{d_j,j}\right\}\right)}{P_{\boldsymbol{\xi}|\boldsymbol{X}_n}\left(\Xi_{\boldsymbol{d}^t,i}\right)} \tag{S-8}$$

Theorem 1 and (A3) together ensures that as $n\to\infty$, $\frac{P_{\boldsymbol{\xi}|\boldsymbol{X}_n}\left(H_{1i}\cap\left\{\cap_{j\neq i}H_{d_j,j}\right\}\right)}{P_{\boldsymbol{\xi}|\boldsymbol{X}_n}\left(\Xi_{\boldsymbol{d}^t,i}\right)}\to 0$ exponentially fast, for all $\boldsymbol{d}\in S_i\setminus\{\boldsymbol{d}^{(i)}\}$. Applying this to the right hand side of (S-8) yields

$$\frac{P_{\boldsymbol{\xi}|\boldsymbol{X}_n}\left(H_{1i}\right)}{P_{\boldsymbol{\xi}|\boldsymbol{X}_n}\left(\Xi_{\boldsymbol{d}^t,i}\right)} \to 1 \tag{S-9}$$

exponentially fast. Now, applying Shalizi's result to $P_{\boldsymbol{\xi}|\boldsymbol{X}_n}\left(H_{1i}\right)$ and $P_{\boldsymbol{\xi}|\boldsymbol{X}_n}\left(\Xi_{\boldsymbol{d}^t,i}\right)$ it follows that if $J\left(\Xi_{\boldsymbol{d}^t,i}\right)\neq J\left(H_{1i}\right)$, then (S-9) is contradicted. Hence, $J\left(\Xi_{\boldsymbol{d}^t,i}\right)=J\left(H_{1i}\right)$, for $i=1,\ldots,m$. ■

From Lemma S-2.2, we see that $\tilde{J}_{\min}=\tilde{H}_{\min}$. Thus, we get the following result:

Theorem S-2.3 *Assume (A1)–(A3). Then, for the non-marginal multiple testing procedure,*

$$\lim_{n\to\infty}\frac{1}{n}\log\left(\frac{\text{mFNR}_{\boldsymbol{X}_n}}{\text{FNR}_{\boldsymbol{X}_n}}\right)=0 \tag{S-10}$$

and

$$\lim_{n\to\infty}\frac{\log\left(\text{mFNR}_{\boldsymbol{X}_n}\right)}{\log\left(\text{FNR}_{\boldsymbol{X}_n}\right)}=1. \tag{S-11}$$

Proof. Note that,

$$\frac{1}{n}\log\left(\frac{\text{mFNR}_{\boldsymbol{X}_n}}{\text{FNR}_{\boldsymbol{X}_n}}\right)=\frac{1}{n}\log\left(\text{mFNR}_{\boldsymbol{X}_n}\right)-\frac{1}{n}\log\left(\text{FNR}_{\boldsymbol{X}_n}\right).$$

Now $\frac{1}{n}\log\left(\text{mFNR}_{\boldsymbol{X}_n}\right)\to-\tilde{J}_{\min}$ and $\frac{1}{n}\log\left(\text{FNR}_{\boldsymbol{X}_n}\right)\to-\tilde{H}_{\min}$ as $n\to\infty$. Again by Lemma S-2.2, $\tilde{J}_{\min}=\tilde{H}_{\min}$. This proves (S-10). The proof of (S-11) follows from (S-6) and (15), using $\tilde{J}_{\min}=\tilde{H}_{\min}$. ■

Theorem S-2.3 remains true for any $\boldsymbol{G}=\{G_1,\ldots,G_m\}$. In other words, given that (A3) holds, (S-10) shows that none of $\text{mFNR}_{\boldsymbol{X}_n}$ or $\text{FNR}_{\boldsymbol{X}_n}$ is asymptotically preferable over the other, while (S-11) shows that $\log\left(\text{mFNR}_{\boldsymbol{X}_n}\right)$ and $\log\left(\text{FNR}_{\boldsymbol{X}_n}\right)$ are asymptotically equivalent, irrespective of how the G_i's are formed.

S-3 Proofs of results in Section 2

Proof of Theorem 2. Let Ξ^{tc} be the complement set of $\Xi(\boldsymbol{d}^t)$. Then by virtue of Theorem 1 we have

$$\lim_{n\to\infty}\frac{1}{n}\log P_{\boldsymbol{\xi}|\boldsymbol{X}_n}\left(\Xi^{tc}\right)=-J\left(\Xi^{tc}\right).$$

This implies that for any $\epsilon>0$, there exists a $n_0(\epsilon)$ such that for all $n>n_0(\epsilon)$

$$\begin{aligned}&\exp\left[-n\left(J\left(\Xi^{tc}\right)+\epsilon\right)\right]<P_{\boldsymbol{\xi}|\boldsymbol{X}_n}\left(\Xi^{tc}\right)<\exp\left[-n\left(J\left(\Xi^{tc}\right)-\epsilon\right)\right]\\ \Rightarrow&1-\exp\left[-n\left(J\left(\Xi^{tc}\right)-\epsilon\right)\right]<P_{\boldsymbol{\xi}|\boldsymbol{X}_n}\left(\Xi^{t}\right)<1-\exp\left[-n\left(J\left(\Xi^{tc}\right)+\epsilon\right)\right].\end{aligned}$$

For notational convenience, we shall henceforth denote $J\left(\Xi^{tc}\right)$ by J.

Observe that if $\boldsymbol{d} \in \mathbb{D}_i^c$, at least one decision is wrong corresponding to some hypothesis in G_i. As $P_{\boldsymbol{\xi}|\boldsymbol{X}_n}\left(\boldsymbol{\Xi}^{tc}\right)$ is the posterior probability of at least one wrong decision in the parameter space, we have

$$w_{in}(\boldsymbol{d}) < P_{\boldsymbol{\xi}|\boldsymbol{X}_n}\left(\boldsymbol{\Xi}^{tc}\right) < \exp\left[-n\left(J-\epsilon\right)\right]. \tag{S-12}$$

Similarly for $\boldsymbol{d} \in \mathbb{D}_i$ and for false H_{0i}

$$w_{in}(\boldsymbol{d}) > P_{\boldsymbol{\xi}|\boldsymbol{X}_n}\left(\boldsymbol{\Xi}^{t}\right) > 1 - \exp\left[-n\left(J-\epsilon\right)\right]. \tag{S-13}$$

From conditions (12) and (13), it follows that there exists n_1 such that for all $n > n_1$

$$\beta_n > \underline{\beta} - \delta,$$
$$\beta_n < 1 - \delta, \text{ such that}$$

$\underline{\beta} - \delta > 0$ and $1 - \bar{\beta} > \delta$, for some $\delta > 0$. It follows using this, (S-12) and (S-13), that

$$\sum_{i:\boldsymbol{d}\in\mathbb{D}_i^c}^m d_i^t w_{in}(\boldsymbol{d}^t) - \sum_{i:\boldsymbol{d}\in\mathbb{D}_i^c}^m d_i w_{in}(\boldsymbol{d}) > \left(1 - e^{-n(J-\epsilon)}\right) \sum_{i:\boldsymbol{d}\in\mathbb{D}_i^c} d_i^t - e^{-n(J-\epsilon)} \sum_{i:\boldsymbol{d}\in\mathbb{D}_i^c} d_i, \text{ and}$$
$$\beta_n \left(\sum_{i:\boldsymbol{d}\in\mathbb{D}_i^c}^m d_i^t - \sum_{i:\boldsymbol{d}\in\mathbb{D}_i^c}^m d_i \right) < (1-\delta) \sum_{i:\boldsymbol{d}\in\mathbb{D}_i^c}^m d_i^t - (\underline{\beta} - \delta) \sum_{i:\boldsymbol{d}\in\mathbb{D}_i^c}^m d_i.$$

Now n_1 can be appropriately chosen such that $e^{-n(J-\epsilon)} < \min\{\delta, \underline{\beta} - \delta\}$. Note that neither n_0 nor n_1 depends on m. Hence, for any value of m and for all $n > \max\{n_0, n_1\}$,

$$\sum_{i:\boldsymbol{d}\in\mathbb{D}_i^c}^m d_i^t w_{in}(\boldsymbol{d}^t) - \sum_{i:\boldsymbol{d}\in\mathbb{D}_i^c}^m d_i w_{in}(\boldsymbol{d}) > \beta_n \left(\sum_{i:\boldsymbol{d}\in\mathbb{D}_i^c}^m d_i^t - \sum_{i:\boldsymbol{d}\in\mathbb{D}_i^c}^m d_i \right), \text{ for all } \boldsymbol{d} \neq \boldsymbol{d}^t, \text{ almost surely}$$
$$\Rightarrow \lim_{n\to\infty} \delta_{\mathcal{NM}}(\boldsymbol{d}^t|\boldsymbol{X}_n) = 1, \text{ almost surely.}$$

■

S-4 Additional results to Section 3 and proofs

Lemma S-4.1 *Assume conditions (A1) and (A2). Then for the non-marginal multiple testing procedure and any $\epsilon > 0$, there exists $n_0(\epsilon) \geq 1$ such that for $n \geq n_0(\epsilon)$, the following holds almost surely:*

$$\exp\left(-n\epsilon\right) \times \frac{\sum_{i=1}^m d_i^t e^{-nJ(\boldsymbol{\Xi}^c_{\boldsymbol{d}^t,i})}}{\sum_{i=1}^m d_i^t} \leq \text{mFDR}_{\boldsymbol{X}_n} \leq \exp\left(n\epsilon\right) \times \frac{\sum_{i=1}^m d_i^t e^{-nJ(\boldsymbol{\Xi}^c_{\boldsymbol{d}^t,i})}}{\sum_{i=1}^m d_i^t}.$$

