SUPPLEMENTARY MATERIAL TO "SELECTIVE INFERENCE FOR FALSE DISCOVERY PROPORTION IN A HIDDEN MARKOV MODEL"

This supplement provides additional materials for the posterior distribution, the algorithms, the computation of the lower bounds, the estimation of f_0 and some theoretical insight. It also presents additional numerical experiments and an additional application to the real data set considered by Sun and Cai (2009).

CONTENTS

S1. Posterior distribution. We provide here the details for the quantities involved in the posterior distribution, see Proposition 2.1. We recall that $\pi = (\pi_0, \pi_1)$ denotes the marginal stationary distribution of the underlying Markov Chain. Let for $i \in \mathbb{N}_m$ and $q \in \{0, 1\}$,

$$\alpha_i(q) = \mathbb{P}_\Gamma(X_{1:i}, \theta_i = q);$$

$$\beta_i(q) = \mathbb{P}_\Gamma(X_{(i+1):m} \mid \theta_i = q);$$

$$\ell_{i,q}(\Gamma) = \mathbb{P}_\Gamma(\theta_i = q \mid X_{1:m}) = \frac{\alpha_i(q)\beta_i(q)}{\sum_{k\in\{0,1\}} \alpha_i(k)\beta_i(k)}, \tag{S1}$$

with the common notation for which $\mathbb{P}_\Gamma(X_{1:i}, \theta_i = q)$ denotes the density of $(X_{1:i}, \theta_i)$ taken at point $(X_{1:i}, q)$ and $\mathbb{P}_\Gamma(X_{(i+1):m} \mid \theta_i = q)$ denotes the density of $X_{(i+1):m}$ conditionally

MSC2020 subject classifications: Primary 00X00, 00X00; secondary 00X00.
Keywords and phrases: First keyword, second keyword.

on $\theta_i = q$ taken at point $X_{(i+1):m}$. These quantities can be obtained through the standard forward-backward algorithm, that is,

(S2) $$\alpha_1(q) = \pi_q f_q(X_1),\ \ \alpha_{i+1}(q) = f_q(X_{i+1}) \sum_{\ell\in\{0,1\}} a_{\ell,q}\alpha_i(\ell),\ \ 1 \le i \le m-1$$

(S3) $$\beta_m(q) = 1,\ \ \beta_{i-1}(q) = \sum_{\ell\in\{0,1\}} f_\ell(X_i) a_{q,\ell}\beta_i(\ell),\ \ 2 \le i \le m.$$

Further, we let for $i \in \{2, \dots, m\}$ and $q, q' \in \{0, 1\}$,

(S4) $$\boldsymbol{\ell}_{i,q,q'}(\Gamma) = \mathbb{P}_\Gamma(\theta_i = q', \theta_{i-1} = q | X_{1:m}) = \frac{\beta_i(q')\alpha_{i-1}(q) f_{q'}(X_i) a_{q,q'}}{\sum_{k\in\{0,1\}} \beta_i(k)\alpha_i(k)};$$

(S5) $$\Pi_{i,q,q'}(\Gamma) = \mathbb{P}_\Gamma(\theta_i = q' | \theta_{i-1} = q, X_{1:m}) = \frac{\boldsymbol{\ell}_{i,q,q'}(\Gamma)}{\boldsymbol{\ell}_{i-1,q}(\Gamma)} = \frac{\beta_i(q') f_{q'}(X_i) a_{q,q'}}{\beta_{i-1}(q)};$$

(S6) $$\Pi_i(\Gamma) = \begin{pmatrix} \Pi_{i,0,0}(\Gamma) & \Pi_{i,0,1}(\Gamma) \\ \Pi_{i,1,0}(\Gamma) & \Pi_{i,1,1}(\Gamma) \end{pmatrix}.$$

S2. Bootstrap bounds.

S2.1. *First bootstrap bound.* Recall for convenience the (deterministic) quantity:

(S7) $$q_{1,\gamma}(\beta, S(\cdot); \Gamma) := \min\left\{ x \in \mathbb{R} \,:\, \mathbb{P}_\Gamma\left(U_\beta(X, S(X); \Gamma) - U_\beta(X, S(X); \widehat{\Gamma}) \le x \right) \ge 1 - \gamma \right\},$$

which corresponds to the $(1-\gamma)$-quantile of the distribution of $U_\beta(X, S(X); \Gamma) - U_\beta(X, S(X); \widehat{\Gamma})$ when X is generated according to the true model parameter Γ. Then, we have by definition, for all $\delta \in (0,1)$,

$$\mathbb{P}_\Gamma\left(\text{FDP}(\theta, S(X)) \le U_{\beta(1-\delta)}(X, S(X); \Gamma) \mid X \right) \ge 1 - \beta(1-\delta);$$

$$\mathbb{P}_\Gamma\left(U_{\beta(1-\delta)}(X, S(X); \Gamma) - U_{\beta(1-\delta)}(X, S(X); \widehat{\Gamma}) \le q_{1,\beta\delta}(\beta(1-\delta), S(\cdot); \Gamma) \right) \ge 1 - \beta\delta.$$

Note that the first bound concerns the distribution of θ conditionally on X while the second one concerns only the marginal distribution of X. Therefore, this immediately implies a bound with respect to the joint distribution of (θ, X).

PROPOSITION S1. *For any selection policy $S(\cdot) : x \in \mathcal{X} \mapsto S(x) \subset \mathbb{N}_m$ and $\delta \in (0,1)$, for any model parameter Γ, we have*

(S8) $$\mathbb{P}_\Gamma\left(\textit{FDP}(\theta, S(X)) \le U_{\beta(1-\delta)}(X, S(X); \widehat{\Gamma}) + q_{1,\beta\delta}(\beta(1-\delta), S(\cdot); \Gamma) \right) \ge 1 - \beta.$$

Hence, the RHS of the event in (S8) is a post selection bound in the sense of (4). However, again, it depends on the unknown parameter Γ, although it is only via the quantity $q_{1,\gamma}(\beta, S(\cdot); \Gamma)$, that we would like to think of as a "second order" term.

Now, plugging the estimate $\widehat{\Gamma}$ of Γ in the latter leads to the bootstrap-type approximation $q_\gamma(\beta, S(\cdot); \Gamma) \approx q_\gamma(\beta, S(\cdot); \widehat{\Gamma})$ which corresponds to consider the $(1-\gamma)$-quantile of the distribution of $U_\beta(X^*, S(X^*); \widehat{\Gamma}) - U_\beta(X^*, S(X^*); \widehat{\Gamma}^*)$ under the model parameter $\widehat{\Gamma}$, where X^* is an independent sample generated under $P_{\widehat{\Gamma}}$ and $\widehat{\Gamma}^* = \widehat{\Gamma}(X^*)$, see (S7). As usual, the bootstrap quantity $q_\gamma(\beta, S(\cdot); \widehat{\Gamma})$ is in turn approximated by a quantity $\tilde{q}_{1,\gamma}^{(B)}(\beta, S(\cdot); \widehat{\Gamma})$ obtained via a Monte-Carlo scheme that generates B times the sequence $(X_i^*)_{1\le i\le m}$ according to the model with parameter $\widehat{\Gamma}$.

We are now in position to introduce our first bootstrap bound:

$$U^{\text{boot1}}_{\beta,\delta}(X,S(\cdot)):=U_{\beta(1-\delta)}(X,S(X);\widehat{\Gamma})+\tilde{q}^{(B)}_{1,\beta\delta}(\beta(1-\delta),S(\cdot);\widehat{\Gamma}), \tag{S9}$$

Algorithm 3 describes how this bound may be computed.

Let us make some comments on this bound:

- This bootstrap bound has a semi-parametric flavor: while the θ^* sequence is generated via a parametric Markov chain using the estimated parameters, the resampled data X_i^* are obtained by weighted smooth bootstrap (Efron, 1979), since they are drawn from a weighted kernel density estimator (see Section S3.3 for more details).
- The term δ balances the way the errors are distributed and should be chosen to derive an appropriate trade-off: a small δ will sharpen the bound $U_{\beta(1-\delta)}(X,S(X);\widehat{\Gamma})$ but makes $\tilde{q}^{(B)}_{1,\beta\delta}(\beta(1-\delta),S(\cdot);\widehat{\Gamma})$ larger. In the numerical experiments, we have observed that the impact of δ is moderate.
- In this bound, we should compute $S(X^*)$ for bootstrap samples, see Algorithm 4. This means that the user should provide their whole selection policy $S: x\in\mathcal{X}\mapsto S(x)\subset\mathbb{N}_m$, and not only the selection $S(X)$ on the observed data set X. This can be seen as a constraint in some cases, and the next paragraph provides a solution to circumvent this limitation.
- Whenever $\tilde{q}^{(B)}_{1,\beta\delta}(\beta(1-\delta),S(\cdot);\widehat{\Gamma})$ is nonnegative (which happens with high probability for suitable values of β and δ), we have $U^{\text{boot1}}_{\beta,\delta}(X,S(\cdot))\geq U_{\beta(1-\delta)}(S(X);\widehat{\Gamma})$ so the obtained bound is in general more conservative than the plug-in bound.

