AdaPT: Interactive Multiple Testing with Side Information

Lihua Lei & Will Fithian

Department of Statistics, UC Berkeley

Simons Workshop 2018
July 25, 2018

First Author

Lihua Lei (UC Berkeley Statistics)

Table of Contents

1. Background and Motivation
2. AdaPT: Adaptive p-Value Thresholding
3. Review of Existing Methods
4. Selecting the Threshold
5. Gene/Drug Response Data

Table of Contents

1. Background and Motivation
2. AdaPT: Adaptive p-Value Thresholding
3. Review of Existing Methods
4. Selecting the Threshold
5. Gene/Drug Response Data

Multiple Hypothesis Testing

Setting: hypotheses $H_1, \ldots, H_n$ with p-values $p_1, \ldots, p_n$

Notation:

- $\mathcal{H}_0 = \{i : H_i \text{ is true}\}$: null hypotheses
- $\mathcal{S} = \{i : H_i \text{ is rejected}\}$: set of rejections (discoveries)
- $R = |\mathcal{S}|$ total rejections (discoveries)
- $V = |\mathcal{S} \cap \mathcal{H}_0|$ incorrect rejections (false discoveries)

False Discovery Proportion $\mathrm{FDP} = \frac{V}{R \vee 1}$

Goal: control False Discovery Rate [Benjamini and Hochberg, 1995]

$$\mathrm{FDR} = \mathbb{E}[\mathrm{FDP}] \leq \alpha$$

Side Information

Observe side information $x_i \in \mathcal{X}$ for each H_i

[Ferkingstad et al., 2008, Ignatiadis et al., 2016, Li and Barber, 2016b]

$x_1, \ldots, x_n$ treated as fixed

Ordered multiple testing

[Foster and Stine, 2008, G'Sell et al., 2015, Li and Barber, 2016a]

- H_1 most "promising," then $H_2, \ldots, H_n$ $(x_i = i)$
- Focus power on early hypotheses

Examples:

- Sample variance / total count
- Data for gene i from one or more similar experiments
- Spatiotemporal location / location on a graph
- "Collaborative filtering" e.g. H_{ij} : drug i kills cancer j

Idea: if we learn a region of $\mathcal{X}$ has many non-nulls, can relax multiplicity correction in that region

Motivating Example: Gene/Drug Response Data

Li and Barber [2016a] proposed ingenious ordered analysis of gene expression data [Coser et al., 2003, Davis and Meltzer, 2007]

Expression in breast cancer cells in response to estrogen

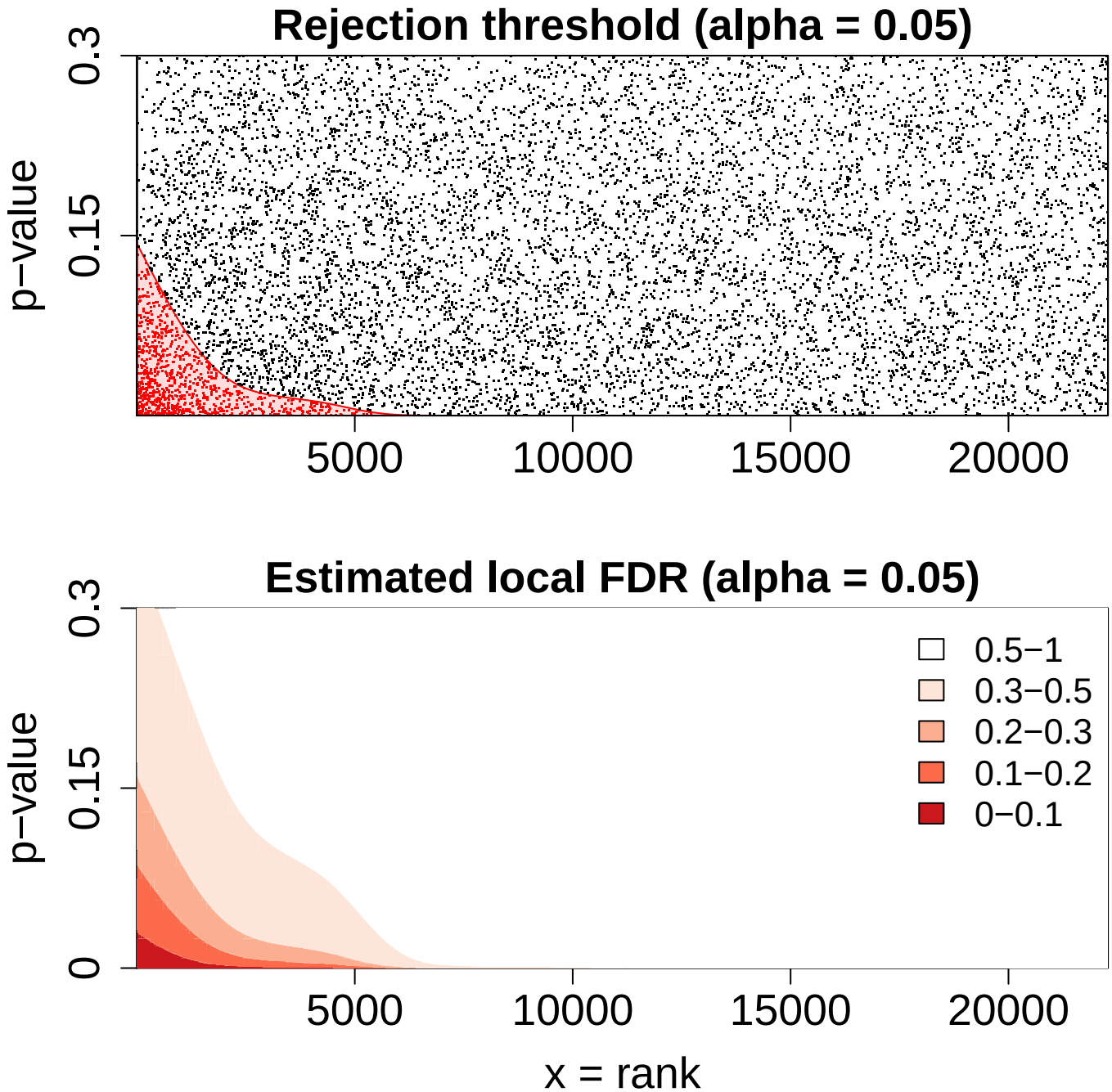
- $n = 22283$ genes, 25 trials at 5 doses incl. control
- H_i : no differential response in low-dose vs. control
- p_i computed via two-sample permutation t-test

Ordered by $\tilde{p}_i$, permutation p-value comparing high-dose vs. pooled sample of low-dose + control

Can show p_i independent of $\tilde{p}_i$ if H_i true

$x_i = \mathsf{Rank}(\tilde{p}_i) \in \{1, \ldots, n\}$

Gene/Drug response Data



Gene/Drug Response Data

Empirical Bayes: estimate $\text{fdr}(p \mid x) = \mathbb{P}(H_i \text{ false} | x_i = x, p_i = p)$

Reject p-values with small $\widehat{\text{fdr}}(p_i \mid x_i)$

Key questions:

- Which level curve?
- Can we control FDR despite "double-dipping"?
- Is there a "price" for adaptivity?
- What if eBayes model is wrong?

Gene/Drug response Data

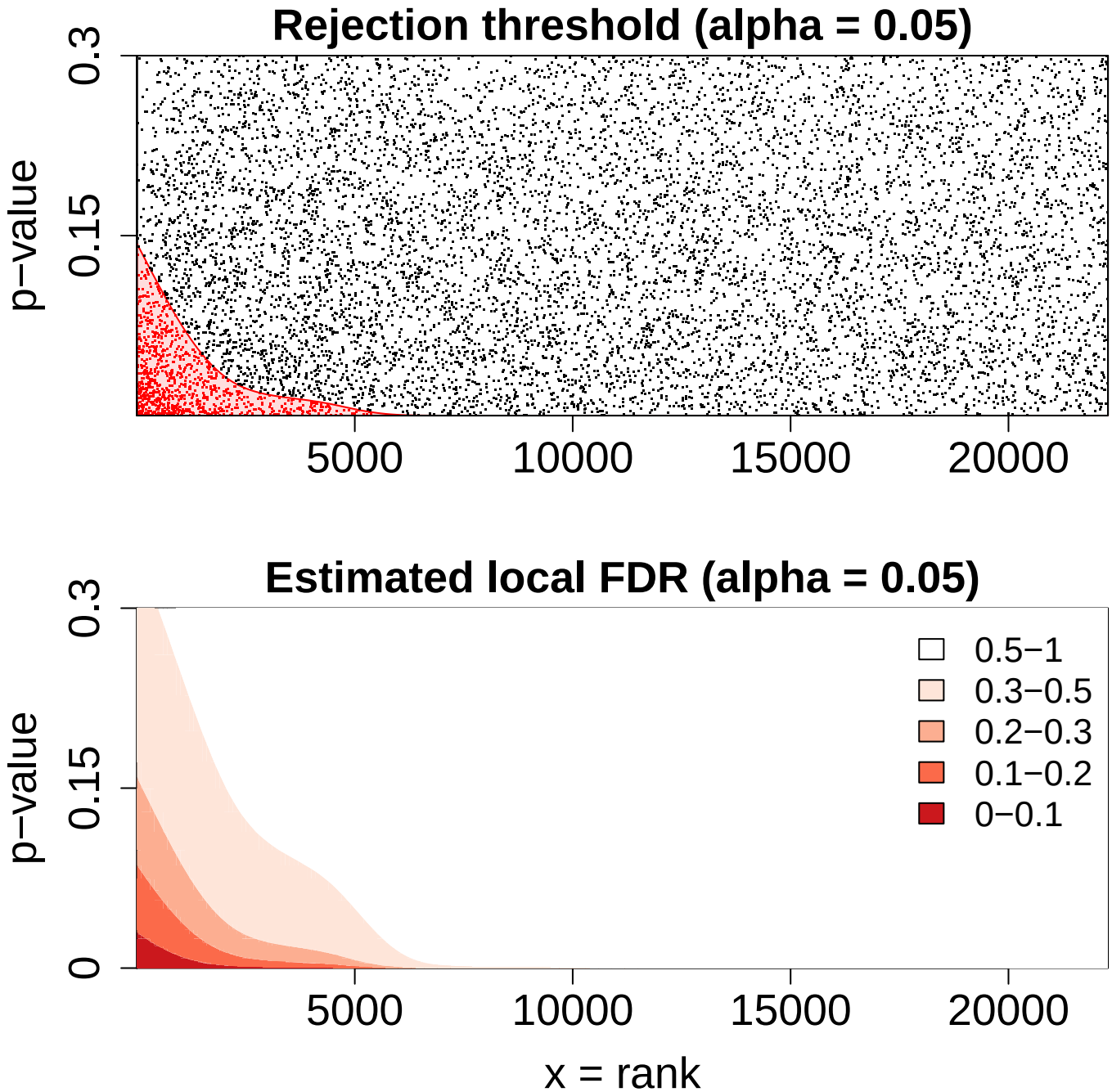


Table of Contents

1. Background and Motivation
2. AdaPT: Adaptive p-Value Thresholding
3. Review of Existing Methods
4. Selecting the Threshold
5. Gene/Drug Response Data

Adaptive p-Value Thresholding (AdaPT)

Iterative method: steps $t = 0, 1, 2, \ldots$

Step t: consider rejection threshold $s_t(x)$, compute:

$$R_t = |\{i : \ p_i \leq s_t(x_i)\}|$$
$$A_t = |\{i : \ p_i \geq 1 - s_t(x_i)\}|$$
$$\widehat{\mathrm{FDP}}_t = \frac{1 + A_t}{R_t \vee 1}$$