Proof. Observe that,

$$\begin{aligned}
&\text{mFDR}_{\boldsymbol{X}_n} \\
&= \sum_{\boldsymbol{d}\neq\boldsymbol{0}} \frac{\sum_{i=1}^m d_i(1 - w_{in}(\boldsymbol{d}))}{\sum_{i=1}^m d_i \vee 1} \delta_{\mathcal{NM}}\left(\boldsymbol{d}|\boldsymbol{X}_n\right) \\
&= \frac{\sum_{i=1}^m d_i^t(1 - w_{in}(\boldsymbol{d}^t))}{\sum_{i=1}^m d_i^t} \delta_{\mathcal{NM}}\left(\boldsymbol{d}^t|\boldsymbol{X}_n\right) + \sum_{\boldsymbol{d}\neq\boldsymbol{d}^t} \frac{\sum_{i=1}^m d_i(1 - w_{in}(\boldsymbol{d}))}{\sum_{i=1}^m d_i \vee 1} \delta_{\mathcal{NM}}\left(\boldsymbol{d}|\boldsymbol{X}_n\right).
\end{aligned} \tag{S-14}$$

From the proof of Theorem 2, we see that under (A1), $\delta_{\mathcal{NM}}(\boldsymbol{d}|\boldsymbol{X}_n) = 0$ for all $\boldsymbol{d} \neq \boldsymbol{d}^t$. Also under

(A2), $\boldsymbol{d}^t \neq \mathbf{0}$. For any $\epsilon > 0$ and $n \geq n_0(\epsilon)$, it follows from (8) and (9) that a lower bound for (S-14) is

$$L_n = \frac{\sum_{i=1}^m d_i^t e^{-n(J(\Xi^c_{\boldsymbol{d}^t,i})+\epsilon)}}{\sum_{i=1}^m d_i^t} \delta_{\mathcal{NM}}(\boldsymbol{d}^t|\boldsymbol{X}_n) = \exp\left(-n\epsilon\right) \times \frac{\sum_{i=1}^m d_i^t e^{-nJ(\Xi^c_{\boldsymbol{d}^t,i})}}{\sum_{i=1}^m d_i^t}.$$

Similarly, an upper bound is given by

$$U_n = \frac{\sum_{i=1}^m d_i^t e^{-n(J(\Xi^c_{\boldsymbol{d}^t,i})-\epsilon)}}{\sum_{i=1}^m d_i^t} \delta_{\mathcal{NM}}\left(\boldsymbol{d}^t|\boldsymbol{X}_n\right) = \exp\left(n\epsilon\right) \times \frac{\sum_{i=1}^m d_i^t e^{-nJ(\Xi^c_{\boldsymbol{d}^t,i})}}{\sum_{i=1}^m d_i^t}.$$

■

Similar asymptotic bounds can also be obtained for $\mathrm{FDR}_{\boldsymbol{X}_n}$ under the same conditions. We state it formally in the following corollary.

Corollary S-4.2 *Assume conditions (A1) and (A2). Then for the non-marginal multiple testing procedure and any $\epsilon > 0$ and large enough n the following holds almost surely:*

$$\exp(-n\epsilon) \times \frac{\sum_{i=1}^m d_i^t e^{-nJ(H_{0i})}}{\sum_{i=1}^m d_i^t} \leq \mathrm{FDR}_{\boldsymbol{X}_n} \leq \exp(n\epsilon) \times \frac{\sum_{i=1}^m d_i^t e^{-nJ(H_{0i})}}{\sum_{i=1}^m d_i^t}.$$

Lemma S-4.3 *Assume conditions (A1) and (A2). Then for the non-marginal multiple testing procedure and any $\epsilon > 0$, there exists a natural number $n_1(\epsilon)$ such that for all $n > n_1(\epsilon)$ the following hold almost surely*

$$\exp(-n\epsilon) \times \frac{\sum_{i=1}^m (1-d_i^t) e^{-nJ(\Xi_{\boldsymbol{d}^t,i})}}{\sum_{i=1}^m (1-d_i^t)} \leq \mathrm{mFNR}_{\boldsymbol{X}_n} \leq \exp(n\epsilon) \times \frac{\sum_{i=1}^m (1-d_i^t) e^{-nJ(\Xi_{\boldsymbol{d}^t,i})}}{\sum_{i=1}^m (1-d_i^t)}.$$

Proof. Note that by Theorem 1, $d_i^t = 0$ implies

$$\exp\left[-n\left(J\left(H_{1i}\right)+\epsilon\right)\right] < v_{in} < \exp\left[-n\left(J\left(H_{1i}\right)-\epsilon\right)\right].$$

From the above bound, similar to the proof of Lemma S-4.1, we obtain asymptotic bounds of $\mathrm{FNR}_{\boldsymbol{X}_n}$.
■

Note that, (A2) is required for both Lemma S-4.1 and S-4.3 to hold. Without the condition the denominators of the bounds would become zero. For proper bounds of the errors and hence for the limits, (A2) is necessary.

Proof of Theorem 3. From Lemma S-4.1 we obtain the following for $n \geq n_0(\epsilon)$,

$$\begin{aligned}
&\exp\left(-n\epsilon\right) \times \frac{\sum_{i=1}^m d_i^t e^{-nJ(\Xi^c_{\boldsymbol{d}^t,i})}}{\sum_{i=1}^m d_i^t} \leq \mathrm{mFDR}_{\boldsymbol{X}_n} \leq \exp\left(n\epsilon\right) \times \frac{\sum_{i=1}^m d_i^t e^{-nJ(\Xi^c_{\boldsymbol{d}^t,i})}}{\sum_{i=1}^m d_i^t} \\
\Longleftrightarrow &-\epsilon + \frac{1}{n}\log\left(\sum_{i=1}^m d_i^t e^{-nJ(\Xi^c_{\boldsymbol{d}^t,i})}\right) - \frac{1}{n}\log\left(\sum_{i=1}^m d_i^t\right) \leq \frac{1}{n}\log \mathrm{mFDR}_{\boldsymbol{X}_n} \\
&\leq \epsilon + \frac{1}{n}\log\left(\sum_{i=1}^m d_i^t e^{-nJ(\Xi^c_{\boldsymbol{d}^t,i})}\right) - \frac{1}{n}\log\left(\sum_{i=1}^m d_i^t\right).
\end{aligned}$$

Applying *L'Hôpital's rule* we observe that

$$\lim_{n\to\infty} \frac{1}{n}\log\left(\sum_{i=1}^m d_i^t e^{-nJ(\Xi^c_{\boldsymbol{d}^t,i})}\right) = -J_{\min}.$$

As ϵ is an arbitrarily small positive quantity, we have

$$\lim_{n\to\infty}\frac{1}{n}\log \mathrm{mFDR}_{\boldsymbol{X}_n} = -J_{\min}.$$

Proceeding in the exact same way, using Corollary S-4.2, we obtain

$$\lim_{n\to\infty}\frac{1}{n}\log \mathrm{FDR}_{\boldsymbol{X}_n} = -H_{\min}.$$

■

Proof of Corollary 2. Note that

$$\begin{aligned}
&\mathrm{mpBFDR}\\
=&E_{\boldsymbol{X}_n}\left[\sum_{\boldsymbol{d}\in\mathbb{D}}\frac{\sum_{i=1}^m d_i(1-w_i(\boldsymbol{d}))}{\sum_{i=1}^m d_i}\delta_\beta(\boldsymbol{d}|\boldsymbol{X}_n)\middle|\delta_{\mathcal{NM}}(\boldsymbol{d}=\boldsymbol{0}|\boldsymbol{X}_n)=0\right]\\
=&E_{\boldsymbol{X}_n}\left[\sum_{\boldsymbol{d}\in\mathbb{D}}\frac{\sum_{i=1}^m d_i(1-w_i(\boldsymbol{d}))}{\sum_{i=1}^m d_i}\delta_{\mathcal{NM}}(\boldsymbol{d}|\boldsymbol{X}_n)\middle|\delta_{\mathcal{NM}}(\boldsymbol{d}=\boldsymbol{0}|\boldsymbol{X}_n)=0\right]\\
=&E_{\boldsymbol{X}_n}\left[\sum_{\boldsymbol{d}\in\mathbb{D}}\frac{\sum_{i=1}^m d_i(1-w_i(\boldsymbol{d}))}{\sum_{i=1}^m d_i}I\left(\sum_{i=1}^m d_i>0\right)\delta_{\mathcal{NM}}(\boldsymbol{d}|\boldsymbol{X}_n)\right]\frac{1}{P_{\boldsymbol{X}_n}\left[\delta_{\mathcal{NM}}(\boldsymbol{d}=\boldsymbol{0}|\boldsymbol{X}_n)=0\right]}\\
=&E_{\boldsymbol{X}_n}\left[\sum_{\boldsymbol{d}\in\mathbb{D}\setminus\{\boldsymbol{0}\}}\frac{\sum_{i=1}^m d_i(1-w_i(\boldsymbol{d}))}{\sum_{i=1}^m d_i}\delta_{\mathcal{NM}}(\boldsymbol{d}|\boldsymbol{X}_n)\right]\frac{1}{P_{\boldsymbol{X}_n}\left[\delta_{\mathcal{NM}}(\boldsymbol{d}=\boldsymbol{0}|\boldsymbol{X}_n)=0\right]}.
\end{aligned}$$