S2.2. *Second bootstrap bound.* As mentioned above, computing $U^{\text{boot1}}_{\beta}(\cdot)$ requires that the user provides $S(X^{*(b)})$ for every single bootstrap sample $X^{*(b)}$, $b=1,\dots,B$. This could be inconvenient if the user does not want to provide the whole selection policy, but only $S(X)$, the desired selection on the observed data set and not on other virtual data sets.

We circumvent this limitation with a heuristic twist: imagine first that we have at hand a sample Y that is an independent copy of X. Then one could estimate Γ by an estimator $\widehat{\Gamma}(Y)$ computed from the sample Y. Therefore,

$$\begin{aligned}U_{\beta(1-\delta)}(X,S(X);\Gamma)&=U_{\beta(1-\delta)}(X,S(X);\widehat{\Gamma}(Y))+\Big(U_{\beta(1-\delta)}(X,S(X);\Gamma)-U_{\beta(1-\delta)}(X,S(X);\widehat{\Gamma}(Y))\Big)\\&\leq U_{\beta(1-\delta)}(X,S(X);\widehat{\Gamma}(Y))+q_{2,\beta\delta}(\beta(1-\delta),S(X);\Gamma),\end{aligned}$$

with probability larger than $1-\beta\delta$, where we denote

$$\begin{aligned}&q_{2,\gamma}(\beta,S(X);\Gamma)\\&:=\min\left\{x\in\mathbb{R}\,:\,\mathbb{P}_{Y\sim P_\Gamma}\Big(U_{\beta(1-\delta)}(X,S(X);\Gamma)-U_{\beta(1-\delta)}(X,S(X);\widehat{\Gamma}(Y))\leq x\,\Big|\,X\Big)\geq 1-\gamma\right\}\end{aligned}$$

the $(1-\gamma)$-quantile of the distribution of $U_{\beta(1-\delta)}(X,S(X);\Gamma)-U_{\beta(1-\delta)}(X,S(X);\widehat{\Gamma}(Y))$ conditionally on X, when $Y\sim P_\Gamma$. Similarly to above, the latter can be approximated via a bootstrapped quantity $q_{2,\gamma}(\beta,S(X);\Gamma)\approx q_{2,\gamma}(\beta,S(X);\widehat{\Gamma}(Y))$ which is the γ-quantile of the distribution of $U_{\beta(1-\delta)}(X,S(X);\widehat{\Gamma}(Y))-U_{\beta(1-\delta)}(X,S(X);\widehat{\Gamma}(Y^*))$ conditionally on X,Y, when $Y^*\sim P_{\widehat{\Gamma}}$, which itself is approximated by applying a Monte-Carlo approximation scheme, giving rise to $\tilde{q}^{(B)}_{2,\gamma}(\beta,S(X);\widehat{\Gamma}(Y))$. Now, since we do not have access to a different sample Y, we propose to use X in place of Y, which leads to the following new bootstrap bound (denoting, again, $\widehat{\Gamma}=\widehat{\Gamma}(X)$):

$$U^{\text{boot2}}_{\beta,\delta}(X,S(X)):=U_{\beta(1-\delta)}(X,S(X);\widehat{\Gamma})+\tilde{q}^{(B)}_{2,\beta\delta}(\beta(1-\delta),S(X);\widehat{\Gamma}), \tag{S10}$$

Algorithm 3 provides full details about the computation of this bound.

Let us make the following comments on the second bootstrap bound:

- $\tilde{q}^{(B)}_{2,\beta\delta}(\beta(1-\delta), S(X); \widehat{\Gamma})$ does not depend on the full selection policy $S(\cdot) : x \in \mathcal{X} \mapsto S(x) \subset \mathbb{N}_m$, but only on the set $S(X)$;
- Whenever $\tilde{q}^{(B)}_{2,\beta\delta}(\beta(1-\delta), S(X); \widehat{\Gamma})$ is nonnegative, the obtained bound is more conservative than the plug-in bound. However, while $U^{\text{boot2}}_{\beta,\delta}$ includes the variability in $\widehat{\Gamma}$, this bound ignores the variations in $S(X)$, so is in general less conservative than $U^{\text{boot1}}_{\beta,\delta}$.

S2.3. *Third bootstrap bound.* An elementary point is that the interval $I_\alpha(S(\cdot)) = [0, q^{\text{naive}}_\beta(S(\cdot); \Gamma)]$ satisfies the unconditional coverage (4) when choosing

$$q_{\text{naive},\beta}(S(\cdot); \Gamma) := \min \left\{ x \in \mathbb{R} \,:\, \mathbb{P}_\Gamma \left(\text{FDP}(\theta, S(X)) \leq x \right) \geq 1 - \beta \right\}.$$

Note that this bound is "unconditional" as it is directly based on the distribution of $\text{FDP}(\theta, S(X))$ when drawing $(\theta, X) \sim P_\Gamma$. Hence, it uses the selection policy $S(\cdot)$. This leads to consider the bound $q_{\text{naive},\beta}(S(\cdot); \widehat{\Gamma})$, which relies on drawing independent couples (θ^*, X^*) from the distribution $P_{\widehat{\Gamma}}$. Then, $q_{\text{naive},\beta}(S(\cdot); \widehat{\Gamma})$ is approximated by $\tilde{q}^{(B)}_{\text{naive},\beta}(S(\cdot); \widehat{\Gamma})$ via a Monte-Carlo scheme that generates B times the couple (θ^*, X^*) according to the model with parameter $\widehat{\Gamma}$. We finally let

$$U^{\text{naive}}_\beta(S(\cdot)) := \tilde{q}^{(B)}_{\text{naive},\beta}(S(\cdot); \widehat{\Gamma}).$$

This is different from the above-mentioned bounds (bootstrap 1, bootstrap 2), which are based only on the marginal X^*.

As demonstrated on experiments in Section 4, using U^{naive}_β is generally too conservative, in particular if the true $\text{FDP}(\theta, S(X))$ has large variability from a realization of (θ, X) to another, since it is based on an unconditional quantile of its distribution. To alleviate this effect, an improvement can be derived by using a proper re-centering by the plug-in bound $U_\beta(X, S(X); \widehat{\Gamma})$ given in (17), that acts like a stabilization:

$$V_\beta(X, S(\cdot); \Gamma) := U_\beta(X, S(X); \widehat{\Gamma}(X)) + q_{3,\beta}(\beta, S(\cdot); \Gamma), \qquad \text{with} \tag{S11}$$

$$q_{3,\beta}(\beta, S(\cdot); \Gamma) := \min \left\{ x \in \mathbb{R} \,:\, \mathbb{P}_{(\theta,X)\sim P_\Gamma} \left(\text{FDP}(\theta, S(X)) - U_\beta(X, S(X); \widehat{\Gamma}(X)) \leq x \right) \geq 1 - \beta \right\}, \tag{S12}$$

where $U_\beta(X, S(X); \Gamma)$ is given by (13) and (16). By definition, $V_\beta(X, S(\cdot); \Gamma) - U_\beta(X, S(X); \widehat{\Gamma}(X))$ is deterministic and we have

$$\begin{aligned} &1 - \beta \\ &\leq \mathbb{P}_{(\theta,X)\sim P_\Gamma} \left(\text{FDP}(\theta, S(X)) - U_\beta(X, S(X); \widehat{\Gamma}(X)) \leq V_\beta(X, S(\cdot); \Gamma) - U_\beta(X, S(X); \widehat{\Gamma}(X)) \right) \\ &= \mathbb{P}_{(\theta,X)\sim P_\Gamma} \left(\text{FDP}(\theta, S(X)) \leq V_\beta(X, S(\cdot); \Gamma) \right). \end{aligned}$$

This gives that $V_\beta(X, S(\cdot); \Gamma)$ is a valid post selection bound, which can be approximated by the bootstrap bound

$$V_\beta(X, S(\cdot); \widehat{\Gamma}) := U_\beta(X, S(X); \widehat{\Gamma}(X)) + q_{3,\beta}(\beta, S(\cdot); \widehat{\Gamma}).$$

Given (S12), this bound also relies on drawing the couple (θ^*, X^*) from the distribution $P_{\widehat{\Gamma}}$. As earlier, $q_{3,\beta}(\beta, S(\cdot); \widehat{\Gamma})$ is approximated by $\tilde{q}^{(B)}_{3,\beta}(\beta, S(\cdot); \widehat{\Gamma})$ via a Monte-Carlo scheme that generates B times the couple (θ^*, X^*) according to the model with parameter $\widehat{\Gamma}$. This leads to the final bound:

$$U_\beta^{\text{boot3}}(X, S(\cdot)) := U_\beta(X, S(X); \widehat{\Gamma}) + \tilde{q}_{3,\beta}^{(B)}(\beta, S(\cdot); \widehat{\Gamma}). \tag{S13}$$

Algorithm 3 gives full details about the computation of this bound.