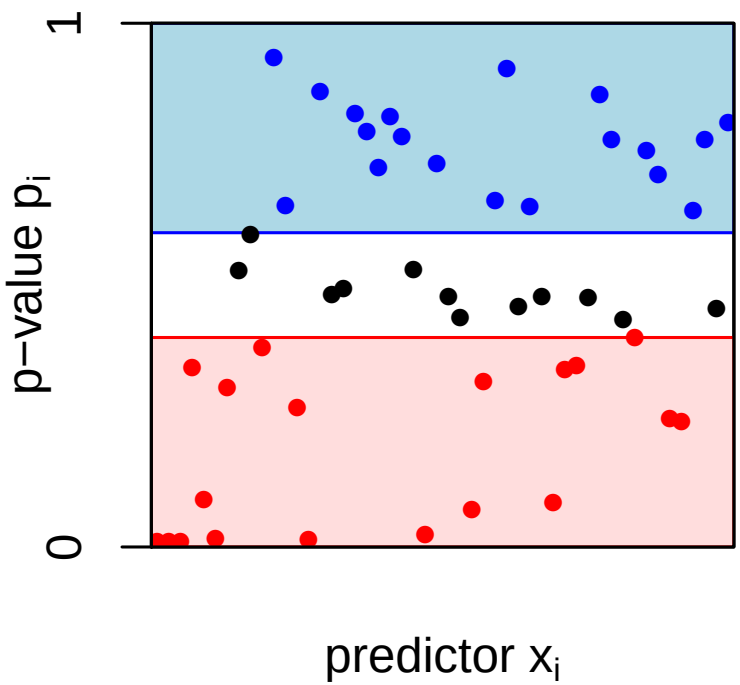
If $\widehat{\mathrm{FDP}}_t \leq \alpha$, stop and reject $\{H_i : \ p_i \leq s_t(x_i)\}$

Else, choose stricter threshold $s_{t+1}(x) \leq s_t(x)$ and continue

Update rule must be chosen based on partially masked data
(Keeps us from cheating!)

AdaPT, Visualized

$$\widehat{\mathrm{FDP}}_t = \frac{A_t + 1}{R_t \vee 1}$$



Covariate-dependent threshold $s_t(x)$

Mirror image $1 - s_t(x)$

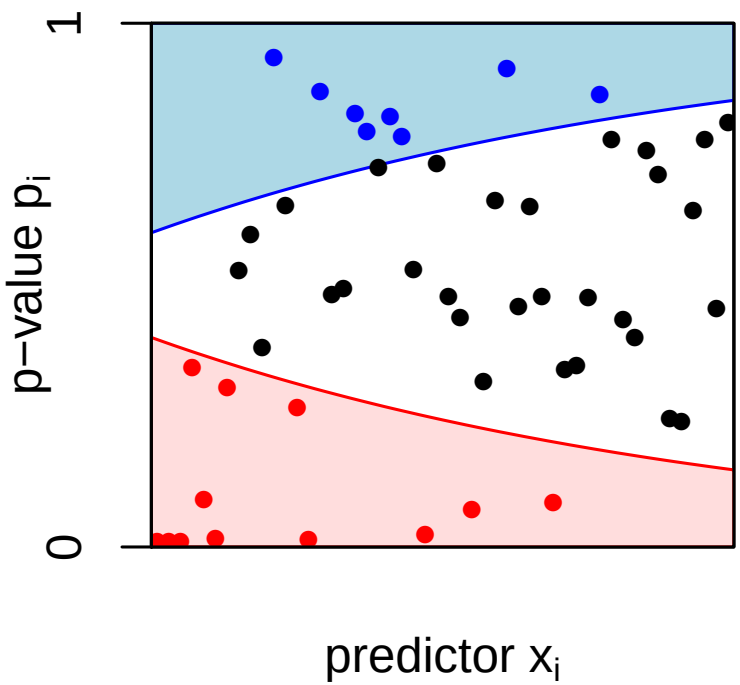
$R_t = \#$ red points

$A_t = \#$ blue points

$A_t \approx V_t = |\{i \in \mathcal{H}_0 : \ p_i \leq s_t(x_i)\}|$

AdaPT, Visualized

$$\widehat{\mathrm{FDP}}_t = \frac{A_t + 1}{R_t \vee 1}$$



Covariate-dependent threshold $s_t(x)$

Mirror image $1 - s_t(x)$

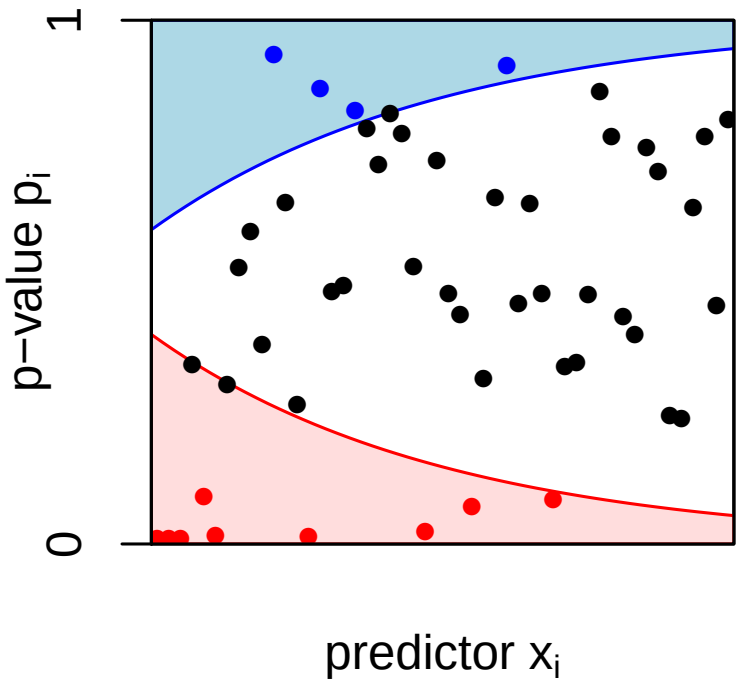
$R_t = \#$ red points

$A_t = \#$ blue points

$A_t \approx V_t = |\{i \in \mathcal{H}_0 : \ p_i \leq s_t(x_i)\}|$

AdaPT, Visualized

$$\widehat{\mathrm{FDP}}_t = \frac{A_t + 1}{R_t \vee 1}$$



Covariate-dependent threshold $s_t(x)$

Mirror image $1 - s_t(x)$

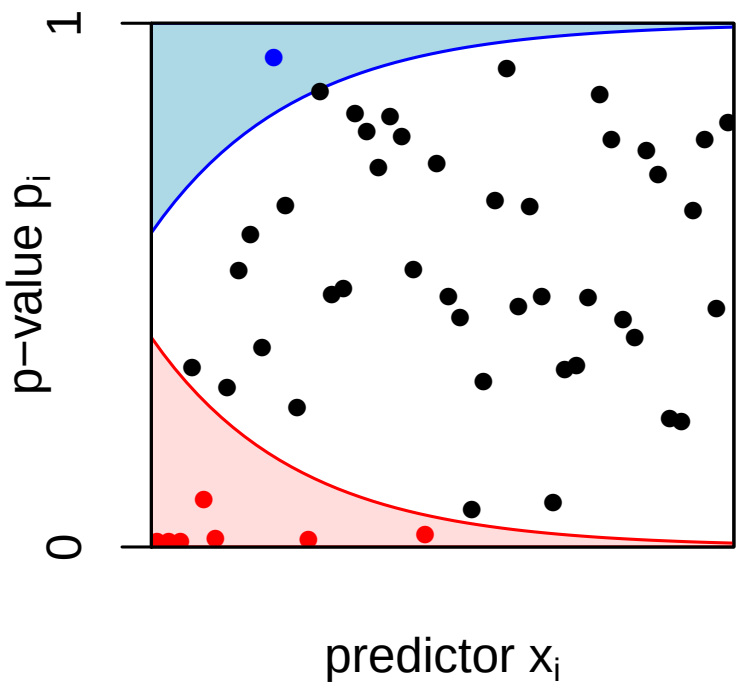
$R_t = \#$ red points

$A_t = \#$ blue points

$A_t \approx V_t = |\{i \in \mathcal{H}_0 : \ p_i \leq s_t(x_i)\}|$

AdaPT, Visualized

$$\widehat{\mathrm{FDP}}_t = \frac{A_t + 1}{R_t \vee 1}$$



Covariate-dependent threshold $s_t(x)$

Mirror image $1 - s_t(x)$

$R_t = \#$ red points

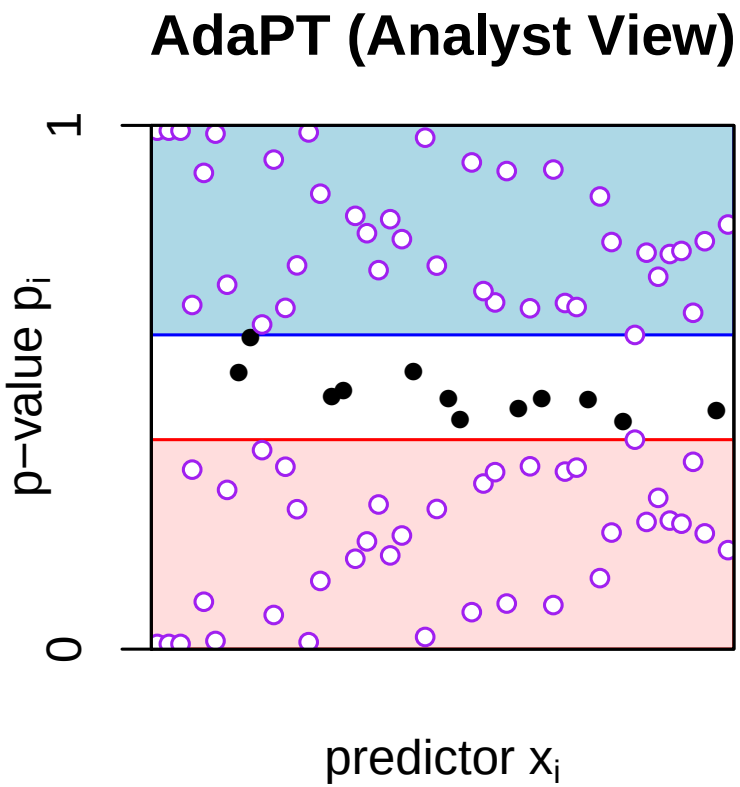
$A_t = \#$ blue points

$A_t \approx V_t = |\{i \in \mathcal{H}_0 : \ p_i \leq s_t(x_i)\}|$

AdaPT, "Analyst View"

Define partially masked p-values:

$$\tilde{p}_{t,i} = \begin{cases} p_i & s_t(x_i) < p_i < 1 - s_t(x_i) \\ \{p_i,\ 1 - p_i\} & \text{otherwise.} \end{cases}$$



To select $s_{t+1}(x)$, we can only use:

- $x_1, \ldots, x_n$
- $\tilde{p}_{t,1}, \ldots, \tilde{p}_{t,n}$
- A_t, R_t

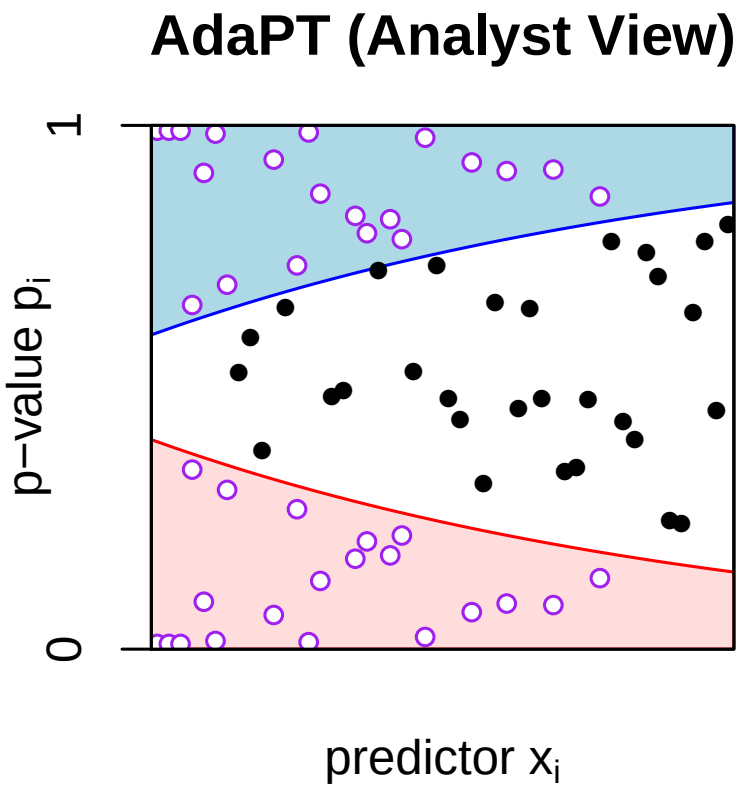
(and same for $t' < t$)

Any such update rule is OK

AdaPT, "Analyst View"

Define partially masked p-values:

$$\tilde{p}_{t,i} = \begin{cases} p_i & s_t(x_i) < p_i < 1 - s_t(x_i) \\ \{p_i,\ 1 - p_i\} & \text{otherwise.} \end{cases}$$



To select $s_{t+1}(x)$, we can only use:

- $x_1, \ldots, x_n$
- $\tilde{p}_{t,1}, \ldots, \tilde{p}_{t,n}$
- A_t, R_t

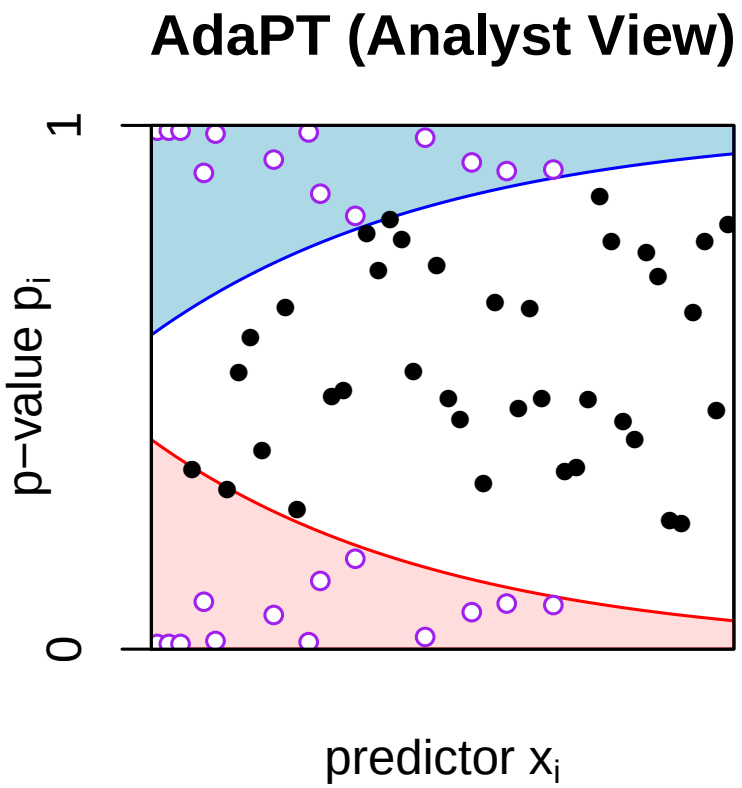
(and same for $t' < t$)

Any such update rule is OK

AdaPT, "Analyst View"

Define partially masked p-values:

$$\tilde{p}_{t,i} = \begin{cases} p_i & s_t(x_i) < p_i < 1 - s_t(x_i) \\ \{p_i,\ 1 - p_i\} & \text{otherwise.} \end{cases}$$



To select $s_{t+1}(x)$, we can only use:

- $x_1, \ldots, x_n$
- $\tilde{p}_{t,1}, \ldots, \tilde{p}_{t,n}$
- A_t, R_t

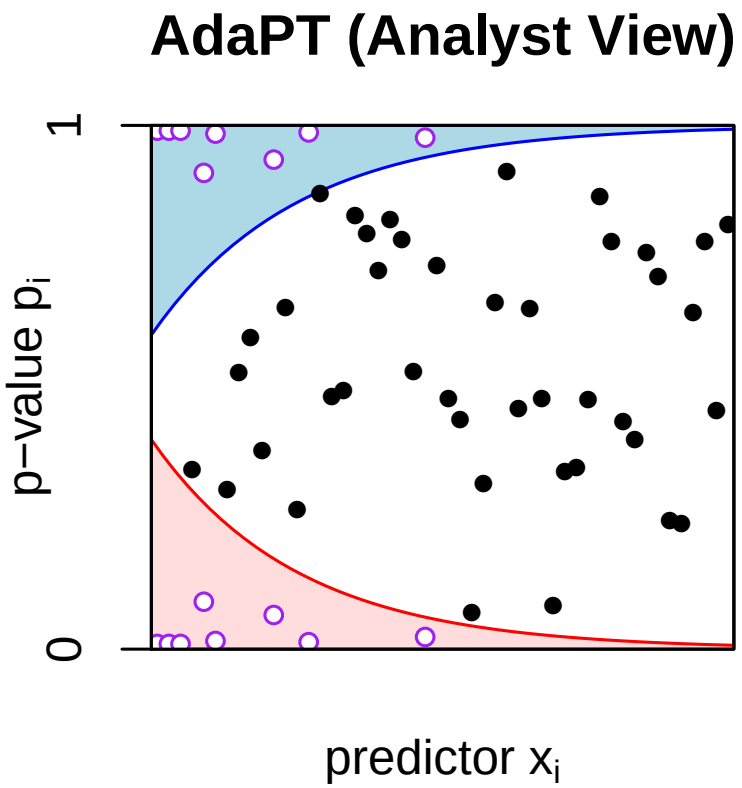
(and same for $t' < t$)

Any such update rule is OK

AdaPT, "Analyst View"

Define partially masked p-values:

$$\tilde{p}_{t,i} = \begin{cases} p_i & s_t(x_i) < p_i < 1 - s_t(x_i) \\ \{p_i,\ 1 - p_i\} & \text{otherwise.} \end{cases}$$



To select $s_{t+1}(x)$, we can only use:

- $x_1, \ldots, x_n$
- $\tilde{p}_{t,1}, \ldots, \tilde{p}_{t,n}$
- A_t, R_t

(and same for $t' < t$)

Any such update rule is OK

Filtrations

Key questions of selective inference

What did the analyst know, and when did she know it?

σ-field formalizes "what we know" at time t:

$$\mathcal{F}_t = \sigma((x_i)_{i=1}^n, (\tilde{p}_{t,i})_{i=1}^n, A_t, R_t)$$
$$\triangleq \{S \subseteq (\mathcal{X} \times [0,1])^n : \text{ analyst knows if } (x_i, p_i) \in S \text{ at time } t\}$$

Filtration formalizes increasing knowledge over time

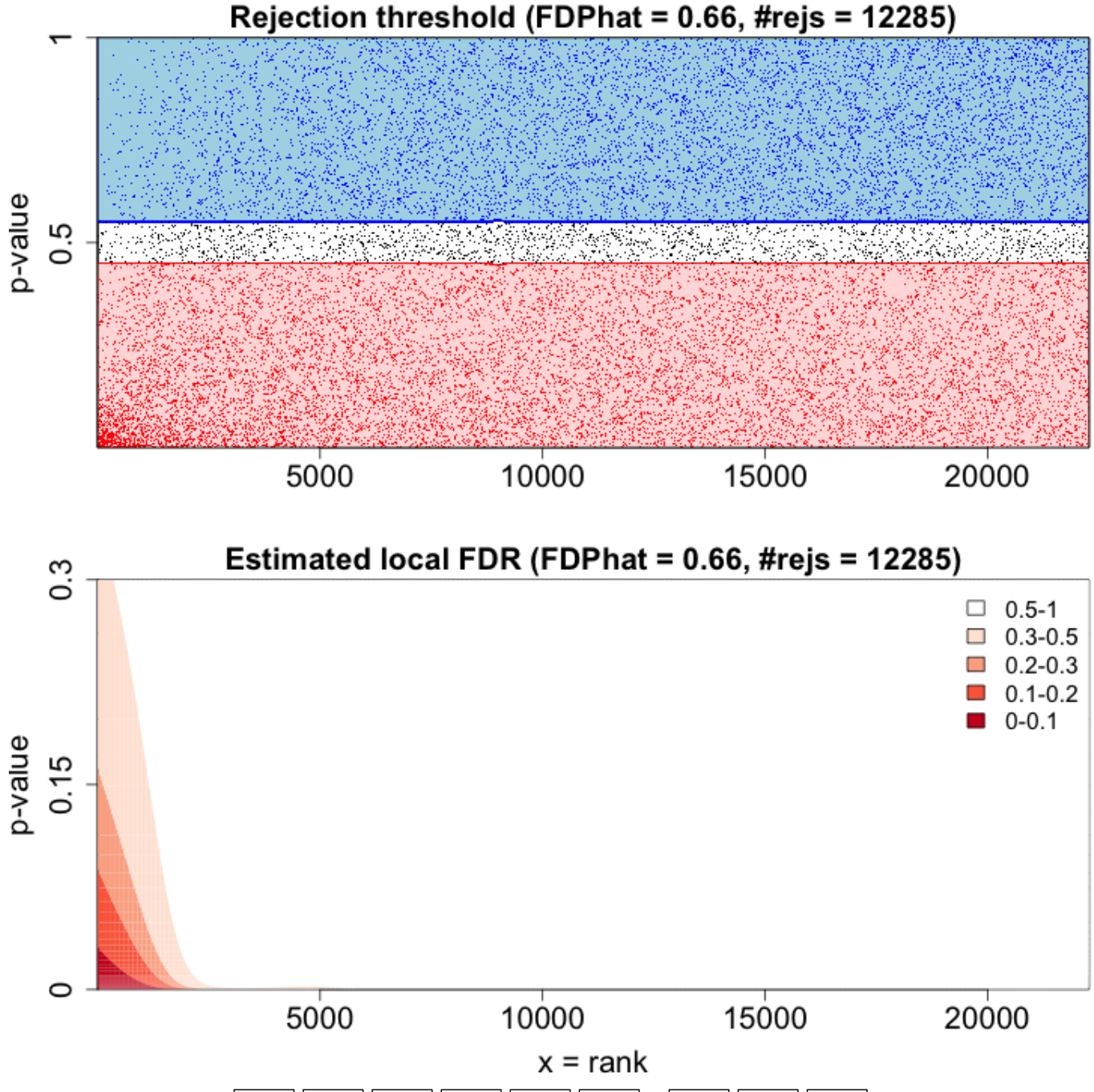
$$\mathcal{F}_0 \subseteq \mathcal{F}_1 \subseteq \cdots$$

Stopping time: $\{\hat{t} \leq t\} \in \mathcal{F}_t$

Supermartingale: M_t known by $\mathcal{F}_t$, $\mathbb{E}[M_{t+1} \mid \mathcal{F}_t] \leq M_t$ a.s.