From Theorem 3, we have $\frac{1}{n}\log \mathrm{mFDR}_{\boldsymbol{X}_n}\to -J_{\min}$, that is, $\mathrm{mFDR}_{\boldsymbol{X}_n}\to 0$, as $n\to\infty$. Also we have

$$0\le \sum_{\boldsymbol{d}\in\mathbb{D}\setminus\{\boldsymbol{0}\}}\frac{\sum_{i=1}^m d_i(1-w_i(\boldsymbol{d}))}{\sum_{i=1}^m d_i}\delta_{\mathcal{NM}}(\boldsymbol{d}|\boldsymbol{X}_n)\le \mathrm{mFDR}_{\boldsymbol{X}_n}\le 1.$$

Therefore by the dominated convergence theorem, $E_{\boldsymbol{X}_n}\left[\sum_{\boldsymbol{d}\in\mathbb{D}\setminus\{\boldsymbol{0}\}}\frac{\sum_{i=1}^m d_i(1-w_i(\boldsymbol{d}))}{\sum_{i=1}^m d_i}\delta_{\mathcal{NM}}(\boldsymbol{d}|\boldsymbol{X}_n)\right]\to 0$, as $n\to\infty$. From (A2) we have $\boldsymbol{d}^t\neq\boldsymbol{0}$ and from Theorem 2 we have $E_{\boldsymbol{X}_n}[\delta_{\mathcal{NM}}(\boldsymbol{d}^t|\boldsymbol{X}_n)]\to 1$. Thus $P_{\boldsymbol{X}_n}\left[\delta_{\mathcal{NM}}(\boldsymbol{d}=\boldsymbol{0}|\boldsymbol{X}_n)=0\right]\to 1$, as $n\to\infty$. This proves the result.

Similarly it can be shown that $\mathrm{pBFDR}\to 0$ as $n\to\infty$. ■

Proof of Theorem 4. The proof is similar to that of Theorem 3. ■

Proof of Corollary 3. Exploiting Theorem 4 and (A2), the theorem can be proved similarly as the proof of Corollary 2. ■

S-5 Proofs of results in Section 4

Proof of Theorem 5. Theorem 3.4 of Chandra and Bhattacharya (2019) shows that mpBFDR is non-increasing in β. Hence, the maximum error that can be incurred is at $\beta=0$ where we actually maximize $\sum_{i=1}^m d_i w_{in}(\boldsymbol{d})$. Let

$$\hat{\boldsymbol{d}} = \underset{\boldsymbol{d}\in\mathbb{D}}{\operatorname{argmax}}\sum_{i=1}^m d_i w_{in}(\boldsymbol{d}) = \underset{\boldsymbol{d}\in\mathbb{D}}{\operatorname{argmax}}\left[\sum_{i=1}^{m_1} d_i w_{in}(\boldsymbol{d}) + \sum_{i=m_1+1}^{m} d_i w_{in}(\boldsymbol{d})\right]$$

Since the groups in $\{G_1, G_2, \cdots, G_{m_1}\}$ have no overlap with those in $\{G_{m_1+1}, \cdots, G_m\}$, $\sum_{i=1}^{m_1} d_i w_{in}(\boldsymbol{d})$ and $\sum_{i=m_1+1}^{m} d_i w_{in}(\boldsymbol{d})$ can be maximized separately.

Let us define the following notations:

$$Q_{\boldsymbol{d}}^{m_1} = Q_{\boldsymbol{d}} \cap \{1, 2, \cdots, m_1\},\ Q_{\boldsymbol{d}}^{m_1 c} = \{1, 2, \cdots, m_1\} \setminus Q_{\boldsymbol{d}}^{m_1}.$$

Now,

$$\begin{aligned}
&\sum_{i=1}^{m_1} d_i w_{in}(\boldsymbol{d}) - \sum_{i=1}^{m_1} d_i^t w_{in}(\boldsymbol{d}^t) \\
&= \left[\sum_{i\in Q_{\boldsymbol{d}}^{m_1}} d_i w_{in}(\boldsymbol{d}) - \sum_{i\in Q_{\boldsymbol{d}}^{m_1}} d_i^t w_{in}(\boldsymbol{d}^t)\right] + \left[\sum_{i\in Q_{\boldsymbol{d}}^{m_1 c}} d_i w_{in}(\boldsymbol{d}) - \sum_{i\in Q_{\boldsymbol{d}}^{m_1 c}} d_i^t w_{in}(\boldsymbol{d}^t)\right] \\
&= \sum_{i\in Q_{\boldsymbol{d}}^{m_1 c}} d_i w_{in}(\boldsymbol{d}) - \sum_{i\in Q_{\boldsymbol{d}}^{m_1 c}} d_i^t w_{in}(\boldsymbol{d}^t),
\end{aligned}$$

since for any $\boldsymbol{d}$, $\sum_{i\in Q_{\boldsymbol{d}}^{m_1}} d_i w_{in}(\boldsymbol{d}) = \sum_{i\in Q_{\boldsymbol{d}}^{m_1}} d_i^t w_{in}(\boldsymbol{d}^t)$ by definition of $Q_{\boldsymbol{d}}^{m_1}$.

Note that $\sum_{i\in Q_{\boldsymbol{d}}^{m_1 c}} d_i^t w_{in}(\boldsymbol{d}^t)$ can not be zero as it contradicts (B1) that "$G_1, G_2, \cdots, G_{m_1}$ have at least one false null hypothesis." From (8) and (9), we have

$$\sum_{i\in Q_{\boldsymbol{d}}^{m_1 c}} d_i w_{in}(\boldsymbol{d}) \to 0 \text{ for all } \boldsymbol{d} \neq \boldsymbol{d}^t, \text{ and } \sum_{i\in Q_{\boldsymbol{d}}^{m_1 c}} d_i^t w_{in}(\boldsymbol{d}^t) \to \sum_{i\in Q_{\boldsymbol{d}}^{m_1 c}} d_i^t > 0.$$

Hence, for large enough n, for $\boldsymbol{d} \neq \boldsymbol{d}^t$,

$$\sum_{i=1}^{m} d_i w_{in}(\boldsymbol{d}) - \sum_{i=1}^{m} d_i^t w_{in}(\boldsymbol{d}^t) < 0.$$

In other words, $\boldsymbol{d}^t$ (or $\boldsymbol{d}$ such that $d_i = d_i^t$ for all $i = 1, \cdots, m_1$) maximizes $\sum_{i=1}^{m_1} d_i w_{in}(\boldsymbol{d})$ when n is large enough.

Let us now consider the term $\sum_{i=m_1+1}^{m} d_i w_{in}(\boldsymbol{d})$. Note that $\sum_{i=m_1+1}^{m} d_i^t w_{in}(\boldsymbol{d}^t) = 0$ by (B1). For any finite n, $\sum_{i=m_1+1}^{m} d_i w_{in}(\boldsymbol{d})$ is maximized for some decision configuration $\tilde{\boldsymbol{d}}$ where $\tilde{d}_i = 1$ for at least one $i \in \{m_1 + 1, \cdots, m\}$. In that case, $\hat{\boldsymbol{d}}^t = (d_1^t, \ldots, d_{m_1}^t, \tilde{d}_{m_1+1}, \tilde{d}_{m_1+2}, \ldots, \tilde{d}_m)$, so that

$$\lim_{n\to\infty} \frac{\sum_{i=1}^{m} \hat{d}_i(1 - w_{in}(\hat{\boldsymbol{d}}))}{\sum_{i=1}^{m} \hat{d}_i} \geq \frac{1}{\sum_{i=1}^{m} d_i^t + 1},$$

almost surely, for all data sequences. Boundedness of $\frac{\sum_{i=1}^{m} d_i(1-w_{in}(\boldsymbol{d}))}{\sum_{i=1}^{m} d_i}$ for all $\boldsymbol{d}$ and $\boldsymbol{X}_n$ ensures uniform integrability, which, in conjunction with the simple observation that for $\beta = 0$, $P\left(\delta_{\mathcal{NM}}(\boldsymbol{d} = \boldsymbol{0}|\boldsymbol{X}_n) = 0\right) = 1$ for all $n \geq 1$, guarantees that under (B1) it is possible to incur $\text{mpBFDR} \geq \frac{1}{\sum_{i=1}^{m} d_i^t + 1}$ asymptotically.