The difference between $U_\beta^{\text{boot3}}(X, S(\cdot))$ and $U_\beta^{\text{naive}}(S(\cdot))$ is that our plug-in bound $U_\beta(X, S(X); \widehat{\Gamma})$ is used to recenter the FDP. As a result, it "stabilizes" the bound $U_\beta^{\text{naive}}(S(\cdot))$: while $U_\beta^{\text{naive}}(S(\cdot))$ depends on X only via $\widehat{\Gamma}$, the bound $U_\beta^{\text{boot3}}(X, S(\cdot))$ also depends on the plug-in bound $U_\beta(X, S(X); \widehat{\Gamma}(X))$. Note that knowledge of the full selection policy $S(\cdot)$ is required to compute both $U^{\text{naive}}(S(\cdot))$ and $U_\beta^{\text{boot3}}(X, S(\cdot))$.

S3. Algorithms.

S3.1. *Algorithm 1: nonparametric HMM estimation.* In this section we give details on the estimation of the transition probabilities for the hidden Markov states and (nonparametric) emission densities of the observables in a HMM, as mentioned in Section 2.2. The initialization $\widehat{\Gamma}^{(0)}$ of $\widehat{\Gamma}$ in Algorithm 1 is done as follows: $\widehat{A}^{(0)}$ is computed first by estimating the stationary distribution, that is, $(\pi_0, 1-\pi_0)$, by using a standard Storey estimator $\widehat{\pi}_0^{(0)}$ $(\lambda = 0.8)$ Storey (2002). Then, the parameter $\widehat{A}_{1,1}^{(0)}$ is taken uniformly over the possible range for this parameter, intersected with $[0.6, 1]$ to ensure that the null class is predominant. On the other hand, $\widehat{f}_1^{(0)} = (\widehat{f} - \widehat{\pi}_0^{(0)} f_0)/(1 - \widehat{\pi}_0^{(0)})$, where $\widehat{f}$ is a Kernel-based estimation of the density of the measurements. The algorithm is stopped if the distance between to successive iteration of $\widehat{\Gamma}$ (measured in infinite norm, restricted on the observations for the densities) is below 10^{-4}.

Algorithm 1: EM-type algorithm for the estimate $\widehat{\Gamma} = (\widehat{A}, \widehat{f}_1)$

Input: $X = (X_1, \ldots, X_m)$
Output: $\widehat{\Gamma} = (\widehat{A}, \widehat{f}_1)$ estimator of $\Gamma = (A, f_1)$

- Initialization: first guess $\widehat{\Gamma}^{(0)} = (\widehat{A}^{(0)}, \widehat{f}_1^{(0)})$ of Γ
- Loop: at step $t \geq 1$, given $\widehat{\Gamma}^{(t-1)}$, do
  - Compute
    1. $\alpha_i(q) = \alpha_i(q; \widehat{\Gamma}^{(t-1)})$ and $\beta_i(q) = \beta_i(q; \widehat{\Gamma}^{(t-1)})$, $1 \leq i \leq m$, $q \in \{0, 1\}$, using the forward-backward algorithm (S2), (S3);
    2. $\boldsymbol{\ell}_{i,q}(\widehat{\Gamma}^{(t-1)})$, $1 \leq i \leq m$, $q \in \{0, 1\}$, by (S1);
    3. $\boldsymbol{\ell}_{i,q,q'}(\widehat{\Gamma}^{(t-1)})$, $1 \leq i \leq m$, $q, q' \in \{0, 1\}$, by (S4).
  - Update
    1. $\widehat{a}_{q,q'}^{(t)} = \sum_{i=1}^{m-1} \boldsymbol{\ell}_{i,q,q'}(\widehat{\Gamma}^{(t-1)}) / \sum_{i=1}^{m-1} \boldsymbol{\ell}_{i,q}(\widehat{\Gamma}^{(t-1)})$, for $0 \leq q, q' \leq 1$;
    2. $\widehat{A}^{(t)} = \begin{pmatrix} \widehat{a}_{0,0}^{(t)} & \widehat{a}_{0,1}^{(t)} \\ \widehat{a}_{1,0}^{(t)} & \widehat{a}_{1,1}^{(t)} \end{pmatrix}$;
    3. $\widehat{f}_1^{(t)}$ by (12);
    4. $\widehat{\Gamma}^{(t)} = (\widehat{A}^{(t)}, \widehat{f}_1^{(t)})$.
  - Stop the loop if convergence, that is, $\widehat{\Gamma}^{(t)}$ close enough to $\widehat{\Gamma}^{(t-1)}$.
- Return $\widehat{\Gamma} = \widehat{\Gamma}^{(t)}$

S3.2. *Algorithm 2: computation of key quantities for FDP control in a HMM.* In this section we give details on the algorithmic computation of the intermediate quantities $B_{k,\ell,0}, B_{k,\ell,1}$ defined by (15) in Section 3.1.

Let us denote for $t \geq 2$,

$$\Pi_t^R(\Gamma) = \begin{pmatrix} \Pi_{t,0,0}^R(\Gamma) \, \Pi_{t,0,1}^R(\Gamma) \\ \Pi_{t,1,0}^R(\Gamma) \, \Pi_{t,1,1}^R(\Gamma) \end{pmatrix}.$$

Algorithm 2: Computation of $B_{k,\ell,0}$ and $B_{k,\ell,1}$.

Input: $\Pi_k^R(\Gamma)$, $1 \leq k \leq s$; $\boldsymbol{\ell}_{j_k,q}(\Gamma)$, $1 \leq k \leq s$, $q \in \{0,1\}$
Output: $B_{k,\ell,0}$ and $B_{k,\ell,1}$, $1 \leq k \leq s$, $0 \leq \ell \leq s$
Initialization $B_{k,0,0} = 0$, $1 \leq k \leq s$, $B_{1,\ell,0} = \boldsymbol{\ell}_{j_1,0}(\Gamma)$, $1 \leq \ell \leq s$, $B_{1,\ell,1} = \boldsymbol{\ell}_{j_1,1}(\Gamma)$, $0 \leq \ell \leq s$.

for $2 \leq k \leq s$ **do**
  $B_{k,0,1} = B_{k-1,0,1}\Pi_{k,1,1}^R(\Gamma)$
  for $1 \leq \ell \leq s$ **do**
    (S14) $B_{k,\ell,0} = B_{k-1,\ell-1,0}\, \Pi_{k,0,0}^R(\Gamma) + B_{k-1,\ell-1,1}\, \Pi_{k,1,0}^R(\Gamma);$
    (S15) $B_{k,\ell,1} = B_{k-1,\ell,0}\, \Pi_{k,0,1}^R(\Gamma) + B_{k-1,\ell,1}\, \Pi_{k,1,1}^R(\Gamma).$
  end
end

In words, (S14) comes from the fact that having at most ℓ zero-occurrences in $\theta_{1:k}^R$ with a zero in the last position means that we have at most $\ell - 1$ zero-occurrences in $\theta_{1:(k-1)}^R$ with either a zero or a one in position $k-1$. As for (S15), the fact that having at most ℓ zero-occurrences in $\theta_{1:k}^R$ with a one in the last position means that we have at most ℓ zero-occurrences in $\theta_{1:(k-1)}^R$ with either a zero or a one in position $k-1$.