Optional stopping theorem: $\mathbb{E}[M_{\hat{t}}] \leq M_0$ a.s.

Animation: Gene/Drug Response Data



AdaPT: Finite-Sample FDR Control

Theorem 1 (Lei and F, 2016).

Assume that, conditional on $(x_i)_{i=1}^n$ and $(p_i)_{i \notin \mathcal{H}_0}$, the null p-values $(p_i)_{i \in \mathcal{H}_0}$ are independent and mirror-conservative. Then AdaPT controls FDR at level α.

Finite-sample, no assumptions on update rule

Mirror-conservative: $f(p) \leq f(1-p), \ \forall\, p \leq 0.5$. Includes:

- Uniform
- Discrete p-values after randomization
- Permutation test p-values
- One-sided tests for
  - MLR families (e.g. log-concave location, exponential family)
  - Symmetric unimodal location families

Proof Sketch (* uniform nulls case)

Define

$$V_t = |\{i \in \mathcal{H}_0 : \ p_i \leq s_t(x_i)\}|$$
$$U_t = |\{i \in \mathcal{H}_0 : \ p_i \geq 1 - s_t(x_i)\}| \leq A_t$$
$$\mathcal{F}_t^+ = \sigma((p_i)_{i \in \mathcal{H}_0^C}, \mathcal{F}_t)$$

Ingredients:

1. $M_t = V_t/(1 + U_t)$ is $\mathcal{F}_t^+$-supermartingale (*)
2. $\hat{t}$ is $\mathcal{F}_t^+$-stopping time
3. $\mathbb{E}M_0 \leq 1$, $\widehat{\mathrm{FDP}}_t \leq \alpha$
4. $M_t \cdot \widehat{\mathrm{FDP}}_t = \frac{1+A_t}{1+U_t} \cdot \mathrm{FDP}_t \geq \mathrm{FDP}_t$

Then:

$$\mathrm{FDR} \ \leq \ \mathbb{E}\left[M_{\hat{t}} \cdot \widehat{\mathrm{FDP}}_{\hat{t}}\right] \ \leq \ \alpha \mathbb{E} M_{\hat{t}} \ \leq \ \alpha$$

Similar argument in Storey et al. [2004], Barber and Candès [2015b]

Prior Work on FDR Control with Side Information

Methods using generic x_i to learn data-adaptive weights for weighted BH:

Independent Hypothesis Weighting (IHW): [Ignatiadis et al., 2016]

- Bin x_i, estimate optimal stepwise rejection thresholds
- Asymptotic FDR control
- Problem: how to choose bin width?

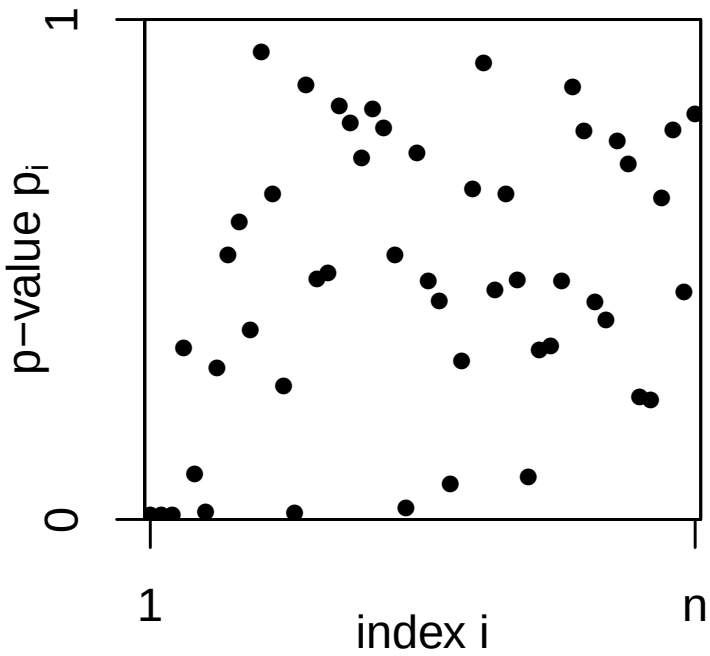
Structure-Adaptive BH Algorithm (SABHA): [Li and Barber, 2016b]

- Estimate $\pi_0(x)$ using truncated $p_i 1\{p_i > \tau\}$
- Can't reject $p_i > \tau$, can't learn from $p_i \leq \tau$
- Correct α for Rademacher complexity of $\hat{\pi}_0$
- Finite-sample control of FDR
- Problem: No way to estimate alternative distribution

Table of Contents

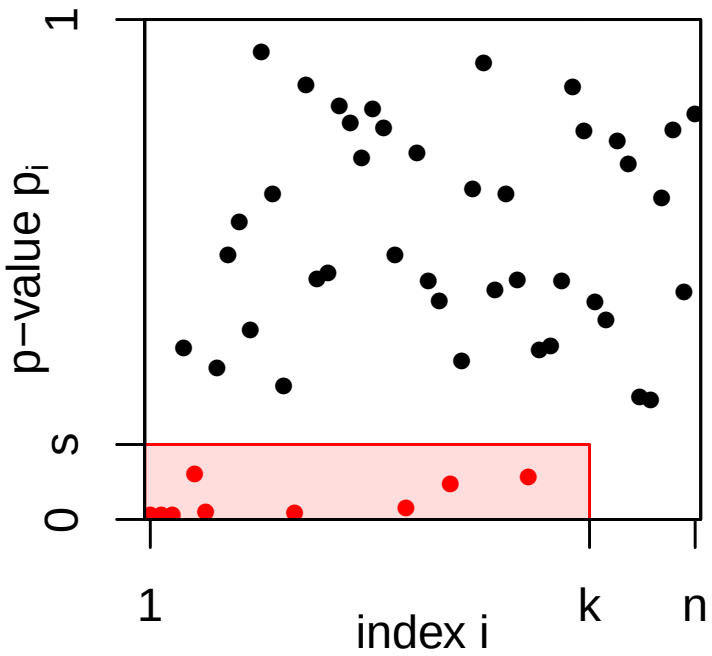
1. Background and Motivation
2. AdaPT: Adaptive p-Value Thresholding
3. Review of Existing Methods
4. Selecting the Threshold
5. Gene/Drug Response Data

Review of Existing Methods: General Recipe



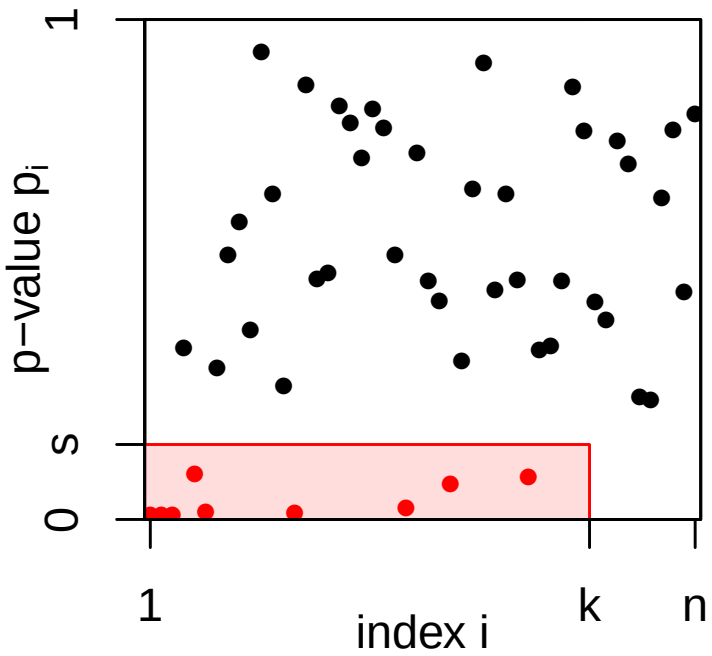
- p-values $p_1, \ldots, p_n$

Review of Existing Methods: General Recipe



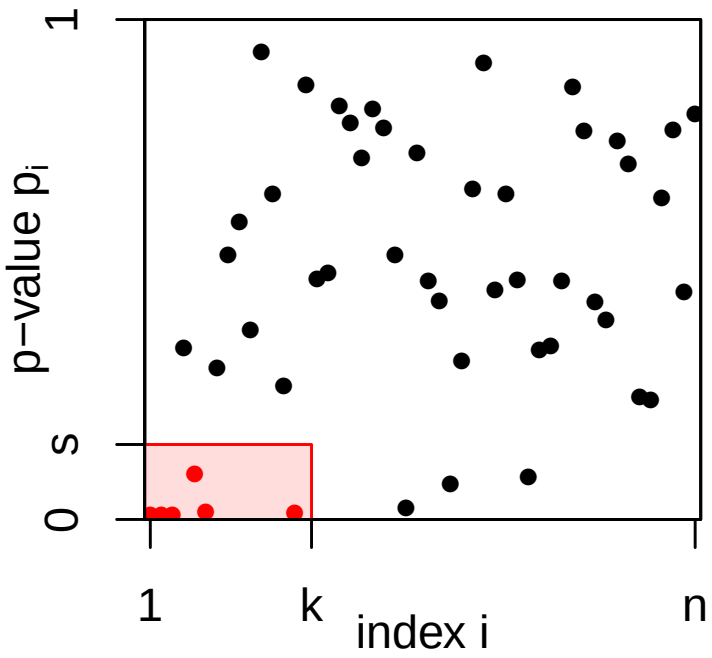
- p-values $p_1, \ldots, p_n$
- Rejection Set: a rectangular region indexed by $s \in [0, 1]$ and $k \in \{1, \ldots, n\}$

Review of Existing Methods: General Recipe



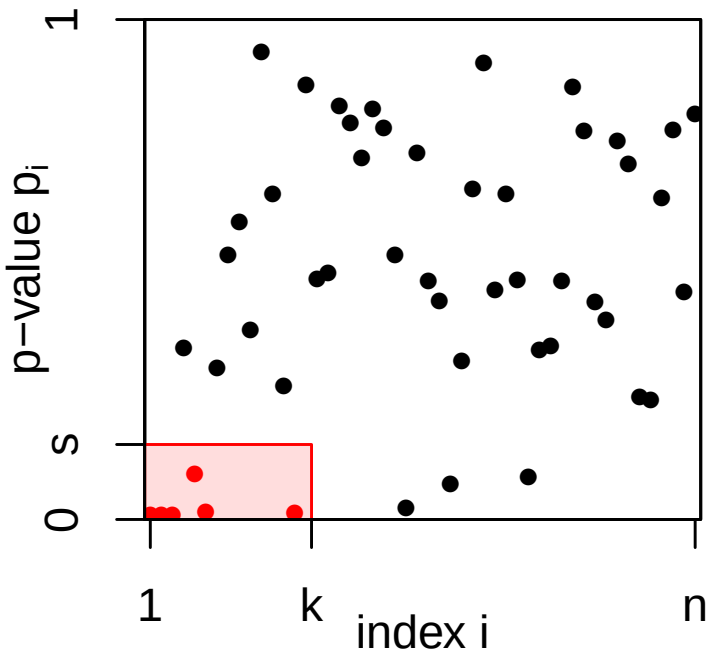
- p-values $p_1, \ldots, p_n$
- Rejection Set: a rectangular region indexed by $s \in [0, 1]$ and $k \in \{1, \ldots, n\}$
- Estimator $\widehat{\mathrm{FDP}}(k, s)$