Now, if $G_{m_1+1}, \cdots, G_m$'s are all disjoint, each consisting of only one true null hypothesis, then $\sum_{i=m_1+1}^{m} d_i w_{in}(\boldsymbol{d})$ will be maximized by $\tilde{\boldsymbol{d}}$ where $\tilde{d}_i = 1$ for all $i \in \{m_1 + 1, \cdots, m\}$. Since d_i^t; $i = 1, \ldots, m_1$ maximizes $\sum_{i=1}^{m_1} d_i w_{in}(\boldsymbol{d})$ for large n, it follows that $\hat{\boldsymbol{d}} = (d_1^t, \ldots, d_{m_1}^t, 1, 1, \ldots, 1)$ is the maximizer of $\sum_{i=1}^{m} d_i w_{in}(\boldsymbol{d})$ for large n. In this case, almost surely for all data sequences,

$$\lim_{n\to\infty} \frac{\sum_{i=1}^{m} \hat{d}_i(1 - w_{in}(\hat{\boldsymbol{d}}))}{\sum_{i=1}^{m} \hat{d}_i} = \frac{m - m_1}{\sum_{i=1}^{m} d_i^t + m - m_1}. \tag{S-15}$$

In this case, the maximum mpBFDR that can be incurred is at $\beta = 0$, and is given by

$$\lim_{n\to\infty} \mathrm{mpBFDR}_{\beta=0} = \frac{m - m_1}{\sum_{i=1}^m d_i^t + m - m_1}.$$

This is also the maximum mpBFDR that can be incurred among all possible configurations of $G_{m_1+1}, \cdots, G_m$. Hence, for any arbitrary configuration of groups, the maximum mpBFDR that can be incurred lies in the interval $\left(\frac{1}{\sum_{i=1}^m d_i^t+1}, \frac{m-m_1}{\sum_{i=1}^m d_i^t+m-m_1}\right)$ asymptotically. ■

Proof of Theorem 6. Let $\epsilon < E - \alpha$. Then from (16), there exists $n(\epsilon)$ such that for all $n > n(\epsilon)$, $\mathrm{mpBFDR}_{\beta=0} > E - \epsilon > \alpha$. Chandra and Bhattacharya (2019) have shown that mpBFDR is continuous and decreasing in β. Hence, for all $n > n(\epsilon)$, there exists $\beta_n \in (0, 1)$ such that $\mathrm{mpBFDR} = \alpha$.

Now, if possible let $\liminf_{n\to\infty} \beta_n > 0$. Then from Theorem S-4.1 we see that mpBFDR decays to 0 exponentially fast, which contradicts the current situation that $\mathrm{mpBFDR} = \alpha$ for $n > n(\epsilon)$. Hence, $\lim_{n\to\infty} \beta_n = 0$. ■

Proof of Theorem 7. Theorems 3.1 and 3.4 of Chandra and Bhattacharya (2019) together state that mpBFDR is continuous and non-increasing in β. It is to be noted that there is no assumption or restriction on the configurations of G_i's. Hence it is easily seen that pBFDR is also continuous and non-increasing in β.

Let $\hat{\boldsymbol{d}}$ be the optimal decision configuration with respect to the additive loss function. Note that for $\beta = 0$, $\hat{d}_i = 1$ for all i. In that case,

$$\lim_{n\to\infty} \frac{\sum_{i=1}^m \hat{d}_i(1 - v_{in})}{\sum_{i=1}^m \hat{d}_i} = \frac{m_0}{m}.$$

Therefore, it is possible to incur error arbitrarily close to m_0/m for large enough sample size. Hence, the remaining part of the proof follows in the same lines as the arguments in the proof of Theorem 6. ■

Proof of Theorem 8. Take $\epsilon < \frac{m_0}{m} - \alpha$. Since for any multiple testing method, $\mathrm{mpBFDR}_\beta > \mathrm{pBFDR}_\beta$, and since $\lim\limits_{n\to\infty} \mathrm{pBFDR}_{\beta=0} = \frac{m_0}{m}$ by the proof of Theorem 7, it follows that there exists $n_0(\epsilon)$ such that for all $n > n_0(\epsilon)$,

$$\mathrm{mpBFDR}_{\beta=0} > \frac{m_0}{m} - \epsilon > \alpha.$$

Since mpBFDR is continuous and non-increasing in β, for for $n > n_0(\epsilon)$, there exists a sequence $\beta_n \in [0, 1]$ such that

$$\mathrm{mpBFDR}_{\beta_n} = \alpha. \tag{S-16}$$

If possible, let $\liminf_{n\to\infty} \beta_n > 0$. This, however, contradicts Theorem S-4.1 which asserts that mpBFDR decays to 0 exponentially fast. Hence, $\lim_{n\to\infty} \beta_n = 0$. ■

Proof of Theorem 9. From Theorem 6 we have that for any feasible choice of α, there exists a sequence $\{\beta_n\}$ such that $\lim_{n\to\infty} \mathrm{mpBFDR}_{\beta_n} = \alpha$. Now, for the sequence $\{\beta_n\}$, let $\hat{\boldsymbol{d}}_n$ be the optimal decision configuration for sample size n, that is, $\delta_{\mathcal{NM}}\left(\hat{\boldsymbol{d}}_n|\boldsymbol{X}_n\right) = 1$ for sufficiently large n. Following the proof of Theorem 5 and 6 we see that $\hat{d}_{in} = d_i^t$ for $i = 1, \cdots, m_1$ and $\sum_{i=m_1+1}^m \hat{d}_{in} > 0$. Now recall from (10) that for any arbitrary $\epsilon > 0$, there exists $n(\epsilon)$ such that for all $n > n(\epsilon)$, $v_{in} <$

$\exp\left[-n\left(J\left(H_{1i}\right)-\epsilon\right)\right]$ if $d_i^t=0$. Therefore,

$$\begin{aligned}
&\frac{\sum_{i=1}^m(1-\hat{d}_{in})v_{in}}{\sum_{i=1}^m(1-\hat{d}_{in})} \leq \frac{\sum_{i=1}^m(1-d_i^t)v_{in}}{\sum_{i=1}^m(1-\hat{d}_{in})} < e^{n\epsilon}\times\frac{\sum_{i=1}^m(1-d_i^t)e^{-nJ(H_{1i})}}{\sum_{i=1}^m(1-\hat{d}_{in})}\\
\Rightarrow\ &\mathrm{FNR}_{\boldsymbol{X}_n} < e^{n\epsilon}\times\frac{\sum_{i=1}^m(1-d_i^t)e^{-nJ(H_{1i})}}{\sum_{i=1}^m(1-\hat{d}_{in})}\\
\Rightarrow\ &\frac{1}{n}\log\left(\mathrm{FNR}_{\boldsymbol{X}_n}\right) < \epsilon+\frac{1}{n}\log\left[\sum_{i=1}^m(1-d_i^t)e^{-nJ(H_{1i})}\right]-\frac{1}{n}\log\left[\sum_{i=1}^m(1-\hat{d}_{in})\right].
\end{aligned}$$

Note that

$$\begin{aligned}
&\lim_{n\to\infty}\frac{1}{n}\log\left[\sum_{i=1}^m(1-\hat{d}_{in})\right]=0 \text{ as } m \text{ is finite, and}\\
&\lim_{n\to\infty}\frac{1}{n}\log\left[\sum_{i=1}^m(1-d_i^t)e^{-nJ(H_{1i})}\right]=-\tilde{H}_{\min} \text{ from L'Hôpital's rule.}
\end{aligned}$$

As ϵ is any arbitrary positive quantity we have

$$\limsup_{n\to\infty}\frac{1}{n}\log\left(\mathrm{FNR}_{\boldsymbol{X}_n}\right)\leq-\tilde{H}_{\min}.$$

■

S-6 Additional results to Section 5 and proofs

Proof of Theorem 10. The proof of this theorem is complete if (S1)-(S7) are verified for the model (18). We do this through the following lemmas and theorems stated and proved in this section. ■

Lemma S-6.1 *Under the model assumption (C1)-(C2), the KL-divergence rate $h(\boldsymbol{\xi})$ defined in (6) exists and is given by*

$$\begin{aligned}
h(\theta)=\log\left(\frac{\sigma}{\sigma_0}\right)&+\left(\frac{1}{2\sigma^2}-\frac{1}{2\sigma_0^2}\right)\left(\frac{\sigma_0^2}{1-\rho_0^2}+\frac{\boldsymbol{\beta}_0'\boldsymbol{\Sigma}_z\boldsymbol{\beta}_0}{1-\rho_0^2}\right)\\
&+\left(\frac{\rho^2}{2\sigma^2}-\frac{\rho_0^2}{2\sigma_0^2}\right)\left(\frac{\sigma_0^2}{1-\rho_0^2}+\frac{\boldsymbol{\beta}_0'\boldsymbol{\Sigma}_z\boldsymbol{\beta}_0}{1-\rho_0^2}\right)+\frac{1}{2\sigma^2}\boldsymbol{\beta}'\boldsymbol{\Sigma}_z\boldsymbol{\beta}-\frac{1}{2\sigma_0^2}\boldsymbol{\beta}_0'\boldsymbol{\Sigma}_z\boldsymbol{\beta}_0\\
&-\left(\frac{\rho}{\sigma^2}-\frac{\rho_0}{\sigma_0^2}\right)\left(\frac{\rho_0\sigma_0^2}{1-\rho_0^2}+\frac{\rho_0\boldsymbol{\beta}_0'\boldsymbol{\Sigma}_z\boldsymbol{\beta}_0}{1-\rho_0^2}\right)-\left(\frac{\boldsymbol{\beta}}{\sigma^2}-\frac{\boldsymbol{\beta}_0}{\sigma_0^2}\right)'\boldsymbol{\Sigma}_z\boldsymbol{\beta}_0. \qquad \text{(S-17)}
\end{aligned}$$