In practice for a large s Algorithm 2 can be time consuming because it requires the estimation of 2 $s \times s$ matrices. To speed up the algorithm we notice that it is not necessary to compute the matrices $B_{k,\ell,0}$ and $B_{k,\ell,1}$ for all $\ell \in \{1,\dots,s\}$ to get $U_\beta(X, S(X), \Gamma)$. Indeed, we can stop for u such that $B_{k,u,1} + B_{k,u,0} \geq 1 - \beta$: as shown in Proposition 3.1 $U_\beta(X, S(X), \Gamma) = u$.

S3.3. *Algorithm 3: bootstrap-based bounds for FDP.* In this section we give algorithmic details on the computation of the bootstrap-based bounds defined in Section S2.

Drawing $\theta^{*(b)}$ and $X^{*(b)}$ from $\widehat{\Gamma}$.

1. Draw a Markov chain $\theta^{*(b)}$ of size m with transition matrix $\widehat{A}$, and initial distribution the stationary distribution, namely $(\frac{\widehat{a}_{1,0}}{\widehat{a}_{0,1}+\widehat{a}_{1,1}}, 1 - \frac{\widehat{a}_{1,0}}{\widehat{a}_{0,1}+\widehat{a}_{1,1}})$.
2. For $1 \leq i \leq m$ draw independently $X_i^{*(b)}$ as (case f_0 known):

$$X_i^{*(b)} \sim \begin{cases} f_0 & \text{if } \theta_i^{*(b)} = 0; \\ \widehat{f}_1 & \text{if } \theta_i^{*(b)} = 1; \end{cases}$$

   Drawing from $\widehat{f}_1$ can be done easily by noting that $\widehat{f}_1$ is the density of a mixture of Gaussian distributions:

$$\widehat{f}_1 = \sum_{i=1}^m w_i \, \mathcal{N}(X_i, h^2), \quad w_i = \frac{\boldsymbol{\ell}_{i,1}(X; \widehat{\Gamma})}{\sum_{j=1}^m \boldsymbol{\ell}_{j,1}(X; \widehat{\Gamma})}, \; 1 \leq i \leq m.$$

If f_0 is unknown, the only difference is that $X_i^{*(b)} \sim \widehat{f}_0$ when $\theta_i^{*(b)} = 0$. Drawing from $\widehat{f}_0$ is made similarly to the case of $\widehat{f}_1$.

Empirical quantiles. To obtain the empirical quantiles $\tilde{q}_{i,\gamma}$ of $D_i = (D_i^{(1)}, \ldots, D_i^{(B)})$ we start by ordering D_i. More precisely let $b_1, \ldots, b_B$ such that $D_i^{(b_1)} \leq D_i^{(b_2)} \leq \cdots \leq D_i^{(b_B)}$ then we define $\tilde{q}_{i,\gamma} = D_i^{(b_j)}$, where j is the smallest integer larger or equal to γB.

Algorithm 3: Core algorithm for computing the bootstrap bounds

Input: Common input for all the bounds : Data X; Estimator $\widehat{\Gamma}$ via Algorithm 1 if f_0 is known (or by Algorithm 4 if f_0 is unknown); $\beta \in (0,1)$
- Input U^{boot1}: Selection policy $S(\cdot) : x \in \mathcal{X} \mapsto S(x) \subset \mathbb{N}_m$; , $\delta \in (0,1)$
- Input U^{boot2}: $\delta \in (0,1)$, $S(X)$
- Input U^{boot3}: Selection policy $S(\cdot) : x \in \mathcal{X} \mapsto S(x) \subset \mathbb{N}_m$

Output: $U_{\beta,\delta}^{\text{boot1}}(S(\cdot), \widehat{\Gamma})$ or $U_{\beta,\delta}^{\text{boot2}}(S(X), \widehat{\Gamma})$ or $U_{\beta}^{\text{boot3}}(S(\cdot), \widehat{\Gamma})$

1. Generate independently B bootstrap samples as follows: for $1 \leq b \leq B$
   a) Draw $\theta^{*(b)}$ and $X^{*(b)}$ from $\widehat{\Gamma}$
   b) Compute $\widehat{\Gamma}^{*(b)}$ using Algorithm 1 with the sequence $X^{*(b)}$ (or using Algorithm 4 if f_0 is unknown);
   c) Compute $D_i^{(b)}$, for the appropriate $i \in (1,2,3)$ and $\lambda \in (\beta(1-\delta), \beta(1-\delta), \beta)$;
      - U^{boot1}: $D_1^{(b)} = U_\lambda(X^{*(b)}, S(X^{*(b)}); \widehat{\Gamma}) - U_\lambda(X^{*(b)}, S(X^{*(b)}); \widehat{\Gamma}^{*(b)})$;
      - U^{boot2}: $D_2^{(b)} = U_\lambda(X, S(X); \widehat{\Gamma}) - U_\lambda(X, S(X); \widehat{\Gamma}^{*(b)})$;
      - U^{boot3}: $D_3^{(b)} = \text{FDP}(\theta^{*(b)}, S(X^{*(b)})) - U_\lambda(X^{*(b)}, S(X^{*(b)}); \widehat{\Gamma}^{*(b)})$,

   where $U_\lambda(X, S(X); \Gamma)$ is given by (13) and (16)
2. Compute $\tilde{q}_{i,\gamma}^{(B)}(\lambda)$ as the empirical γ-quantile of $D_i = (D_i^{(1)}, \ldots, D_i^{(B)})$ for the appropriate $\gamma \in (\beta\delta, \beta\delta, \beta)$;
3. Return the corresponding bound
   - $U_{\beta,\delta}^{\text{boot1}}(X, S(\cdot), \widehat{\Gamma}) = U_{\beta(1-\delta)}(X, S(X), \widehat{\Gamma}) + \tilde{q}_{1,\beta\delta}^{(B)}(\beta(1-\delta))$
   - $U_{\beta,\delta}^{\text{boot2}}(X, S(X), \widehat{\Gamma}) = U_{\beta(1-\delta)}(X, S(X), \widehat{\Gamma}) + \tilde{q}_{2,\beta\delta}^{(B)}(\beta(1-\delta))$
   - $U_{\beta}^{\text{boot3}}(X, S(X), \widehat{\Gamma}) = U_{\beta}(X, S(X), \widehat{\Gamma}) + \tilde{q}_{3,\beta}^{(B)}(\beta)$

S4. Lower confidence bounds. Expressions for lower bounds can be obtained similarly to those of upper bounds. We provide them in this section for completeness.

Oracle and plug-in lower bounds. According to (9), we have (using the notation of Section 3.1, e.g., $R = S(X)$, $s = |S(X)|$):

$$L_\beta(X, R; \Gamma) = s^{-1} \max \left\{ n \in \{0, \ldots, m\} \, : \, \mathbb{P}_\Gamma \left(\sum_{t=1}^{s} (1 - \theta_t^R) \geq n \,\middle|\, X \right) \geq 1 - \beta \right\}.$$

Similarly to Proposition 3.1 the quantity $L_\beta(X, S(X); \Gamma)$ can be computed as

$$L_\beta(X, R; \Gamma) = s^{-1} \max \{ n \in \{0, \ldots, m\} \, : \, B_{s,n-1,0} + B_{s,n-1,1} \leq \beta \}.$$

Adding the convention $B_{s,-1,0} = B_{s,-1,1} = 0$. The plug-in lower bound is given by $L_\beta(X, S(X); \widehat{\Gamma})$.