Review of Existing Methods: General Recipe



- p-values $p_1, \ldots, p_n$
- Rejection Set: a rectangular region indexed by $s \in [0, 1]$ and $k \in \{1, \ldots, n\}$
- Estimator $\widehat{\mathrm{FDP}}(k, s)$
- Gradually reduce k or s until $\widehat{\mathrm{FDP}} \leq \alpha$

Review of Existing Methods: General Recipe

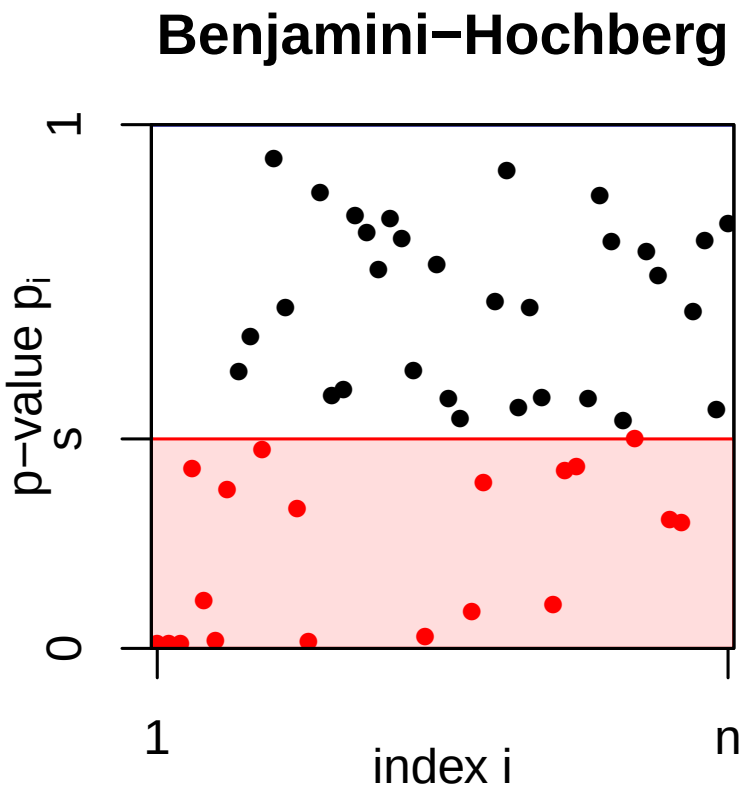


- p-values $p_1, \ldots, p_n$
- Rejection Set: a rectangular region indexed by $s \in [0, 1]$ and $k \in \{1, \ldots, n\}$
- Estimator $\widehat{\mathrm{FDP}}(k, s)$
- Gradually reduce k or s until $\widehat{\mathrm{FDP}} \leq \alpha$
- Reject $\{H_i : \ i \leq \hat{k}, p_i < \hat{s}\}$ (red points)

Methods differ on sequence of rectangles, formula for $\widehat{\mathrm{FDP}}$

Benjamini–Hochberg Procedure [Benjamini and Hochberg, 1995]

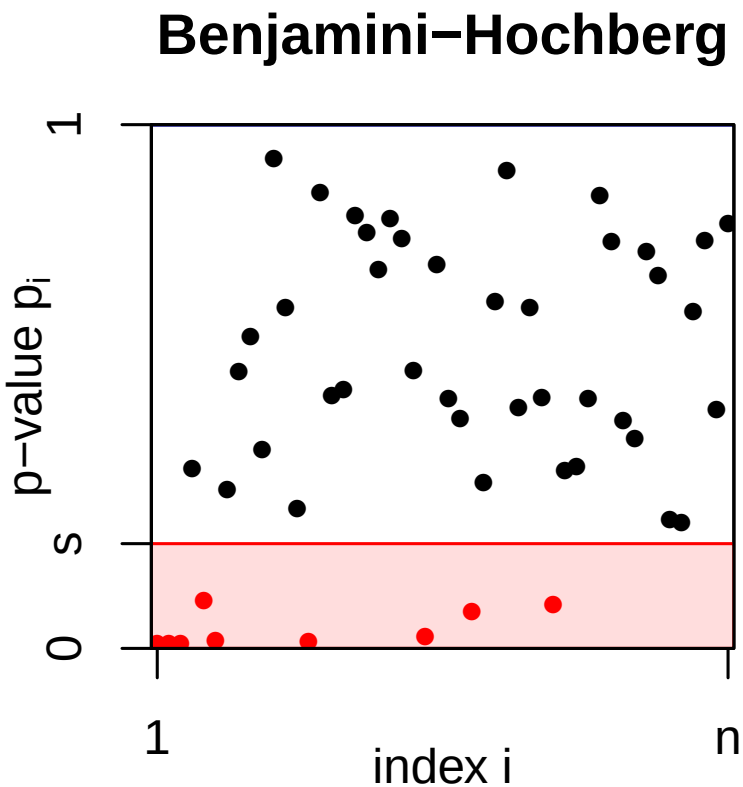
$$\widehat{\mathrm{FDP}}_{\mathrm{BH}} = \frac{ns}{R(s)}$$



- $R(s) = |\{i : p_i \leq s\}|$
- Optional stopping argument [Storey et al., 2004]
- Conservative: FDR $\leq \alpha\pi_0$ $(\pi_0 = |\mathcal{H}_0|/n)$

Benjamini–Hochberg Procedure [Benjamini and Hochberg, 1995]

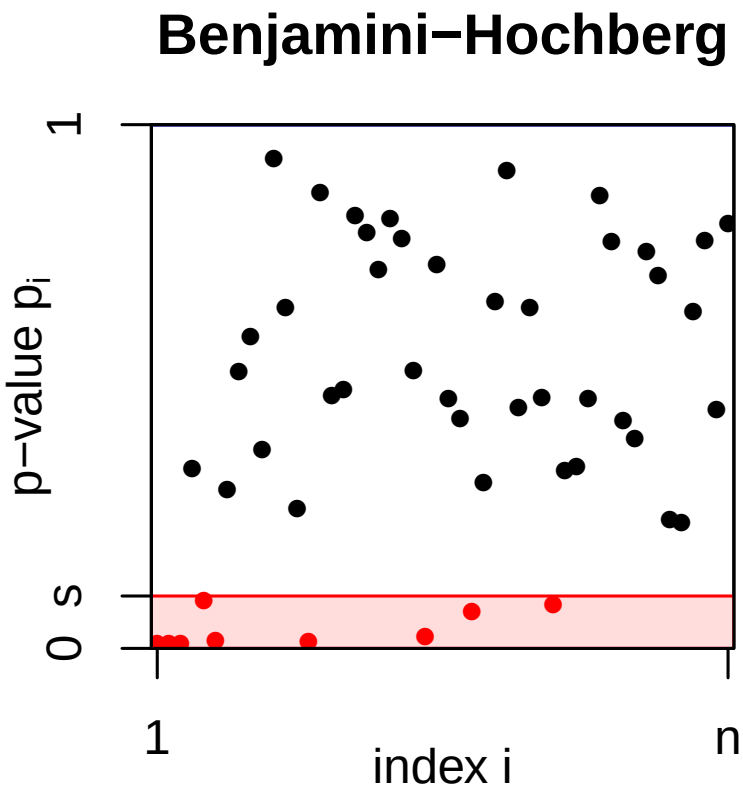
$$\widehat{\mathrm{FDP}}_{\mathsf{BH}} = \frac{ns}{R(s)}$$



- $R(s) = |\{i : p_i \leq s\}|$
- Optional stopping argument [Storey et al., 2004]
- Conservative: FDR $\leq \alpha\pi_0$ $(\pi_0 = |\mathcal{H}_0|/n)$

Benjamini–Hochberg Procedure [Benjamini and Hochberg, 1995]

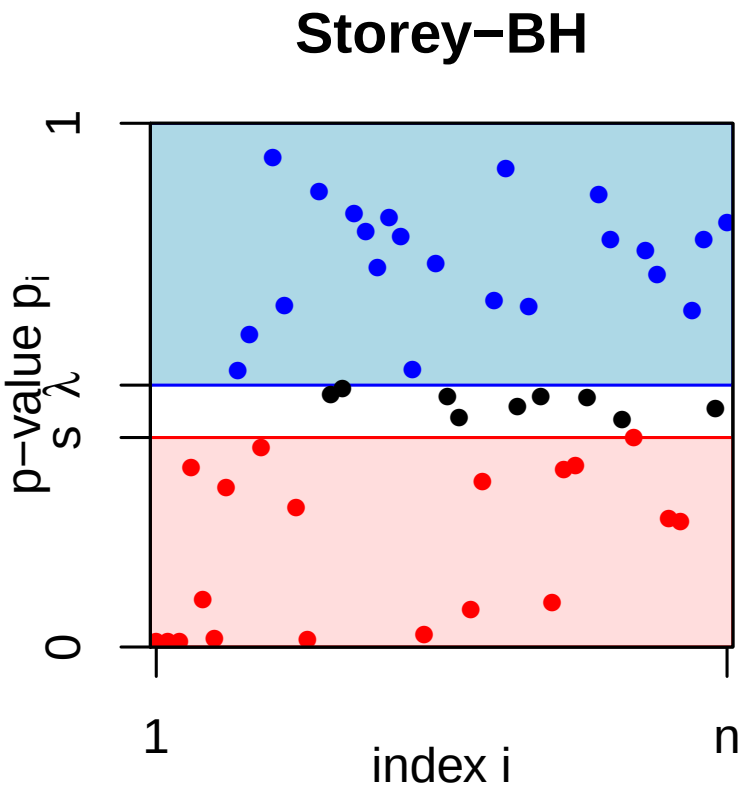
$$\widehat{\mathrm{FDP}}_{\mathrm{BH}} = \frac{ns}{R(s)}$$



- $R(s) = |\{i : p_i \leq s\}|$
- Optional stopping argument [Storey et al., 2004]
- Conservative: FDR $\leq \alpha\pi_0$ $(\pi_0 = |\mathcal{H}_0|/n)$

Storey–BH Procedure [Storey et al., 2004]

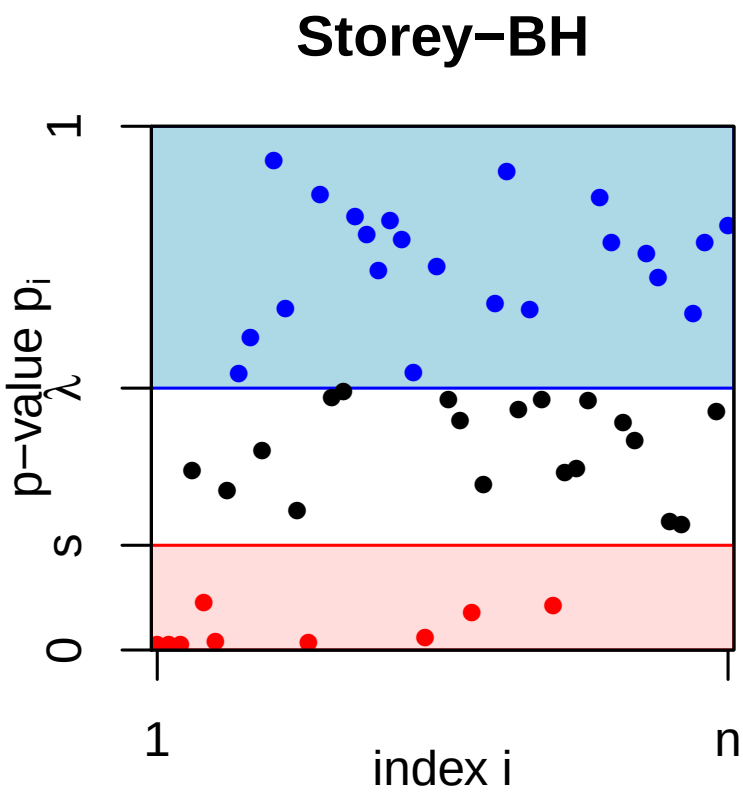
$$\widehat{\mathrm{FDP}}_{\mathsf{SBH}} = \frac{ns}{R(s)} \cdot \frac{A(\lambda)+1}{(1-\lambda)n}$$