Proof. It is easy to see that under the true model P,

$$\begin{aligned}
E(x_t)&=\sum_{k=1}^t\rho_0^{t-k}\boldsymbol{z}_k'\boldsymbol{\beta}_0; \qquad \text{(S-18)}\\
E(x_{t+h}x_t)&\sim\frac{\sigma_0^2\rho_0^h}{1-\rho_0^2}+E(x_{t+h})E(x_t);\ h\geq 0,
\end{aligned}$$

where for any two sequences $\{a_t\}_{t=1}^\infty$ and $\{b_t\}_{t=1}^\infty$, $a_t\sim b_t$ stands for $a_t/b_t\to 1$ as $t\to\infty$. Hence,

$$E(x_t^2)\sim\frac{\sigma_0^2}{1-\rho_0^2}+\left(\sum_{k=1}^t\rho_0^{t-k}\boldsymbol{z}_k'\boldsymbol{\beta}_0\right)^2. \qquad \text{(S-19)}$$

Now let

$$\varrho_t = \sum_{k=1}^{t} \rho_0^{t-k} \boldsymbol{z}_k' \boldsymbol{\beta}_0$$

and for $t > t_0$,

$$\tilde{\varrho}_t = \sum_{k=t-t_0}^{t} \rho_0^{t-k} \boldsymbol{z}_k' \boldsymbol{\beta}_0,$$

where, for any $\varepsilon > 0$, t_0 is so large that

$$\frac{C\rho_0^{t_0+1}}{(1-\rho^{t_0})} \leq \varepsilon. \tag{S-20}$$

It follows, using (C2) and (S-20), that for $t > t_0$,

$$|\varrho_t - \tilde{\varrho}_t| \leq \sum_{k=1}^{t-t_0-1} \rho_0^{t-k} \left|\boldsymbol{z}_k' \boldsymbol{\beta}_0\right| \leq \frac{C\rho_0^{t_0+1}(1-\rho_0^{t-t_0+1})}{1-\rho_0} \leq \varepsilon. \tag{S-21}$$

Hence, for $t > t_0$,

$$\tilde{\varrho}_t - \varepsilon \leq \varrho_t \leq \tilde{\varrho}_t + \varepsilon. \tag{S-22}$$

Now,

$$\begin{aligned}\frac{\sum_{t=1}^{n} \tilde{\varrho}_t}{n} &= \rho_0^{t_0} \left(\frac{\sum_{t=1}^{n} \boldsymbol{z}_t}{n}\right)' \boldsymbol{\beta}_0 + \rho_0^{t_0-1} \left(\frac{\sum_{t=2}^{n} \boldsymbol{z}_t}{n}\right)' \boldsymbol{\beta}_0 + \rho_0^{t_0-2} \left(\frac{\sum_{t=3}^{n} \boldsymbol{z}_t}{n}\right)' \boldsymbol{\beta}_0 + \cdots \\ &\quad \cdots + \rho_0 \left(\frac{\sum_{t=t_0}^{n} \boldsymbol{z}_t}{n}\right)' \boldsymbol{\beta}_0 + \left(\frac{\sum_{t=t_0+1}^{n} \boldsymbol{z}_t}{n}\right)' \boldsymbol{\beta}_0 \\ &\rightarrow 0, \text{ as } n \rightarrow \infty, \text{ by virtue of (C1).}\end{aligned}$$

Similarly, it is easily seen, using (C1), that

$$\frac{\sum_{t=1}^{n} \tilde{\varrho}_t^2}{n} \rightarrow \left(\frac{1-\rho^{2(2t_0+1)}}{1-\rho_0^2}\right) \boldsymbol{\beta}_0' \boldsymbol{\Sigma}_z \boldsymbol{\beta}_0, \text{ as } n \rightarrow \infty.$$

Since (S-21) implies that for $t > t_0$, $\tilde{\varrho}_t^2 + \varepsilon^2 - 2\varepsilon\tilde{\varrho}_t \leq \varrho_t^2 \leq \tilde{\varrho}_t^2 + \varepsilon^2 + 2\varepsilon\tilde{\varrho}_t$, it follows that

$$\lim_{n\rightarrow\infty} \frac{\sum_{t=1}^{n} \varrho_t^2}{n} = \lim_{n\rightarrow\infty} \frac{\sum_{t=1}^{n} \tilde{\varrho}_t^2}{n} + \varepsilon^2 = \left(\frac{1-\rho^{2(2t_0+1)}}{1-\rho_0^2}\right) \boldsymbol{\beta}_0' \boldsymbol{\Sigma}_z \boldsymbol{\beta}_0 + \varepsilon^2,$$

and since $\epsilon > 0$ is arbitrary, it follows that

$$\lim_{n\rightarrow\infty} \frac{\sum_{t=1}^{n} \varrho_t^2}{n} = \frac{\boldsymbol{\beta}_0' \boldsymbol{\Sigma}_z \boldsymbol{\beta}_0}{1-\rho_0^2}. \tag{S-23}$$

Hence, it also follows from (S-18), (S-19), (C1) and (S-23), that

$$\frac{\sum_{t=1}^{n} E(x_t^2)}{n} \rightarrow \frac{\sigma_0^2}{1-\rho_0^2} + \frac{\boldsymbol{\beta}_0' \boldsymbol{\Sigma}_z \boldsymbol{\beta}_0}{1-\rho_0^2}; \frac{\sum_{t=1}^{n} E(x_{t-1}^2)}{n} \rightarrow \frac{\sigma_0^2}{1-\rho_0^2} + \frac{\boldsymbol{\beta}_0' \boldsymbol{\Sigma}_z \boldsymbol{\beta}_0}{1-\rho_0^2} \text{ as } n \rightarrow \infty.$$

Now note that

$$x_t x_{t-1} = \rho_0 x_{t-1}^2 + \boldsymbol{z}_t' \boldsymbol{\beta}_0 x_{t-1} + \epsilon_t x_{t-1}. \tag{S-24}$$

Using (20), (S-22) and arbitrariness of $\varepsilon > 0$ it is again easy to see that

$$\frac{\sum_{t=1}^{n} \boldsymbol{z}_t' \boldsymbol{\beta}_0 E(x_{t-1})}{n} \rightarrow 0, \text{ as } n \rightarrow \infty.$$

Also, since for $t = 1, 2, \ldots$, $E(\epsilon_t x_{t-1}) = E(\epsilon_t) E(x_{t-1})$ by independence, and since $E(\epsilon_t) = 0$ for $t = 1, 2, \ldots$, it holds that

$$\frac{\sum_{t=1}^{n} E\left(\epsilon_t x_{t-1}\right)}{n} = 0, \text{ for all } n = 1, 2, \ldots. \tag{S-25}$$

Combining (S-6)-(S-25) we obtain

$$\frac{\sum_{t=1}^{n} E\left(x_t x_{t-1}\right)}{n} \rightarrow \frac{\rho_0 \sigma_0^2}{1-\rho_0^2} + \frac{\rho_0 \boldsymbol{\beta}_0' \boldsymbol{\Sigma}_z \boldsymbol{\beta}_0}{1-\rho_0^2}.$$

Also (C1) along with (S-22) and arbitrariness of $\varepsilon > 0$ yields

$$\frac{\sum_{t=1}^{n} \boldsymbol{z}_t E(x_t)}{n} \rightarrow \boldsymbol{\Sigma}_z \boldsymbol{\beta}_0, \ \frac{\sum_{t=1}^{n} \boldsymbol{z}_t E(x_{t-1})}{n} \rightarrow \boldsymbol{0} \text{ as } n \rightarrow \infty.$$

Using assumptions (C1) and (C2) and the above results, it follows that

$$\begin{aligned} h(\boldsymbol{\xi}) = \lim_{n\rightarrow\infty} \frac{1}{n} E\left[-\log R_n(\boldsymbol{\xi})\right] &= \log\left(\frac{\sigma}{\sigma_0}\right) + \left(\frac{1}{2\sigma^2} - \frac{1}{2\sigma_0^2}\right)\left(\frac{\sigma_0^2}{1-\rho_0^2} + \frac{\boldsymbol{\beta}_0' \boldsymbol{\Sigma}_z \boldsymbol{\beta}_0}{1-\rho_0^2}\right) \\ &+ \left(\frac{\rho^2}{2\sigma^2} - \frac{\rho_0^2}{2\sigma_0^2}\right)\left(\frac{\sigma_0^2}{1-\rho_0^2} + \frac{\boldsymbol{\beta}_0' \boldsymbol{\Sigma}_z \boldsymbol{\beta}_0}{1-\rho_0^2}\right) + \frac{1}{2\sigma^2} \boldsymbol{\beta}' \boldsymbol{\Sigma}_z \boldsymbol{\beta} - \frac{1}{2\sigma_0^2} \boldsymbol{\beta}_0' \boldsymbol{\Sigma}_z \boldsymbol{\beta}_0 \\ &- \left(\frac{\rho}{\sigma^2} - \frac{\rho_0}{\sigma_0^2}\right)\left(\frac{\rho_0 \sigma_0^2}{1-\rho_0^2} + \frac{\rho_0 \boldsymbol{\beta}_0' \boldsymbol{\Sigma}_z \boldsymbol{\beta}_0}{1-\rho_0^2}\right) - \left(\frac{\boldsymbol{\beta}}{\sigma^2} - \frac{\boldsymbol{\beta}_0}{\sigma_0^2}\right)' \boldsymbol{\Sigma}_z \boldsymbol{\beta}_0. \end{aligned}$$

In other words, (S2) holds, with $h(\boldsymbol{\xi})$ given by equation (S-17). ■

Theorem S-6.2 *For each $\boldsymbol{\xi} \in \boldsymbol{\Xi}$, the generalized or relative asymptotic equipartition property holds, and so*

$$\lim_{n\rightarrow\infty} \frac{1}{n} \log R_n(\boldsymbol{\xi}) = -h(\boldsymbol{\xi}).$$

The convergence is uniform over any compact subset of $\boldsymbol{\Xi}$.