First bootstrap lower bound. Proceeding as for the upper bound (see Section S2.1), we obtain

$$\mathbb{P}_\Gamma(\text{FDP}(\theta, S(X)) \geq L_{\beta(1-\delta)}(X, S(X); \Gamma) \mid X) \geq 1 - \beta(1-\delta)$$

$$\mathbb{P}_\Gamma(L_{\beta(1-\delta)}(X, S(X); \Gamma) - L_{\beta(1-\delta)}(X, S(X); \widehat{\Gamma}) \geq q^\ell_{1,\beta\delta}(\beta(1-\delta), S(\cdot); \Gamma)) \geq 1 - \beta\delta.$$

by letting

$$q^\ell_{1,\gamma}(\beta, S(\cdot); \Gamma)) = \max\{x \in \mathbb{R} : \ \mathbb{P}_\Gamma(L_\beta(X, S(X), \Gamma) - L_\beta(X, S(X), \widehat{\Gamma}) \geq x) \geq 1 - \gamma\}.$$

This gives

$$\mathbb{P}_\Gamma(\text{FDP}(\theta, S(X)) \geq L_{\beta(1-\delta)}(X, S(X); \widehat{\Gamma}) + q^\ell_{1,\beta\delta}(\beta(1-\delta), S(\cdot); \Gamma)) \geq 1 - \beta. \tag{S16}$$

and leads to the boot1 lower bound:

$$L^{\text{boot1}}_{\beta,\delta}(X, S(\cdot)) = L_{\beta(1-\delta)}(X, S(X); \widehat{\Gamma}) + \tilde{q}^{\ell,(B)}_{1,\beta\delta}(\beta(1-\delta), S(\cdot); \widehat{\Gamma}),$$

where $\tilde{q}^{\ell,(B)}_{1,\beta\delta}(\beta(1-\delta), S(\cdot); \widehat{\Gamma})$ is the Monte-Carlo approximation of the bootstrap quantile $q^\ell_{1,\beta\delta}(\beta(1-\delta), S(\cdot); \widehat{\Gamma})$, itself being an approximation of $q^\ell_{1,\beta\delta}(\beta(1-\delta), S(\cdot); \Gamma)$. In practice, $\tilde{q}^{\ell,(B)}_{1,\beta\delta}(\beta(1-\delta), S(\cdot); \widehat{\Gamma})$ can be derived as $E_1^{(b_i)}$ with $i = \lfloor B\beta\delta \rfloor + 1$ where $E_1^{(b_1)} \leq E_1^{(b_2)} \leq \cdots \leq E_1^{(b_B)}$ with

$$E_1^{(b)} = L_{\beta(1-\delta)}(X^{*(b)}, S(X^{*(b)}), \widehat{\Gamma}) - L_{\beta(1-\delta)}(X^{*(b)}, S(X^{*(b)}), \widehat{\Gamma}^{*(b)}), 1 \leq b \leq B.$$

Second bootstrap lower bound. The same heuristic as for U^{boot2} (see Section S2.2) gives:

$$L^{\text{boot2}}_{\beta,\delta}(X, S(X)) = L_{\beta(1-\delta)}(X, S(X); \widehat{\Gamma}) + \tilde{q}^{\ell,(B)}_{2,\beta\delta}(\beta(1-\delta), S(X); \widehat{\Gamma}),$$

where $\tilde{q}^{\ell,(B)}_{2,\beta\delta}(\beta(1-\delta), S(X); \widehat{\Gamma}) = E_2^{(b_i)}$ with $i = \lfloor B\beta\delta \rfloor + 1$ where $E_2^{(b_1)} \leq E_2^{(b_2)} \leq \cdots \leq E_2^{(b_B)}$ with

$$E_2^{(b)} = L_{\beta(1-\delta)}(X, S(X), \widehat{\Gamma}) - L_{\beta(1-\delta)}(X, S(X), \widehat{\Gamma}^{*(b)}), 1 \leq b \leq B.$$

Third lower bootstrap bound. Using similar arguments as for U^{boot3} (see Section S2.3), we obtain the bound

$$L^{\text{boot3}}_{\beta}(X, S(\cdot)) = L_\beta(X, S(X), \widehat{\Gamma}) + \tilde{q}^{\ell,(B)}_{3,\beta}(\beta, S(\cdot); \widehat{\Gamma}),$$

where $\tilde{q}^{\ell,(B)}_{3,\beta}(\beta, S(X); \widehat{\Gamma}) = E_3^{(b_i)}$ with $i = \lfloor B\beta \rfloor + 1$ where $E_3^{(b_1)} \leq E_3^{(b_2)} \leq \cdots \leq E_3^{(b_B)}$ with

$$E_3^{(b)} = \text{FDP}(X^{*(b)}, S(X^{*(b)})) - L_\beta(X^{*(b)}, S(X^{*(b)}), \widehat{\Gamma}^{*(b)}), 1 \leq b \leq B.$$

S5. Estimation of the null distribution.

Estimation of f_0 for our post selection bounds. When f_0 is unknown, we should estimate it along with A, f_1 in Algorithm 1. The following algorithm builds an estimator $\widehat{f_0}$ in the same spirit as $\widehat{f_1}$ (12), by replacing the $\boldsymbol{\ell}_{i,1}$ by $\boldsymbol{\ell}_{i,0}$, that is,

$$\widehat{f_0}^{(t)}(x) = \sum_{i=1}^m \boldsymbol{\ell}_{i,0}(\widehat{\Gamma}^{(t-1)}) \frac{K((x - X_i)/h)}{h} / \sum_{i=1}^m \boldsymbol{\ell}_{i,0}(\widehat{\Gamma}^{(t-1)}). \tag{S17}$$

Algorithm 4: EM-type algorithm to derive $\widehat{\Gamma} = (\widehat{A}, \widehat{f}_0, \widehat{f}_1)$ (with f_0 unknown).

Input: $X = (X_1, \dots, X_m)$
Output: $\widehat{\Gamma} = (\widehat{A}, \widehat{f}_0, \widehat{f}_1)$ estimator of $\Gamma = (A, f_0, f_1)$

- Initialization: first guess $\widehat{\Gamma}^{(0)} = (\widehat{A}^{(0)}, \widehat{f}_0^{(0)}, \widehat{f}_1^{(0)})$ of Γ
- Loop: at step $t \geq 1$, given $\widehat{\Gamma}^{(t-1)}$, do
  - Compute
    1. $\alpha_i(q) = \alpha_i(q; \widehat{\Gamma}^{(t-1)})$ and $\beta_i(q) = \beta_i(q; \widehat{\Gamma}^{(t-1)})$, $1 \leq i \leq m$, $q \in \{0,1\}$, using the forward-backward algorithm (S2), (S3);
    2. $\boldsymbol{\ell}_{i,q}(X; \widehat{\Gamma}^{(t-1)})$, $1 \leq i \leq m$, $q \in \{0,1\}$, by using (S1);
    3. $\boldsymbol{\ell}_{i,q,q'}(X; \widehat{\Gamma}^{(t-1)})$, $1 \leq i \leq m$, $q, q' \in \{0,1\}$, by using (S4).
  - Update
    1. $\widehat{a}_{q,q'}^{(t)} = \sum_{i=1}^{m-1} \boldsymbol{\ell}_{i,q,q'}(\widehat{\Gamma}^{(t-1)}) / \sum_{i=1}^{m-1} \boldsymbol{\ell}_{i,q}(\widehat{\Gamma}^{(t-1)})$, for $0 \leq q, q' \leq 1$;
    2. $\widehat{A}^{(t)} = \begin{pmatrix} \widehat{a}_{0,0}^{(t)} & \widehat{a}_{0,1}^{(t)} \\ \widehat{a}_{1,0}^{(t)} & \widehat{a}_{1,1}^{(t)} \end{pmatrix}$;
    3. $\widehat{f}_0^{(t)}$ using (S17);
    4. $\widehat{f}_1^{(t)}$ using (12);
    5. $\widehat{\Gamma}^{(t)} = (\widehat{A}^{(t)}, \widehat{f}_0^{(t)}, \widehat{f}_1^{(t)})$.
  - Stop the loop if convergence, that is, $\widehat{\Gamma}^{(t)}$ close enough to $\widehat{\Gamma}^{(t-1)}$.
- Return $\widehat{\Gamma} = \widehat{\Gamma}^{(t)}$

Note that, at the end of Algorithm 4, the user can define the “null” state according to their preference. For instance:

- define the "null" state as the predominant one, that is, the most probable one according to the stationary distribution of $\widehat{A}$;
- define the "null" state according to the density among $\{\widehat{f}_0, \widehat{f}_1\}$ whose mean is closer to 0.

Estimation of F_0 for p-value-based post selection bounds. The estimated p-values are given by $\widehat{p}_i = 2\Big(\min(1 - \widehat{F}_0(X_i), \widehat{F}_0(X_i)\Big)$, where

$$\widehat{F}_0(x) = \frac{\displaystyle\sum_{i=1}^{m} \mathbb{1}\{X_i \leq x\} \boldsymbol{\ell}_{i,0}(\widehat{\Gamma})}{\displaystyle\sum_{i=1}^{m} \boldsymbol{\ell}_{i,0}(\widehat{\Gamma})}.$$

S6. Towards plug-in consistency. The aim of this section is to provide sufficient conditions in order to ensure that the plug-in bound is asymptotically valid, as m tends to infinity. This supports the discussion made in Section 6.3.