- $R(s) = |\{i : p_i \leq s\}|$
- $A(\lambda) = |\{i : p_i \geq \lambda\}|$
- A estimates $|\mathcal{H}_0|/n$, FDR $\leq \alpha$

Storey–BH Procedure [Storey et al., 2004]

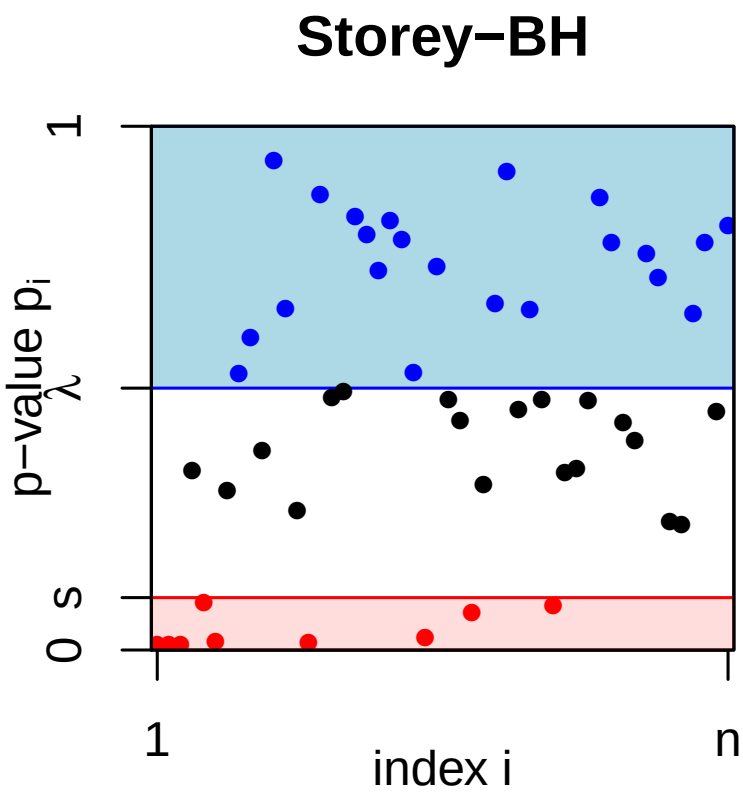
$$\widehat{\mathrm{FDP}}_{\mathsf{SBH}} = \frac{ns}{R(s)} \cdot \frac{A(\lambda)+1}{(1-\lambda)n}$$



- $R(s) = |\{i : p_i \leq s\}|$
- $A(\lambda) = |\{i : p_i \geq \lambda\}|$
- A estimates $|\mathcal{H}_0|/n$, FDR $\leq \alpha$

Storey–BH Procedure [Storey et al., 2004]

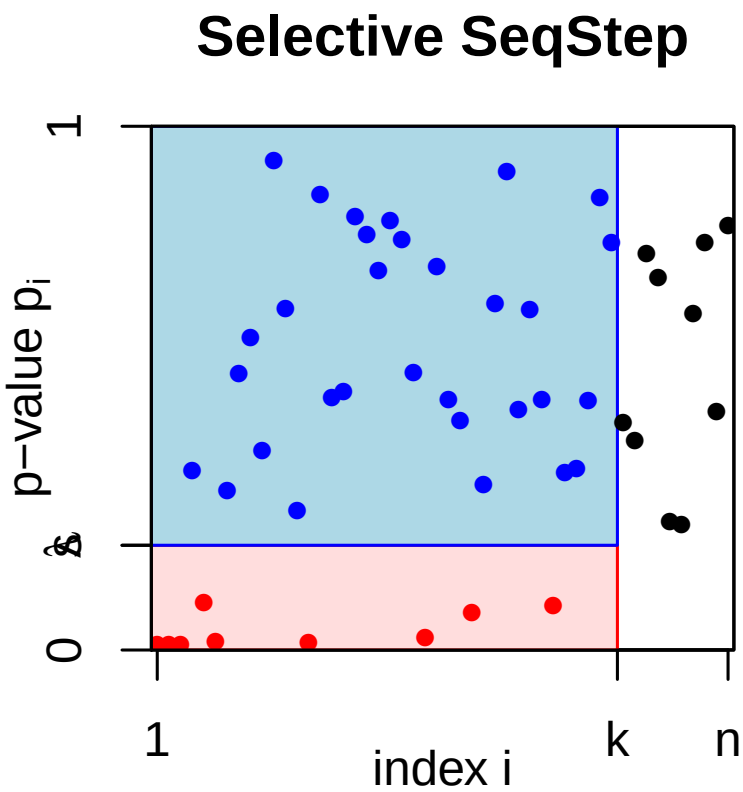
$$\widehat{\mathrm{FDP}}_{\mathsf{SBH}} = \frac{ns}{R(s)} \cdot \frac{A(\lambda)+1}{(1-\lambda)n}$$



- $R(s) = |\{i : p_i \leq s\}|$
- $A(\lambda) = |\{i : p_i \geq \lambda\}|$
- A estimates $|\mathcal{H}_0|/n$, FDR $\leq \alpha$

Selective SeqStep [Barber and Candès, 2015a]

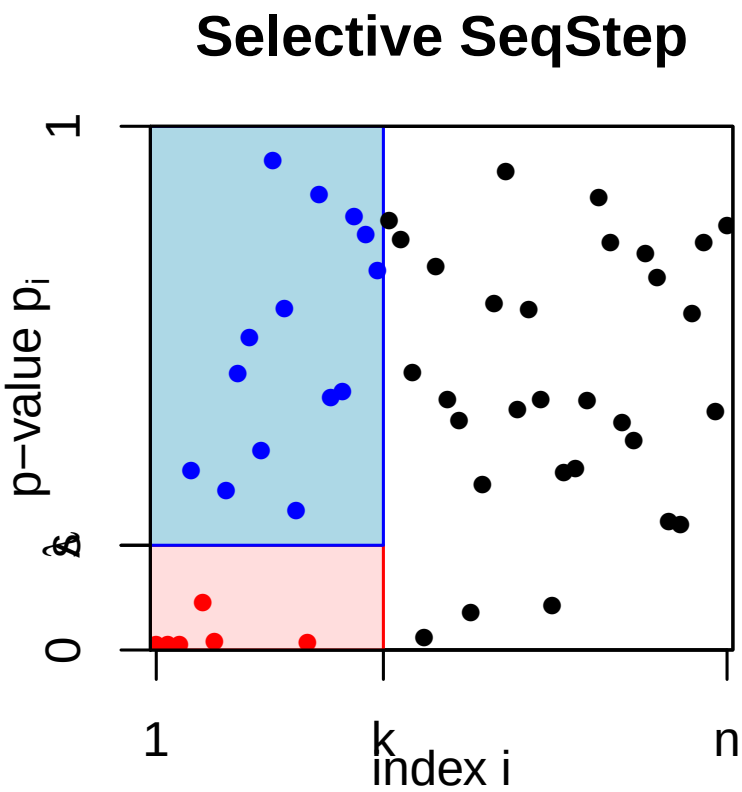
$$\widehat{\mathrm{FDP}}_{\mathrm{SS}} = \frac{ks}{R(k,s)} \cdot \frac{A(k,s)+1}{k(1-s)}$$



- $R(k,s) = |\{i \leq k : p_i \leq s\}|$
- $A(k,s) = |\{i \leq k : p_i \geq s\}|$

Selective SeqStep [Barber and Candès, 2015a]

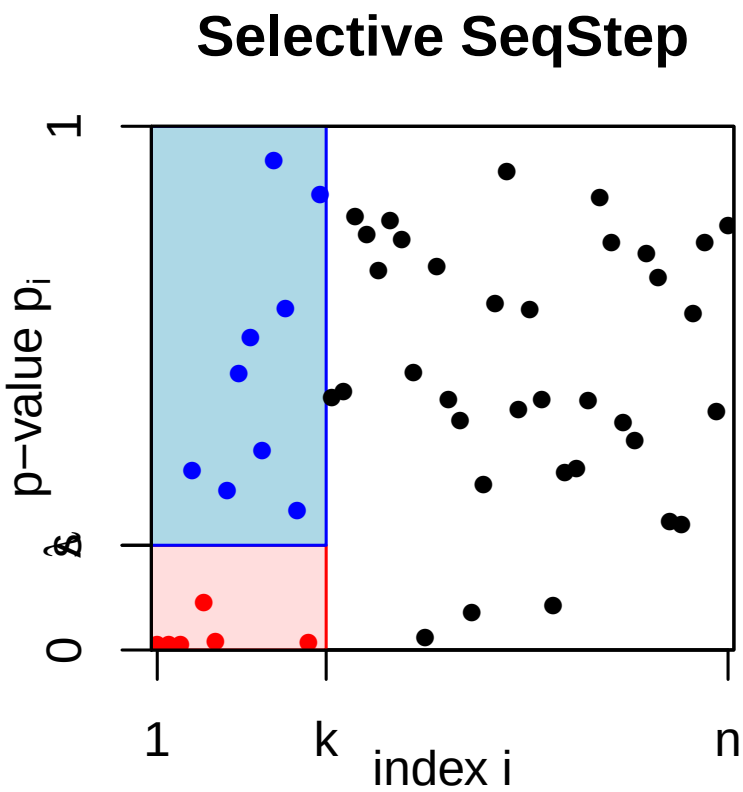
$$\widehat{\mathrm{FDP}}_{\mathrm{SS}} = \frac{ks}{R(k,s)} \cdot \frac{A(k,s)+1}{k(1-s)}$$



- $R(k,s) = |\{i \le k : p_i \le s\}|$
- $A(k,s) = |\{i \le k : p_i \ge s\}|$

Selective SeqStep [Barber and Candès, 2015a]

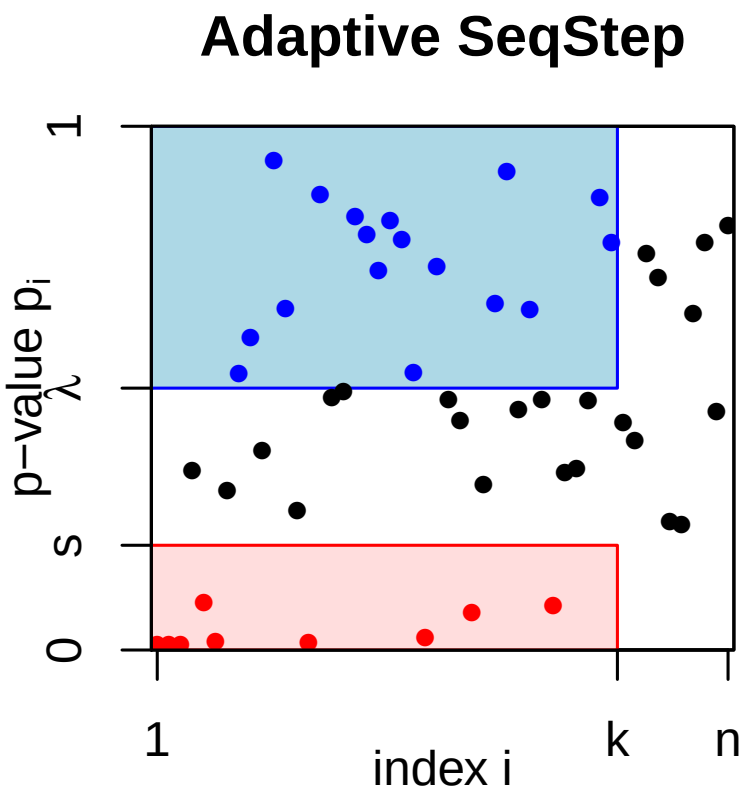
$$\widehat{\mathrm{FDP}}_{\mathrm{SS}} = \frac{ks}{R(k,s)} \cdot \frac{A(k,s)+1}{k(1-s)}$$