Proof. Note that

$$x_t = \sum_{k=1}^{t} \rho_0^{t-k} \boldsymbol{z}_k' \boldsymbol{\beta}_0 + \sum_{k=1}^{t} \rho_0^{t-k} \epsilon_k,$$

where $\tilde{\epsilon}_t = \sum_{k=1}^{t} \rho_0^{t-k} \epsilon_k$ is an asymptotically stationary Gaussian process with mean zero and covariance

$$cov(\tilde{\epsilon}_{t+h}, \tilde{\epsilon}_t) \sim \frac{\sigma_0^2 \rho_0^h}{1-\rho_0^2}, \text{ where } h \geq 0.$$

Then

$$\frac{\sum_{t=1}^{n} x_t^2}{n} = \frac{\sum_{t=1}^{n} \varrho_t^2}{n} + \frac{\sum_{t=1}^{n} \tilde{\epsilon}_t^2}{n} + \frac{2\sum_{t=1}^{n} \tilde{\epsilon}_t \varrho_t}{n}. \tag{S-26}$$

By (S-23), the first term of the right hand side of (S-26) converges to $\frac{\boldsymbol{\beta}_0' \boldsymbol{\Sigma}_z \boldsymbol{\beta}_0}{1-\rho_0^2}$, as $n \rightarrow \infty$, and since $\tilde{\epsilon}_t$; $t = 1, 2, \ldots$ is also an irreducible and aperiodic Markov chain, by the ergodic theorem it follows that the second term of (S-26) converges to $\sigma_0^2/(1-\rho_0^2)$ almost surely, as $n \rightarrow \infty$. Also observe that ϱ_t; $t = 1, 2, \ldots$, is also a sample path of an irreducible and aperiodic stationary Markov chain, with univariate stationary distribution having mean 0 and variance $\boldsymbol{\beta}_0'\boldsymbol{\Sigma}_z\boldsymbol{\beta}_0 \times \lim_{t\to\infty} \sum_{k=1}^{t} \rho_0^{2(t-k)} = \frac{\boldsymbol{\beta}_0'\boldsymbol{\Sigma}_z\boldsymbol{\beta}_0}{1-\rho_0^2}$. Since for each t, $\tilde{\epsilon}_t$ and ϱ_t are independent, $\tilde{\epsilon}_t\varrho_t$; $t = 1, 2, \ldots$, is also an irreducible and aperiodic Markov chain having a stationary distribution with mean 0 and variance $\frac{\sigma_0^2\boldsymbol{\beta}_0'\boldsymbol{\Sigma}_z\boldsymbol{\beta}_0}{(1-\rho_0^2)^2}$. Hence, by the ergodic theorem, the third term of (S-26) converges to zero, almost surely, as $n \to \infty$. It follows that

$$\frac{\sum_{t=1}^{n} x_t^2}{n} \to \frac{\sigma_0^2}{1-\rho_0^2} + \frac{\boldsymbol{\beta}_0'\boldsymbol{\Sigma}_z\boldsymbol{\beta}_0}{1-\rho_0^2}, \tag{S-27}$$

and similarly,

$$\frac{\sum_{t=1}^{n} x_{t-1}^2}{n} \to \frac{\sigma_0^2}{1-\rho_0^2} + \frac{\boldsymbol{\beta}_0'\boldsymbol{\Sigma}_z\boldsymbol{\beta}_0}{1-\rho_0^2}. \tag{S-28}$$

Now, since $x_t = \varrho_t + \tilde{\epsilon}_t$, it follows using (C1) and (S-22) that

$$\lim_{n\to\infty} \frac{\sum_{t=1}^{n} \boldsymbol{z}_t x_t}{n} = \left(\lim_{n\to\infty} \frac{\sum_{t=1}^{n} \boldsymbol{z}_t \boldsymbol{z}_t'}{n}\right)\boldsymbol{\beta}_0 + \lim_{n\to\infty} \frac{\sum_{t=1}^{n} \boldsymbol{z}_t \tilde{\epsilon}_t}{n}. \tag{S-29}$$

By (C1), the first term on the right hand side of (S-29) is $\boldsymbol{\Sigma}_z\boldsymbol{\beta}_0$. For the second term, note that it follows from (C1) that $\boldsymbol{z}_t\tilde{\epsilon}_t$; $t = 1, 2, \ldots$, is sample path of an irreducible and aperiodic Markov chain with a stationary distribution having zero mean. Hence, by the ergodic theorem, it follows that the second term of (S-29) is $\mathbf{0}$, almost surely. In other words, almost surely,

$$\frac{\sum_{t=1}^{n} \boldsymbol{z}_t x_t}{n} \to \boldsymbol{\Sigma}_z\boldsymbol{\beta}_0, \text{ as } n \to \infty, \tag{S-30}$$

and similar arguments show that, almost surely,

$$\frac{\sum_{t=1}^{n} \boldsymbol{z}_t x_{t-1}}{n} \to \mathbf{0}, \text{ as } n \to \infty. \tag{S-31}$$

We now calculate the limit of $\sum_{t=1}^{n} x_t x_{t-1}/n$, as $n \to \infty$. By (S-24),

$$\lim_{n\to\infty} \frac{\sum_{t=1}^{n} x_t x_{t-1}}{n} = \lim_{n\to\infty} \frac{\rho_0 \sum_{t=1}^{n} x_{t-1}^2}{n} + \lim_{n\to\infty} \frac{\boldsymbol{\beta}_0' \sum_{t=1}^{n} \boldsymbol{z}_t x_{t-1}}{n} + \lim_{n\to\infty} \frac{\sum_{t=1}^{n} \epsilon_t x_{t-1}}{n}. \tag{S-32}$$

By (S-28), the first term on the right hand side of (S-32) is given, almost surely, by $\frac{\rho_0\sigma_0^2}{1-\rho_0^2} + \frac{\rho_0\boldsymbol{\beta}_0'\boldsymbol{\Sigma}_z\boldsymbol{\beta}_0}{1-\rho_0^2}$, and the second term is almost surely zero due to (S-31). For the third term, note that $\epsilon_t x_{t-1} = \epsilon_t\varrho_{t-1} + \epsilon_t\tilde{\epsilon}_{t-1}$. Both $\epsilon_t\varrho_{t-1}$; $t = 1, 2, \ldots$ and $\epsilon_t\tilde{\epsilon}_{t-1}$; $t = 1, 2, \ldots$, are sample paths of irreducible and aperiodic Markov chains having stationary distributions with mean zero. Hence, by the ergodic theorem, the third term of (S-32) is zero, almost surely. That is,

$$\lim_{n\to\infty} \frac{\sum_{t=1}^{n} x_t x_{t-1}}{n} = \frac{\rho_0\sigma_0^2}{1-\rho_0^2} + \frac{\rho_0\boldsymbol{\beta}_0'\boldsymbol{\Sigma}_z\boldsymbol{\beta}_0}{1-\rho_0^2}. \tag{S-33}$$

The limits (S-27), (S-28), (S-30), (S-31), (S-33) applied to $\log R_n(\boldsymbol{\xi})$ given by Theorem S-6.2, shows that $\frac{\log R_n(\theta)}{n}$ converges to $-h(\theta)$ almost surely as $n \to \infty$. In other words, (S3) holds.

Now $\frac{1}{n}\log R_n(\boldsymbol{\xi})$ has continuous partial derivatives implying that $\frac{\partial}{\partial\boldsymbol{\xi}}\left[\frac{1}{n}\log R_n(\boldsymbol{\xi})\right]$ is bounded in any compact set. Hence $\frac{1}{n}\log R_n(\boldsymbol{\xi})$ is Lipschitz continuous and hence stochastic equicontinuous in $\boldsymbol{\xi}$. Thus by applying the stochastic Ascoli theorem we have that the convergence is uniform over $\boldsymbol{\xi}$ in that compact set (for details about stochastic equicontinuity, see, for example, Billingsley 2013). ■

The meaning of Theorem S-6.2 is that, relative to the true distribution, the likelihood of each $\boldsymbol{\xi}$ goes to zero exponentially, the rate being the Kullback-Leibler divergence rate. Roughly speaking, an integral of exponentially-shrinking quantities will tend to be dominated by the integrand with the slowest rate of decay. Lemma S-6.1 and Theorem S-6.2 imply that (S1)-(S3) hold. For any $\boldsymbol{\xi} \in \Xi$, $h(\boldsymbol{\xi})$ is finite, which implies that $(S4)$ also holds. As regards (S5), we can always make (S-3) to hold by considering $\mathcal{G}_n$s as credible regions of the prior distribution and these can be chosen increasing compact sets without loss of generality. Since $h(\cdot)$ is continuous in $\boldsymbol{\xi}$ the second and third parts of $(S5)$ will also hold.