LEMMA S1. *Assume that $\widehat{\Gamma}$ is an estimator of Γ such that*

$$\forall \Gamma,\ d_{tv}\left(\mathcal{D}_{\widehat{\Gamma}}\left(\theta \,\middle|\, X\right), \mathcal{D}_{\Gamma}\left(\theta \,\middle|\, X\right)\right) \text{ converges in probability to } 0 \text{ under } \Gamma, \tag{S18}$$

where $\mathcal{D}_{\Gamma}\left(\theta \,\middle|\, X\right)$ denotes the conditional distribution of θ conditionally on X under P_Γ and d_{tv} denotes the total variation distance. Then the plug-in bound $U_\beta^{PI}(X, S(X))$ (17) satisfies that

$$\forall \Gamma,\ \liminf_m \left\{\mathbb{P}_\Gamma\left(\mathit{FDP}(\theta, S(X)) \leq U_\beta^{PI}(X, S(X))\right)\right\} \geq 1 - \beta.$$

PROOF. Since $U_\beta^{\text{PI}}(X,S(X)) = U_\beta(X,S(X);\widehat{\Gamma})$, we have point-wise in X, for all Γ,

$$\begin{aligned}
&\mathbb{P}_\Gamma\left(\text{FDP}(\theta,S(X)) \le U_\beta^{\text{PI}}(X,S(X)) \,\middle|\, X\right)\\
&= \mathbb{P}_\Gamma\left(\text{FDP}(\theta,S(X)) \le U_\beta(X,S(X);\widehat{\Gamma}) \,\middle|\, X\right)\\
&\ge \mathbb{P}_{\widehat{\Gamma}}\left(\text{FDP}(\theta,S(X)) \le U_\beta(X,S(X);\widehat{\Gamma}) \,\middle|\, X\right)\\
&\quad - \sup_{n\in[0,m]}\left\{\mathbb{P}_\Gamma\left(\sum_{i\in S(X)}(1-\theta_i)\le n \,\middle|\, X\right) - \mathbb{P}_{\widehat{\Gamma}}\left(\sum_{i\in S(X)}(1-\theta_i)\le n \,\middle|\, X\right)\right\}\\
&\ge 1-\beta - d_{tv}\left(\mathcal{D}_{\widehat{\Gamma}}\left(\theta \,\middle|\, X\right), \mathcal{D}_\Gamma\left(\theta \,\middle|\, X\right)\right),
\end{aligned}$$

by definition (13) of the functional U_β. This entails the result by integrating over X. □

S7. Additional numerical experiments. In this section we present additional numerical experiments. We have modified the simulation described Section 4 by changing either the model parameters $\widehat{\Gamma}$ or the selection policies $S(\cdot)$. In all these additional experiments, the number of hypotheses is $m = 3200$, the number of runs is 300 and the risk is $\beta = 10\%$.

S7.1. *Invalid selection policies.* The selection policies used in Figure (S1) produce invalid bounds (even the oracle bound) as they depend of some knowledge of $\mathcal{H}_0$, the set of true null hypotheses.

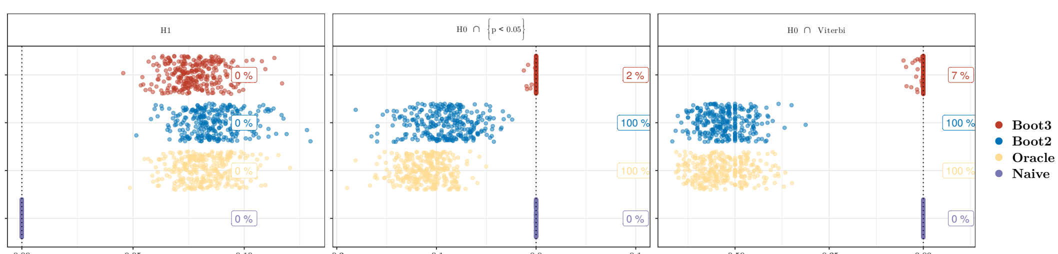

FIG S1. *Plot similar to Figure 3 (A) (same model and simulation parameters) but when the selection policy is of the form $S(X,\mathcal{H}_0)$, that is, depends on some knowledge of $\mathcal{H}_0$.Recall that the targeted level β is* 10%.

S7.2. *Independent states or small determinant.* In Figure S2, we set $a_{0,0} = 0.95$, and the value of $a_{1,1}$ is modified in order to achieve the desired value of $\det(A)$.

S7.3. *Unknown f_0.* Figure S3 presents the results of numerical experiments in which f_0 is unknown and has to be estimated from the data as well.

S7.4. *Different values of m.* Figure S4 presents the results of numerical experiments for different values of m. The simulation settings is the one described in Section 4.2.

S7.5. *Stationarity.* Figure S5 presents the results of numerical experiments where the settings is the one described in Section 4.2 expept that we force $\theta_1 = 1$.

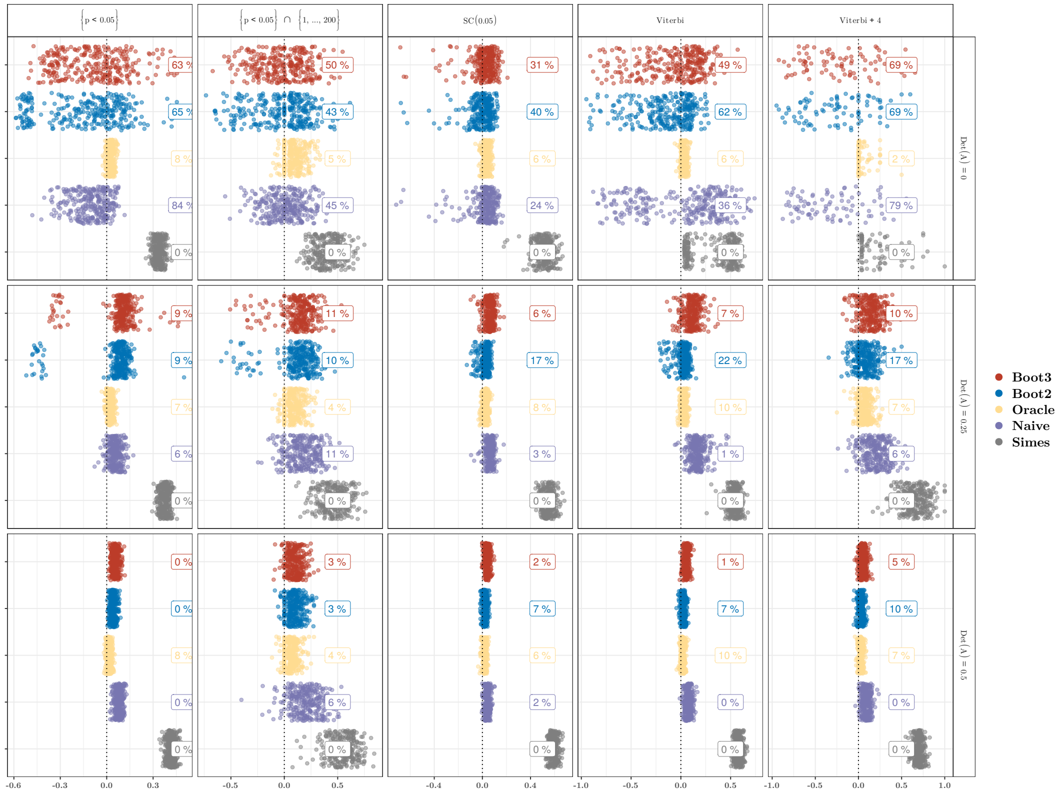

FIG S2. *Plot similar to Figure 3 (A) (same densities f_0 and f_1, and same simulation parameters), for different model parameters making $\det(A)$ small (rows). Recall that the targeted level β is 10%.*