- $R(k,s) = |\{i \leq k : p_i \leq s\}|$
- $A(k,s) = |\{i \leq k : p_i \geq s\}|$

Adaptive SeqStep [Lei and F, 2016]

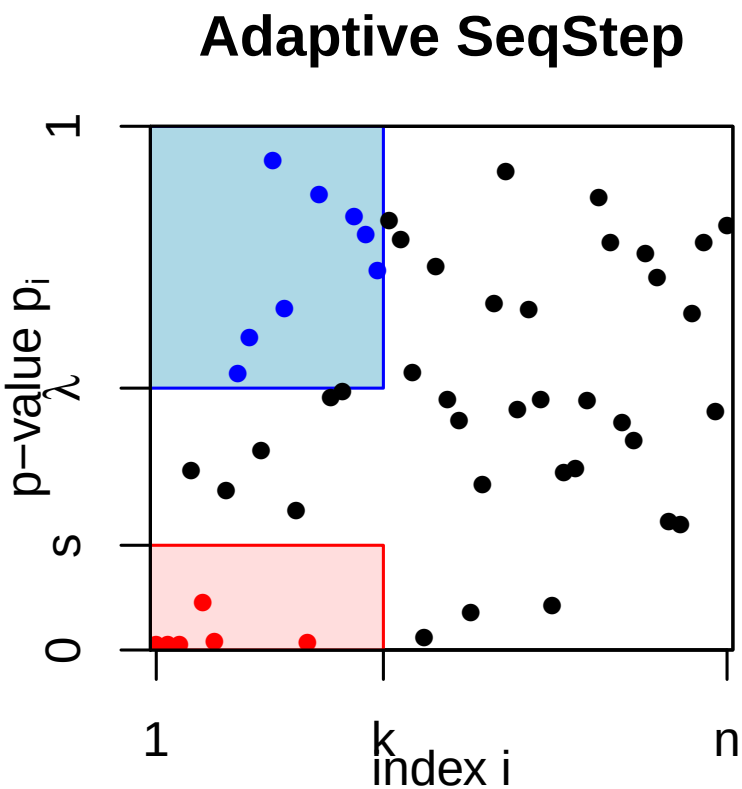
$$\widehat{\mathrm{FDP}}_{\mathrm{AS}} = \frac{ks}{R(k,s)} \cdot \frac{A(k,\lambda)+1}{k(1-\lambda)}$$



- $R(k,s) = |\{i \le k : p_i \le s\}|$
- $A(k,s) = |\{i \le k : p_i \ge s\}|$

Adaptive SeqStep [Lei and F, 2016]

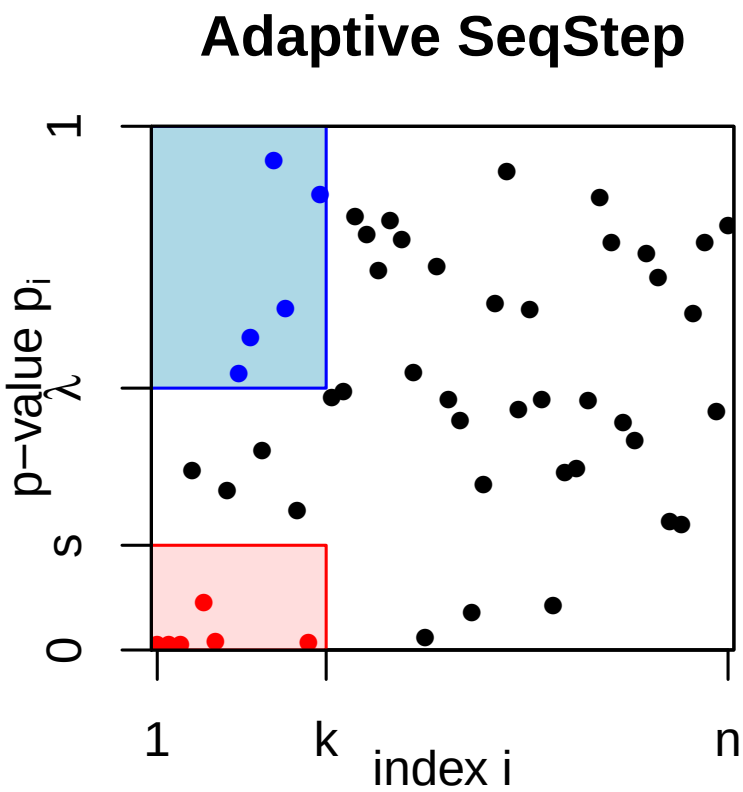
$$\widehat{\mathrm{FDP}}_{\mathsf{AS}} = \frac{ks}{R(k,s)} \cdot \frac{A(k,\lambda)+1}{k(1-\lambda)}$$



- $R(k,s) = |\{i \le k : p_i \le s\}|$
- $A(k,s) = |\{i \le k : p_i \ge s\}|$

Adaptive SeqStep [Lei and F, 2016]

$$\widehat{\mathrm{FDP}}_{\mathsf{AS}} = \frac{ks}{R(k,s)} \cdot \frac{A(k,\lambda)+1}{k(1-\lambda)}$$



- $R(k,s) = |\{i \le k : p_i \le s\}|$
- $A(k,s) = |\{i \le k : p_i \ge s\}|$

Table of Contents

1. Background and Motivation
2. AdaPT: Adaptive p-Value Thresholding
3. Review of Existing Methods
4. Selecting the Threshold
5. Gene/Drug Response Data

Conditional Two-Groups Model

Frame threshold choice in terms of conditional two-groups model:

$$H_i \mid x_i \sim \text{Bernoulli}(\pi_1(x_i))$$

$$p_i \mid H_i, x_i \sim \begin{cases} \text{Unif} & \text{if } H_i = 0 \\ f_1(p \mid x_i) & \text{if } H_i = 1 \end{cases}.$$

Mixture density $f(p \mid x) \to$ conditional local fdr [Efron et al., 2001]

$$\text{fdr}(p \mid x) \;=\; \mathbb{P}(H_i \text{ is null } \mid x_i = x, p_i = p) \;=\; \frac{1 - \pi_1(x)}{f(p \mid x)}$$

Estimate $\widehat{\text{fdr}}(p \mid x)$ using models for $\pi_1(x), f_1(\cdot \mid x)$

Impute masked p_i using (e.g.) EM

AdaPT–GLM

Simple choice of model, given featurization $\phi : \mathcal{X} \to \mathbb{R}^d$

Logistic regression for π_1, 1-parameter beta GLM for f_1

$$\log \frac{\mathbb{P}(H = 1 \mid x)}{\mathbb{P}(H = 0 \mid x)} = \beta' \phi(x)$$

$$f_1(p \mid x, H = 1) = \frac{1}{\gamma' \phi(x)} p^{\gamma' \phi(x) - 1} \quad (0 < p < 1)$$

For 1-dimensional examples, ϕ typically a natural cubic spline basis

Expectation–Maximization: alternate between

- Impute H_i and masked p_i (both binary, knowing $\mathcal{F}_t$)
- Estimate β, γ via `glm` or `glmnet` (regularized)

`adaptMT` R package: `adapt_glm`, `adapt_glmnet`

Updating the Threshold

Theorem 2 (Lei and F, 2016).

Under mild assumptions, the optimal threshold $s(x)$ is a level surface of local FDR.

Step 1. Use **your favorite method** to fit **your favorite model**

Step 2. Estimate level surfaces of local FDR

Step 3. Move the threshold to a smaller level surface, re-fit

Table of Contents

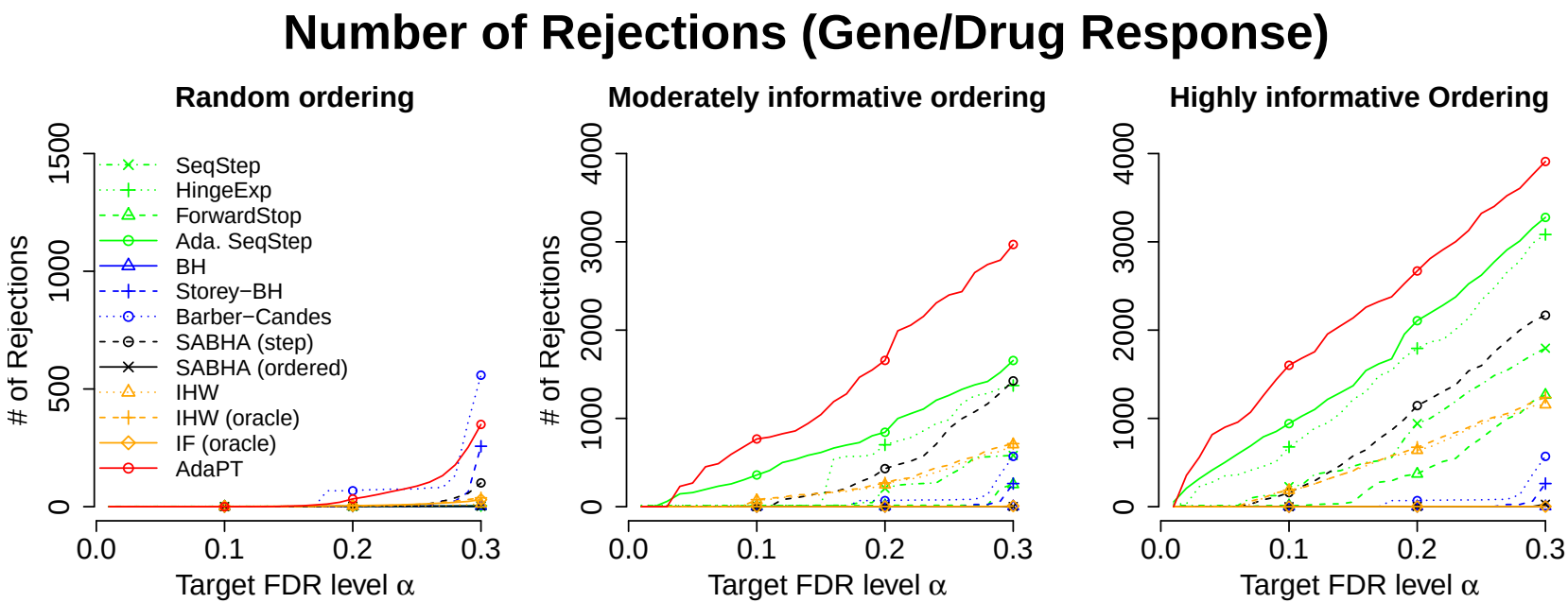
1. Background and Motivation
2. AdaPT: Adaptive p-Value Thresholding
3. Review of Existing Methods
4. Selecting the Threshold
5. Gene/Drug Response Data