Note that the maximizer of $R_n(\boldsymbol{\xi})$ is the maximum likelihood estimator (mle) of $\boldsymbol{\xi}$. Let $\hat{\boldsymbol{\xi}}_n = \sup_{\boldsymbol{\xi}\in\mathcal{G}_n} R_n(\boldsymbol{\xi})$. Then

$$\frac{1}{n}\log\int_{\mathcal{G}_n} R_n(\boldsymbol{\xi})\pi(\boldsymbol{\xi})d\boldsymbol{\xi} \leq \frac{1}{n}\log\left[R_n(\hat{\boldsymbol{\xi}}_n)\pi(\mathcal{G}_n)\right]. \tag{S-34}$$

If we can show that $\hat{\boldsymbol{\xi}}_n$ is a consistent estimator of $\boldsymbol{\xi}_0$, then this will validate (S6). Importantly, the conditions for mle consistency generally require iid observations (Lehmann and Casella, 1998). In this model the data sequence $\{x_t\}_{t=1}^{\infty}$ have dependence structure and regular asymptotic theory will not hold. Hence, we provide a direct proof of consistency; below we provide the main results leading to the desired consistency result. The equipartition property plays a crucial role in the proceeding.

Theorem S-6.3 *The function $\frac{1}{n}\log R_n(\boldsymbol{\xi})$ is asymptotically concave in $\boldsymbol{\xi}$.*

Proof. Note that

$$\sup_{\boldsymbol{\xi}\in\Xi}\frac{1}{n}\log R_n(\boldsymbol{\xi}) = \sup_{\rho,\boldsymbol{\beta}}\sup_{\sigma^2}\frac{1}{n}\log R_n(\boldsymbol{\xi}) = -\inf_{\rho,\boldsymbol{\beta}}\ \log\left[\frac{1}{n}\sum_{t=1}^{n}\left(x_t-\rho x_{t-1}-\boldsymbol{\beta}'\boldsymbol{z}_t\right)^2\right]-\frac{1}{2}.$$

Since log is a monotonic function, minimizing $\log\left[\frac{1}{n}\sum_{t=1}^{n}\left(x_t-\rho x_{t-1}-\boldsymbol{\beta}'\boldsymbol{z}_t\right)^2\right]$ is equivalent to minimizing $\frac{1}{n}\sum_{t=1}^{n}\left(x_t-\rho x_{t-1}-\boldsymbol{\beta}'\boldsymbol{z}_t\right)^2 = g_n(\rho,\boldsymbol{\beta})$, say. Now the Jacobian matrix J of $g_n(\rho,\boldsymbol{\beta})$ is given by

$$J = \begin{bmatrix}\frac{1}{n}\sum x_{t-1}^2 & \frac{1}{n}\sum x_{t-1}\boldsymbol{z}_t' \\ \frac{1}{n}\sum x_{t-1}\boldsymbol{z}_t & \frac{1}{n}\sum \boldsymbol{z}_t\boldsymbol{z}_t'\end{bmatrix}.$$

(S-27), (S-31) together with the model assumptions (C1)-(C2) clearly shows that for large enough n, J is positive-definite. Hence $g_n(\rho,\boldsymbol{\beta})$ is convex implying that $\frac{1}{n}\log R_n(\boldsymbol{\xi})$ is a concave function for large n. ■

The above theorem ensures that for large enough n, the likelihood equation have unique mle. Rest we need to ensure the strong consistency of the mle for this dependent setup.

Theorem S-6.4 *Given any $\eta > 0$, the log-likelihood ratio $\frac{1}{n}\log R_n(\boldsymbol{\xi})$ has its unique root in the η-neighbourhood of $\boldsymbol{\xi}_0$ almost surely for large n.*

Proof. (C3) ensures that $\boldsymbol{\xi}_0$ is an interior point in Ξ, implying that there exists a compact set $G\subset\Xi$ such that $\boldsymbol{\xi}_0$ is an interior point of G also. From Theorem S-6.2, for each $\boldsymbol{\xi}\in\Xi$, we have

$$\lim_{n\to\infty}\frac{1}{n}\log R_n(\boldsymbol{\xi}) = -h(\boldsymbol{\xi}), \tag{S-35}$$

and the convergence in (S-35) is uniform over $\boldsymbol{\xi}$ in G. Thus,

$$\lim_{n\to\infty}\sup_{\boldsymbol{\xi}\in G}\left|\frac{1}{n}\log R_n(\boldsymbol{\xi})+h(\boldsymbol{\xi})\right| = 0. \tag{S-36}$$

For any $\eta > 0$, we define

$$N_\eta(\boldsymbol{\xi}_0) = \{\boldsymbol{\xi} : \|\boldsymbol{\xi}_0-\boldsymbol{\xi}\| < \eta\};\ N_\eta'(\boldsymbol{\xi}_0) = \{\boldsymbol{\xi} : \|\boldsymbol{\xi}_0-\boldsymbol{\xi}\| = \eta\};\ \overline{N}_\eta(\boldsymbol{\xi}_0) = \{\boldsymbol{\xi} : \|\boldsymbol{\xi}_0-\boldsymbol{\xi}\| \leq \eta\}.$$

Note that for sufficiently small η, $\overline{N}_\eta(\boldsymbol{\xi}_0)\subset G$. Let $H = \inf_{\boldsymbol{\xi}\in N_\eta'(\boldsymbol{\xi}_0)} h(\boldsymbol{\xi})$. By the properties of KL-divergence $h(\boldsymbol{\xi})$ is minimum at $\boldsymbol{\xi}=\boldsymbol{\xi}_0$ and therefore, $H > 0$. Let us fix an ε such that $0<\varepsilon<H$.

Then by (S-36), for large enough n all $\boldsymbol{\xi} \in N'_\eta(\boldsymbol{\xi}_0)$, $\frac{1}{n}\log R_n(\boldsymbol{\xi}) < -h(\boldsymbol{\xi}) + \varepsilon < 0$. Now by definition $\frac{1}{n}\log R_n(\boldsymbol{\xi}_0) = 0$ and thus for all $\boldsymbol{\xi} \in N'_\eta(\boldsymbol{\xi}_0)$

$$\frac{1}{n}\log R_n(\boldsymbol{\xi}) < \frac{1}{n}\log R_n(\boldsymbol{\xi}_0) \tag{S-37}$$

for large enough n. Now, $\overline{N}_\eta(\boldsymbol{\xi}_0)$ is a compact set with $N'_\eta(\boldsymbol{\xi}_0)$ being its boundary. Since $\frac{1}{n}\log R_n(\boldsymbol{\xi})$ is continuous in $\boldsymbol{\xi}$, it is bounded in $\overline{N}_\eta(\boldsymbol{\xi}_0)$. From (S-37) we see that the maximum is attained at some interior point of $\overline{N}_\eta(\boldsymbol{\xi}_0)$ and not on the boundary. Since the supremum is attained at an interior point of $\overline{N}_\eta(\boldsymbol{\xi}_0)$, the supremum is also a local maximum. Now, Theorem S-6.3 ensures that for large n the maximizer of $\frac{1}{n}\log R_n(\boldsymbol{\xi})$ is unique. This proves the result. ■

Theorem S-6.4 essentially entails the strong consistency of the mle. This also leads to the verification of (S6) required for posterior consistency. We formally state it in the following lemma.

Lemma S-6.5 *For any proper prior distribution $\pi(\cdot)$ over the parameter space Ξ, we have*

$$\limsup_{n\to\infty} \frac{1}{n}\log \int_{\mathcal{G}_n} R_n(\boldsymbol{\xi})\pi(\boldsymbol{\xi})d\boldsymbol{\xi} \le 0.$$

Proof. From Theorems S-6.2 and S-6.4 we have

$$\lim_{n\to\infty} \frac{1}{n}\log R_n(\hat{\boldsymbol{\xi}}_n) = 0,$$

and hence

$$\limsup_{n\to\infty} \frac{1}{n}\log \int_{\mathcal{G}_n} R_n(\boldsymbol{\xi})\pi(\boldsymbol{\xi})d\boldsymbol{\xi} \le \lim_{n\to\infty} \frac{1}{n}\left[\log R_n(\hat{\boldsymbol{\xi}}_n) + \log \pi(\mathcal{G}_n)\right] \le 0.$$

■

Lemma S-6.5 signifies that (S6) holds. About (S7), it trivially holds since $h(\cdot)$ is a continuous function.