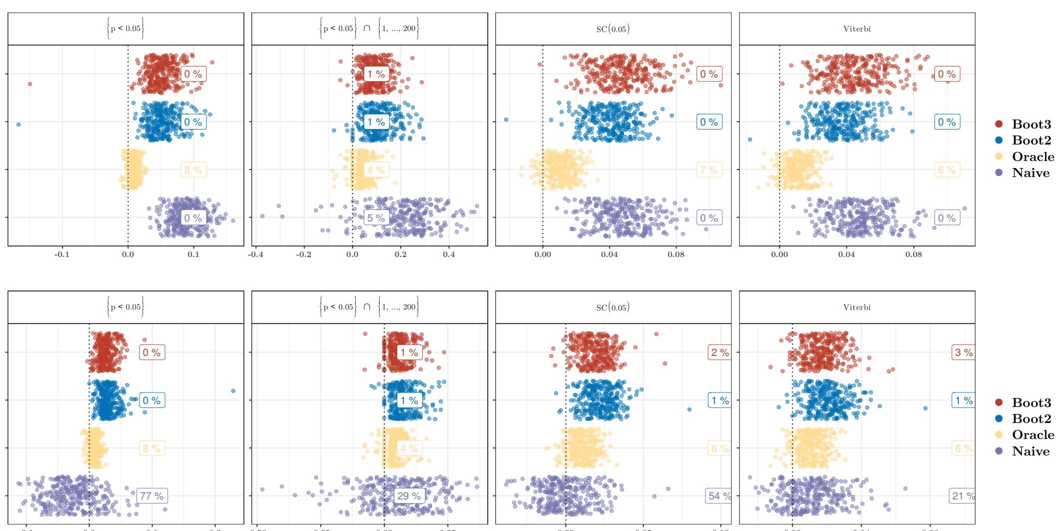

FIG S3. *Plot similar to Figure 3 (A) (same model and simulation parameters) when f_0 is unknown, that is, when the bounds also use a f_0 estimator, see Algorithm 4. Top panel: initialization with the true f_0; Bottom panel: initialization using local FDR algorithm (Efron, 2004).*

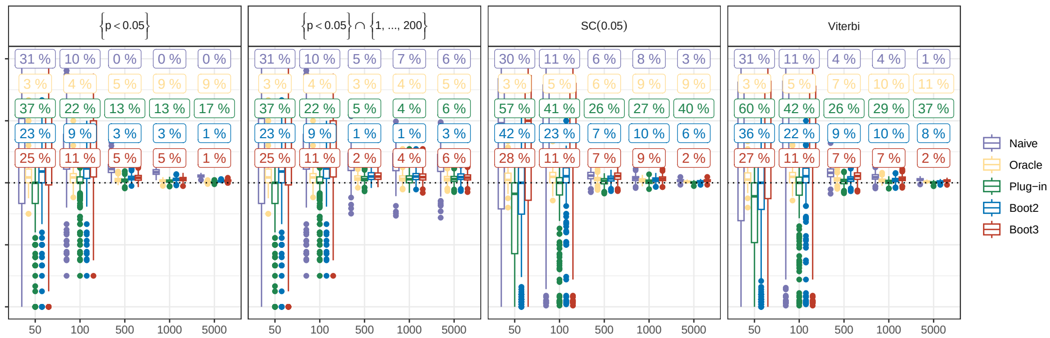

FIG S4. *Plot similar to Figure 3 (A) (same model and simulation parameters) for different values of m. The percentage display are the proportion of times when the bounds are lower than the true FDP value.*

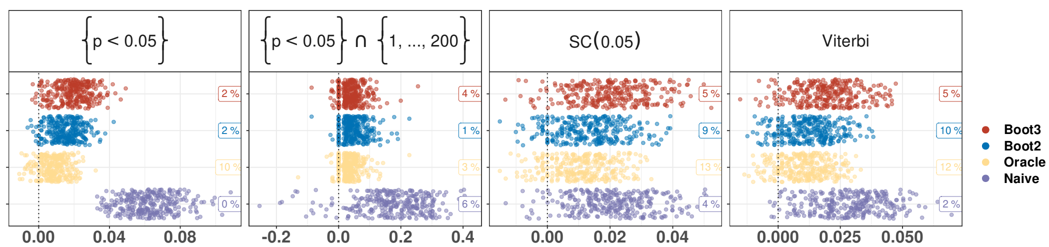

FIG S5. *Changer la legende!!!!!!Plot similar to Figure 3 (A) (same model and simulation parameters) for different values of m. The percentage display are the proportion of times when the bounds are lower than the true FDP value.*

S7.6. *Semi-simulated copy-number data.* Figure S6 illustrates the semi-simulated data set analyzed in Section 4.4, and Figure S7 shows the distribution of the corresponding test statistics across genomic regions.

S7.7. *Power in the semi-simulated case.* Figure S8 displays the power of the different bounds in the semi-simulated data set analyzed in Section 4.4.

S8. Application to influenza-like illness (ILI). We apply the proposed method to the weekly incidence rates of influenza-like illness (ILI). This data were collected from the Sentinelles Network, a national surveillance system in France. Sun and Cai (2009) studied this data set between January 1985 and February 2008, to be comparable we will restrict ourselves to this period and modeled these data by an HMM as in Section 2.1. They stated that the incidence rates can be classified into one of the two categories: aberration or usual.
The usual is the null states and the aberration is the alternatives one. The weekly ILI incidence rates are standardized according to the sizes of the underlying population and the representativeness of the participating physicians. Sun and Cai (2009) also applied a log transformation to reduce the skewness of the original data, hence so will we. In this particular example, the law under the null hypothesis (P_0) is unknown. Therefore, we estimate $\widehat{f}_0$ using Algorithm 4, with $\mathcal{N}(2.37, 0.76^2)$, the estimation of Sun and Cai (2009), as initial value. Accordingly, the p-values are replaced by the empirical p-values $\widehat{p}$ defined in Section S7.3.

To emphasize the advantages of bounds on the FDP compared to the FDR estimate, we displayed the 80% FDP post selection intervals on the different sets $S(X)$ described in the numerical experiments. For instance, for the selection $SC(0.05)$, it is interesting to know not

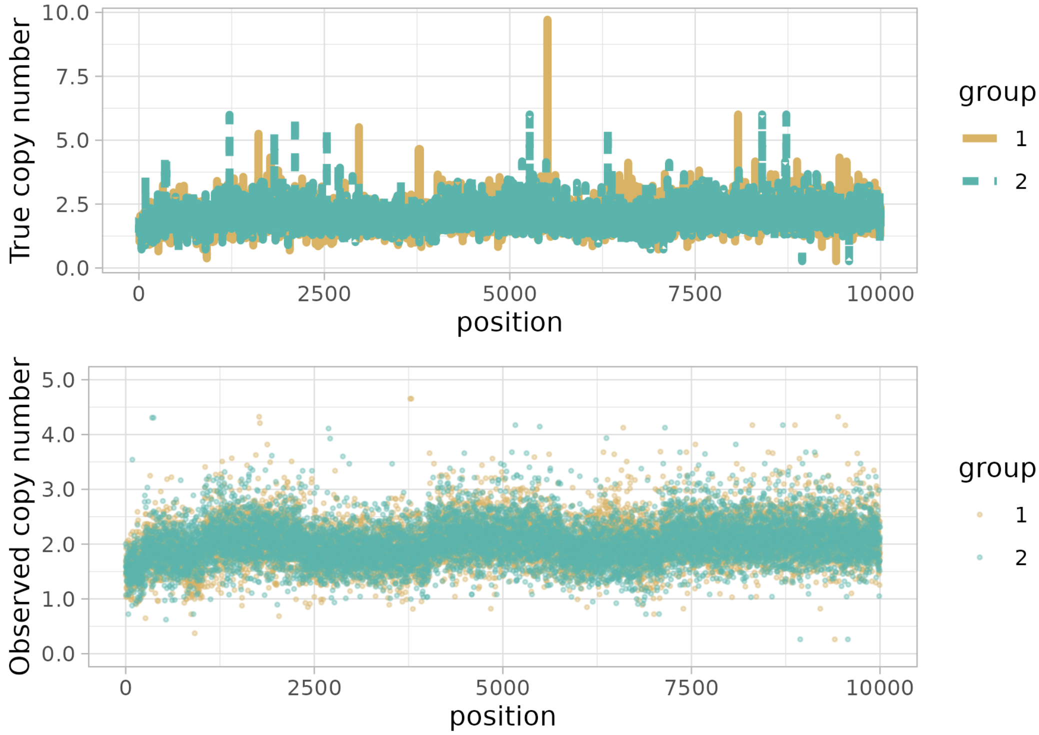

FIG S6. *Example of semi-simulated data used in Section 4.4. Top: true CN regions; bottom: CN signal for one sample for group 1 and one sample for group 2. The proportion of tumor cells was set to* 70%.

only that the estimated FDR is 5% but also that the 90% upper bound is smaller than 0.08. This application also underlines the interest of developing lower bounds, for instance in the first set ($\{\widehat{p} < 0.05\}$) we not only know, with high probability, that the FDP is smaller than 0.054 but also that it is higher than 0.028, which narrows down the probable true value of the FDP.

For completeness, we also added two selection policies :

- $SC(FDR_p)$ which uses Sun and Cai (2009) algorithm (with our parameter estimator) at a level corresponding to the estimated FDR of the set $\{\widehat{p} < 0.05\}$. Doing so, the selections $SC(FDR_p)$ and $\{\widehat{p} < 0.05\}$ have the same estimated FDR, see (10).
- $\{\widehat{p} < th\}$ which selects the $|SC(0.05)|$ smallest $\widehat{p}$-values. Doing so, $\{\widehat{p} < th\}$ and $SC(0.05)$ select the same number of null hypotheses.

We added the selection policies $SC(FDR_p)$ to compare two selection sets that have the same FDR estimate but not the same size. The selection $\{\widehat{p} < th\}$ has been added to compare two sets that have the same size but not the same estimated FDR. The result are displayed Figure S9. All the intervals are sharp around the FDP, which means that the FDP variance is estimated to be low. However, the post selection intervals of $SC(FDR_p)$ is wider than the one of $\{\widehat{p} < 0.05\}$. This information is not provided by FDR estimates. The set $\{\widehat{p} < th\}$ has the same size that $|SC(0.05)|$ but a larger FDR. It emphasizes that Sun and Cai (2009) is a better selection policy. Overall, this reinforces the interest in using a post selection interval rather than a point-wise FDR estimate.

REFERENCES

EFRON, B. (1979). Bootstrap Methods: Another Look at the Jackknife. *The Annals of Statistics* **7** 1 – 26.

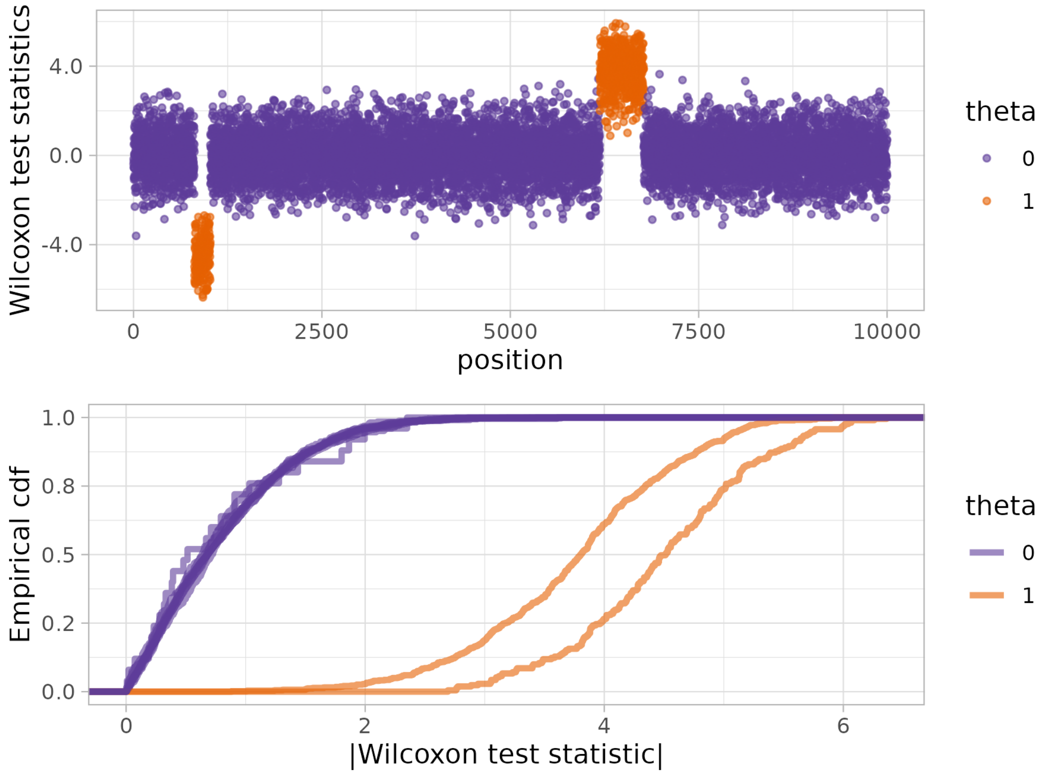

FIG S7. *Example of semi-simulated data used in Section 4.4 (continued). Top: Wilcoxon test statistics for the comparison of $n_1 = 50$ samples from group 1 and $n_2 = 50$ samples from group 2; bottom: empirical cumulative distribution function of the absolute value of the Wilcoxon test statistics for each of the 5 regions in the top plot. The proportion of tumor cells is set to* 70%.

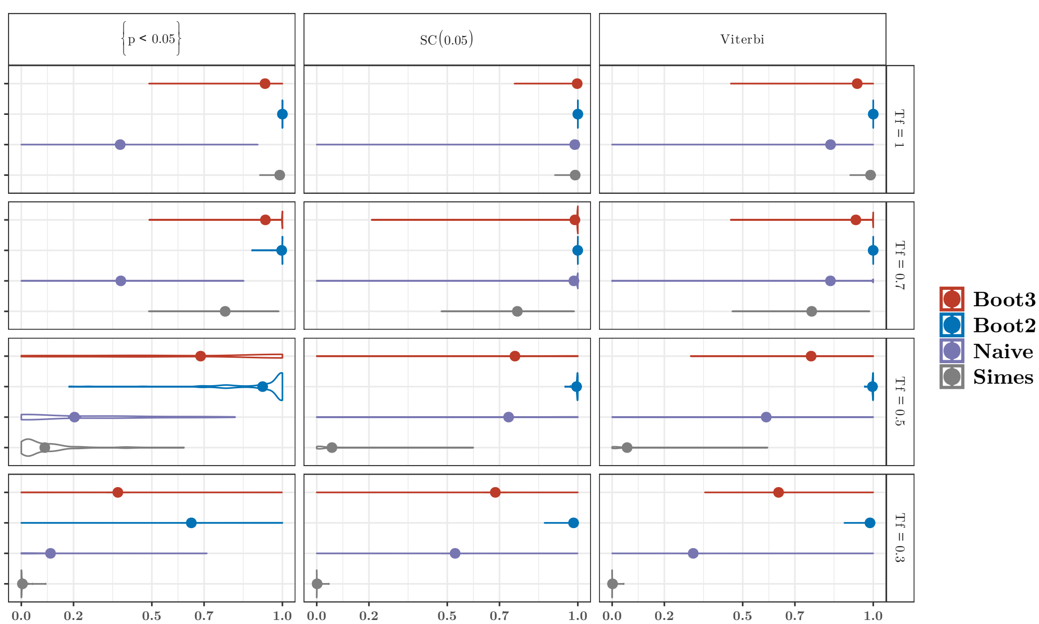

FIG S8. *Summary of the power for the semi-simulated data set analyzed in Section 4.4.*

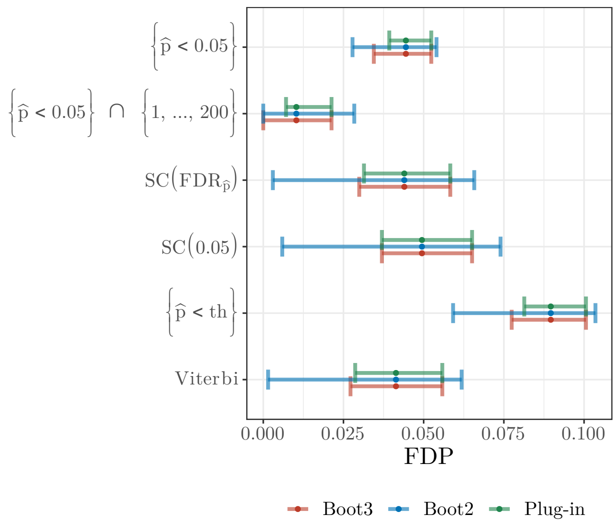

FIG S9. *80% FDP post selection interval intervals for different selection policies on the ILI incidence rates.*

EFRON, B. (2004). Large-scale simultaneous hypothesis testing: the choice of a null hypothesis. *J. Am. Stat. Assoc.* **99** 96–104.

STOREY, J. D. (2002). A direct approach to false discovery rates. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **64** 479–498.

SUN, W. and CAI, T. T. (2009). Large-scale multiple testing under dependence. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **71** 393–424.