S-7 Supplementary to real data analysis

Table S-1: Causal SNPs for different populations

Popu-lation	Causal SNP

1	m1, m12, m13, m114, m135, m146, m147, m236, m249, m274, m275, m276, m407, m422, m449, m537, m620, m665, m674, m680, m709, m765, m887, m894, m895, m899, m934, m951, m955, m1076, m1161, m1234, m1249, m1291, m1328, m1412, m1436, m1437, m1445, m1456, m1575, m1646, m1733, m1761, m1762, m1763, m1764, m1765, m1766, m1767, m1768, m1946, m2043, m2093, m2169, m2174, m2175, m2205, m2287, m2348, m2349, m2374, m2403, m2451, m2452, m2467, m2468, m2508, m2610, m2677, m2678, m2679, m2680, m2681, m2682, m2687, m2688, m2689, m2692, m2817, m2906, m2907, m2943, m2951, m2952, m2953, m2954, m2955, m2956, m2962, m2996, m2997, m3106, m3279, m3280, m3281, m3282, m3283, m3358, m3418, m3457, m3489, m3490, m3491, m3545, m3571, m3644, m3735, m3738, m3795, m3931, m3951, m3952, m4015, m4038, m4144, m4188, m4281, m4297, m4372, m4373, m4374, m4375, m4499, m4500, m4504, m4506, m4538, m4674, m4766, m4767, m4768, m4919, m4924, m4925, m4973, m4974, m5041, m5149, m5199, m5228, m5318, m5352, m5353, m5411, m5437, m5505, m5515, m5516, m5517, m5646, m5688, m5728, m5766, m5926, m5927, m6025, m6066, m6116, m6117, m6158, m6159, m6160, m6161, m6296, m6359, m6365, m6394, m6395, m6396, m6397, m6398, m6399, m6400, m6401, m6402, m6434, m6466, m6473, m6505, m6507, m6573, m6574, m6599, m6617, m6723, m6757, m6765, m6766, m6816, m6817, m6818, m6851, m6852, m6853, m6858, m6859, m6860, m6872, m6903, m6995, m7085, m7089, m7156, m7202, m7253, m7325, m7338, m7348
2	m1, m147, m432, m440, m458, m589, m597, m598, m599, m600, m741, m1010, m1011, m1039, m1046, m1047, m1048, m1049, m1050, m1051, m1052, m1053, m1120, m1350, m1362, m1620, m1670, m1812, m2014, m2027, m2028, m2143, m2144, m2200, m2201, m2203, m2213, m2295, m2421, m2439, m2521, m2569, m2573, m2586, m2795, m2797, m3216, m3412, m3560, m3615, m3727, m3728, m3729, m3730, m3956, m4141, m4273, m4328, m4421, m4453, m4454, m4510, m4742, m4776, m4777, m4809, m4826, m4827, m4828, m4988, m5229, m5375, m5542, m5544, m5590, m5674, m5803, m5804, m5805, m5885, m5886, m5887, m5936, m5997, m6004, m6016, m6017, m6018, m6019, m6312, m6320, m6342, m6343, m6457, m6485, m6486, m6492, m6493, m6652, m7178, m7189, m7220, m7269, m7270
3	m1, m113, m162, m172, m176, m196, m198, m443, m446, m538, m678, m777, m796, m896, m917, m945, m946, m947, m1042, m1238, m1318, m1324, m1325, m1468, m1740, m1905, m1906, m2039, m2062, m2148, m2162, m2163, m2197, m2202, m2265, m2409, m2435, m2502, m2627, m2628, m2629, m2691, m2881, m2988, m3414, m3415, m3438, m3439, m3440, m3441, m3442, m3443, m3446, m3447, m3448, m3484, m3543, m3544, m3811, m3848, m3849, m3850, m3851, m4297, m4318, m4425, m4493, m4578, m5030, m5031, m5094, m5223, m5379, m5380, m5448, m5685, m5706, m5799, m5808, m5911, m5931, m6095, m6096, m6395, m6396, m6397, m6401, m6434, m6473, m6481, m6527, m6729, m7059, m7066

4	m1, m61, m71, m72, m143, m144, m181, m182, m183, m225, m294, m295, m555, m707, m799, m805, m816, m937, m1020, m1156, m1262, m1312, m1316, m1318, m1323, m1335, m1377, m1408, m1448, m1483, m1484, m1485, m1486, m1487, m1488, m1586, m1714, m1715, m1716, m1717, m1776, m1822, m1840, m1909, m1946, m1947, m1948, m1949, m2013, m2120, m2121, m2129, m2130, m2131, m2139, m2179, m2191, m2227, m2228, m2263, m2286, m2413, m2433, m2434, m2481, m2507, m2621, m2785, m2816, m2832, m2895, m2902, m2922, m2923, m2937, m2964, m2965, m3168, m3170, m3211, m3290, m3312, m3378, m3412, m3493, m3495, m3496, m3604, m3614, m3680, m3819, m3820, m3837, m3838, m3839, m3840, m3978, m4069, m4150, m4165, m4166, m4216, m4217, m4218, m4219, m4220, m4221, m4222, m4234, m4235, m4351, m4374, m4375, m4377, m4378, m4504, m4580, m4581, m4582, m4583, m4584, m4585, m4586, m4587, m4588, m4612, m4637, m4648, m4692, m4712, m4713, m4714, m4715, m4776, m4818, m4833, m4918, m5016, m5079, m5152, m5153, m5154, m5233, m5234, m5237, m5326, m5379, m5453, m5454, m5455, m5456, m5457, m5747, m5789, m5794, m5825, m5854, m5859, m5893, m5894, m5904, m5923, m5948, m6084, m6152, m6153, m6154, m6155, m6431, m6432, m6438, m6445, m6502, m6503, m6509, m6560, m6756, m6853, m6869, m6870, m6897, m6898, m6899, m6900, m6967, m7071, m7072, m7257, m7267, m7268, m7272
5	m1, m96, m329, m363, m431, m904, m951, m1406, m1893, m2150, m2357, m2359, m2360, m2463, m2547, m2551, m2570, m2621, m2622, m2623, m3287, m3983, m3984, m3985, m4822, m5168, m5186, m5222, m5223, m5404, m5405, m5416, m5425, m5699, m5706, m5880, m5881, m5914, m5925, m5926, m5927, m5928, m5929, m5930, m5931, m6408, m6440, m6494, m6533, m6538, m6562, m7067, m7076, m7078, m7080, m7081, m7082, m7083, m7091, m7092, m7093, m7188, m7223, m7224, m7227, m7248, m7249

References

Billingsley P (2013) Convergence of Probability Measures. John Wiley & Sons

Chandra NK, Bhattacharya S (2019) Non-marginal decisions: A novel Bayesian multiple testing procedure. *Electronic Journal of Statistics* 13(1):489–535, DOI 10.1214/19-EJS1535

Lehmann EL, Casella G (1998) Theory of Point Estimation. Springer-Verlag, New York, Inc

Shalizi CR (2009) Dynamics of Bayesian updating with dependent data and misspecified models. *Electronic Journal of Statistics* 3:1039–1074, DOI 10.1214/09-EJS485

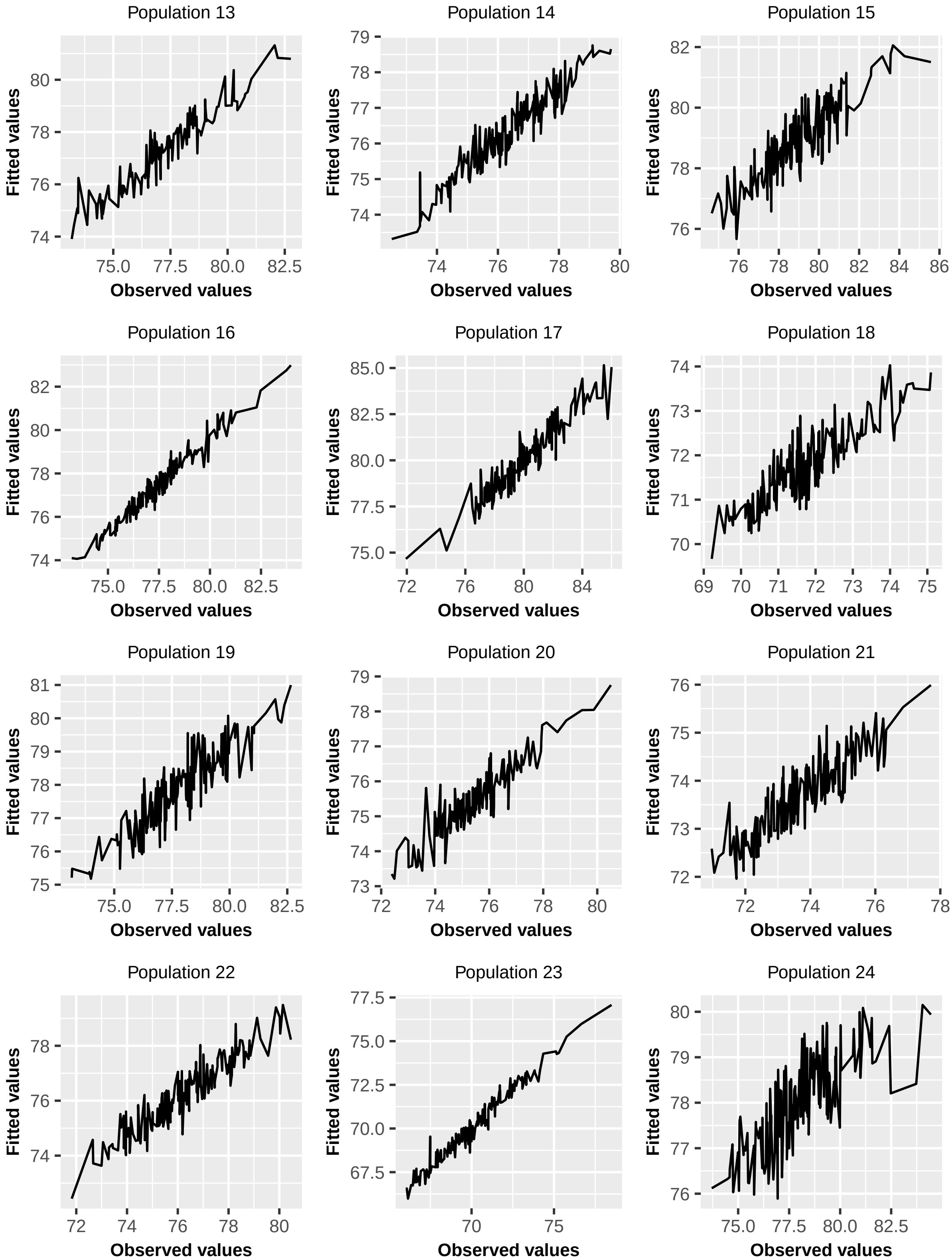

Figure S-1: Fitted values versus observed values.