Gene/Drug response Data: Power Comparison

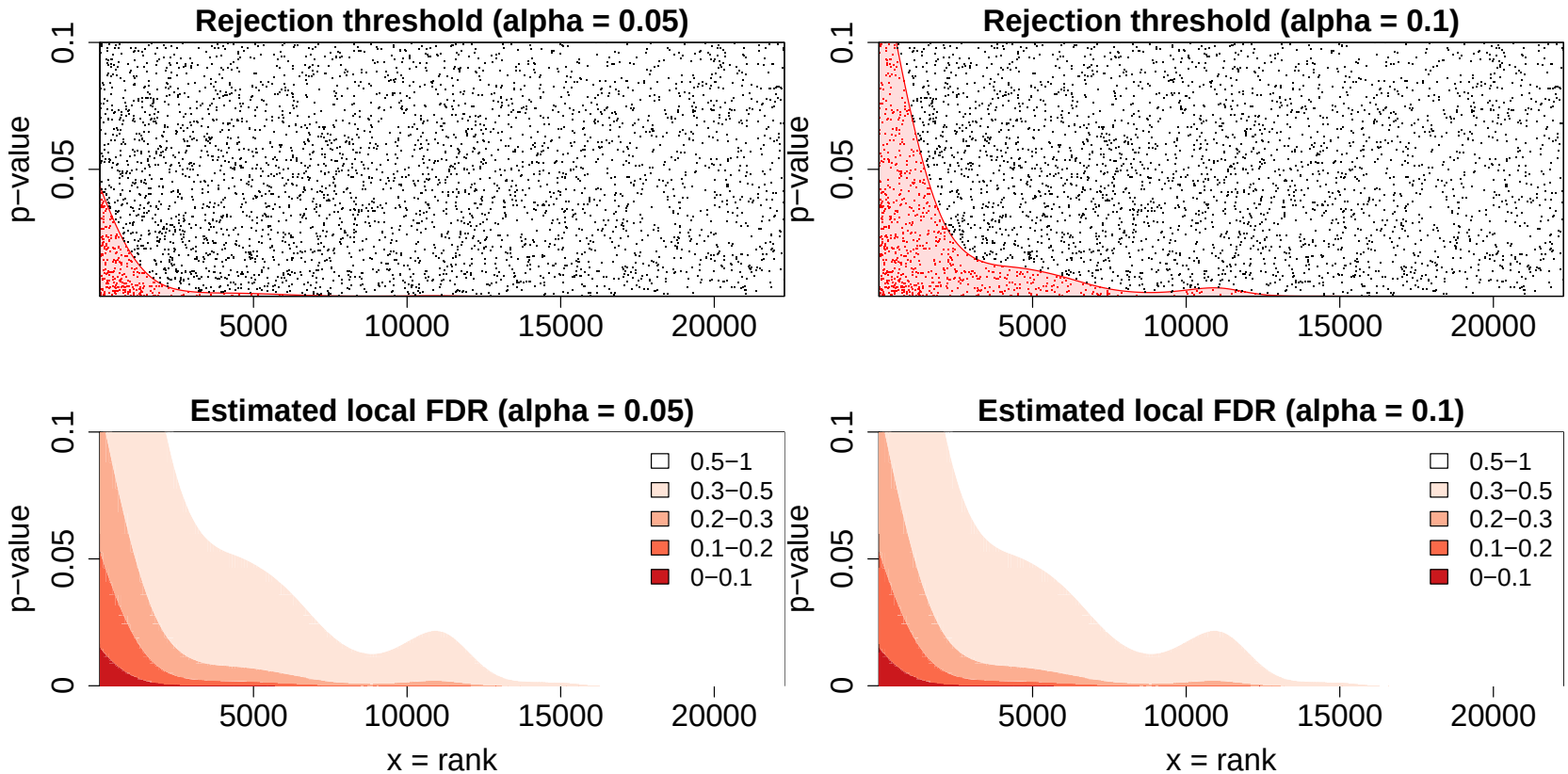
Compared AdaPT to competing methods using three orderings:

- Random ordering
- Moderate dose ordering (dose 2 vs. pooled doses 0 & 1)
- High dose ordering (dose 4 vs. pooled doses 0 & 1)

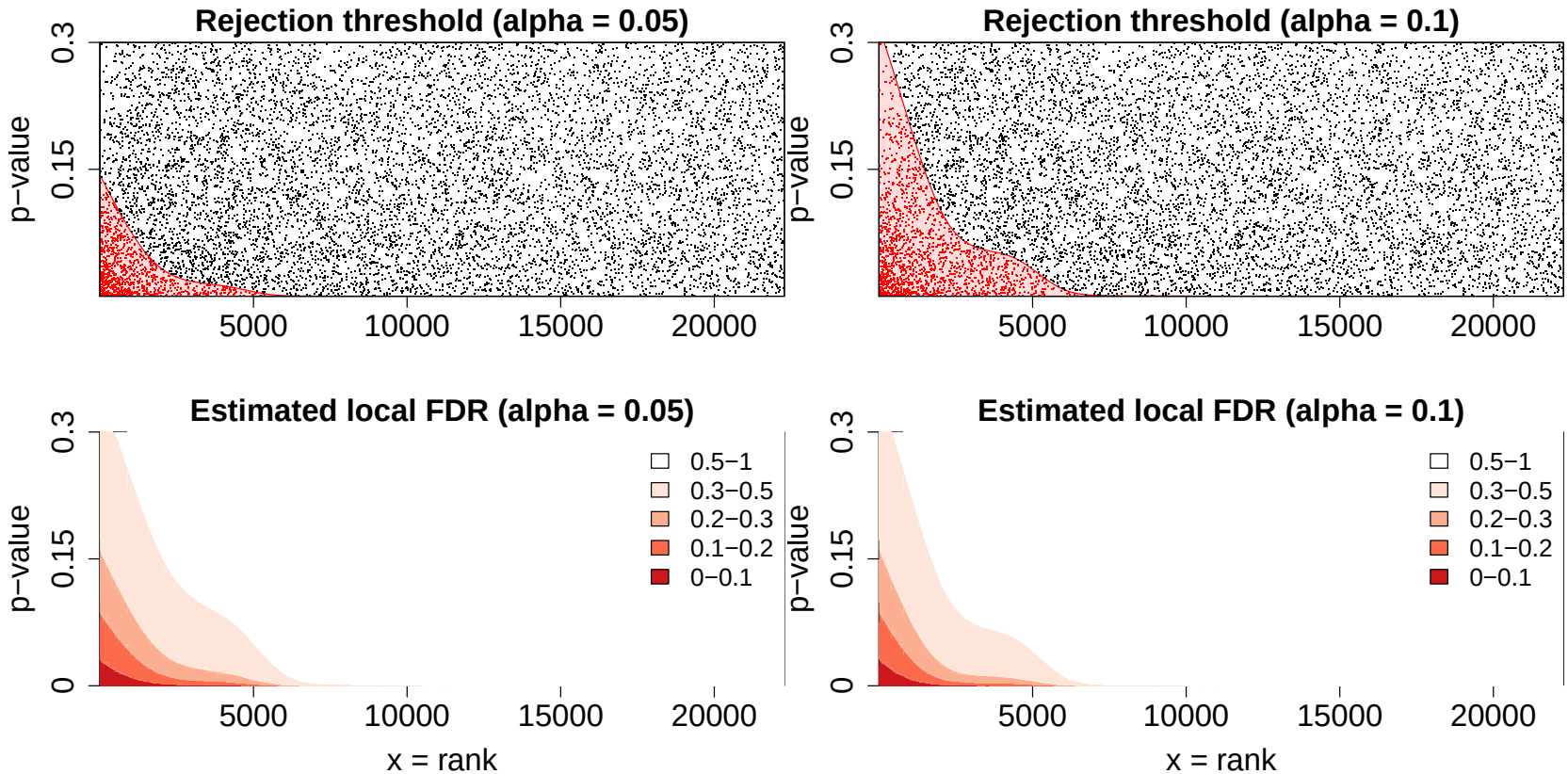
Natural-spline GLM for $\pi_1(x)$, $f_1(p \mid x)$ (fit by EM)



Gene/Drug response Data: Moderate Dose Ordering



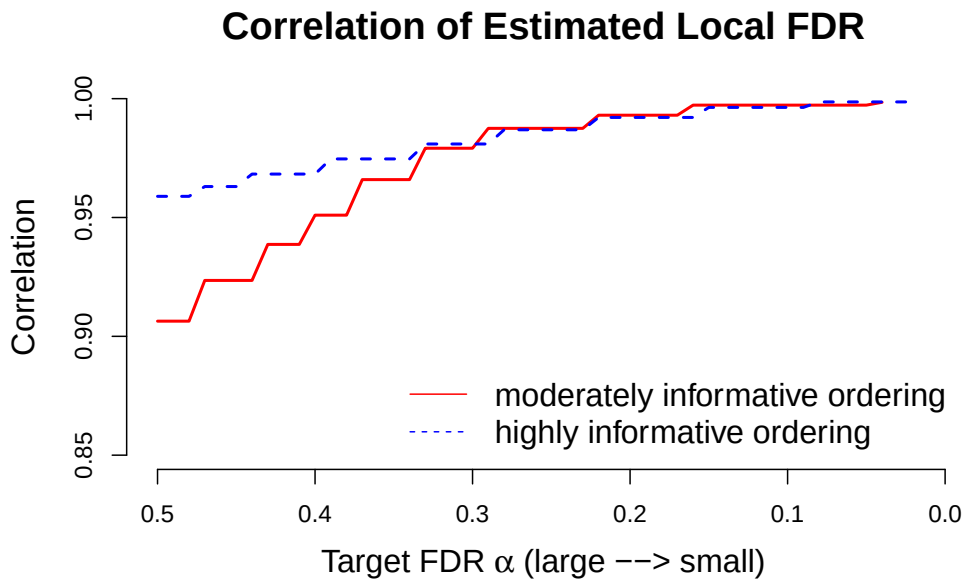
Gene/Drug response Data: High Dose Ordering



Convergence of fdr estimates

$\widehat{\text{fdr}}(p_i \mid x_i)$ at step t $\rightarrow$ $\widehat{\text{fdr}}(p_i \mid x_i)$ estimated using all data

Corr $\geq 90\%$ once $\widehat{\text{FDP}} \leq 0.3$.



Dependent p-values

Can deal with some forms of dependence using ideas from knockoff filter [Barber and Candès, 2015a]

Assume we observe correlated z-values $Z \sim N_d(\mu, \Sigma)$

Find diagonal $D \succeq \Sigma$, generate noise $\epsilon \sim N_d(0, D - \Sigma)$:

$$\begin{pmatrix} Z + \epsilon \\ W \end{pmatrix} \sim N\left(\begin{pmatrix} \mu \\ (I - \Sigma D^{-1})\mu \end{pmatrix}, \begin{pmatrix} D & 0 \\ 0 & * \end{pmatrix}\right)$$

where $W = Z - \Sigma D^{-1}(Z + \epsilon)$

Compute $p_i = 2 - 2\Phi(|Z_i + \epsilon_i|)$, use filtration

$$\mathcal{F}_t = \sigma(W, (x_i)_{i=1}^n, (\tilde{p}_{t,i})_{i=1}^n, A_t, R_t)$$

STAR: Multiple Testing With Structural Constraint

Lihua Lei (Berkeley Statistics)

Aaditya Ramdas (CMU Statistics)

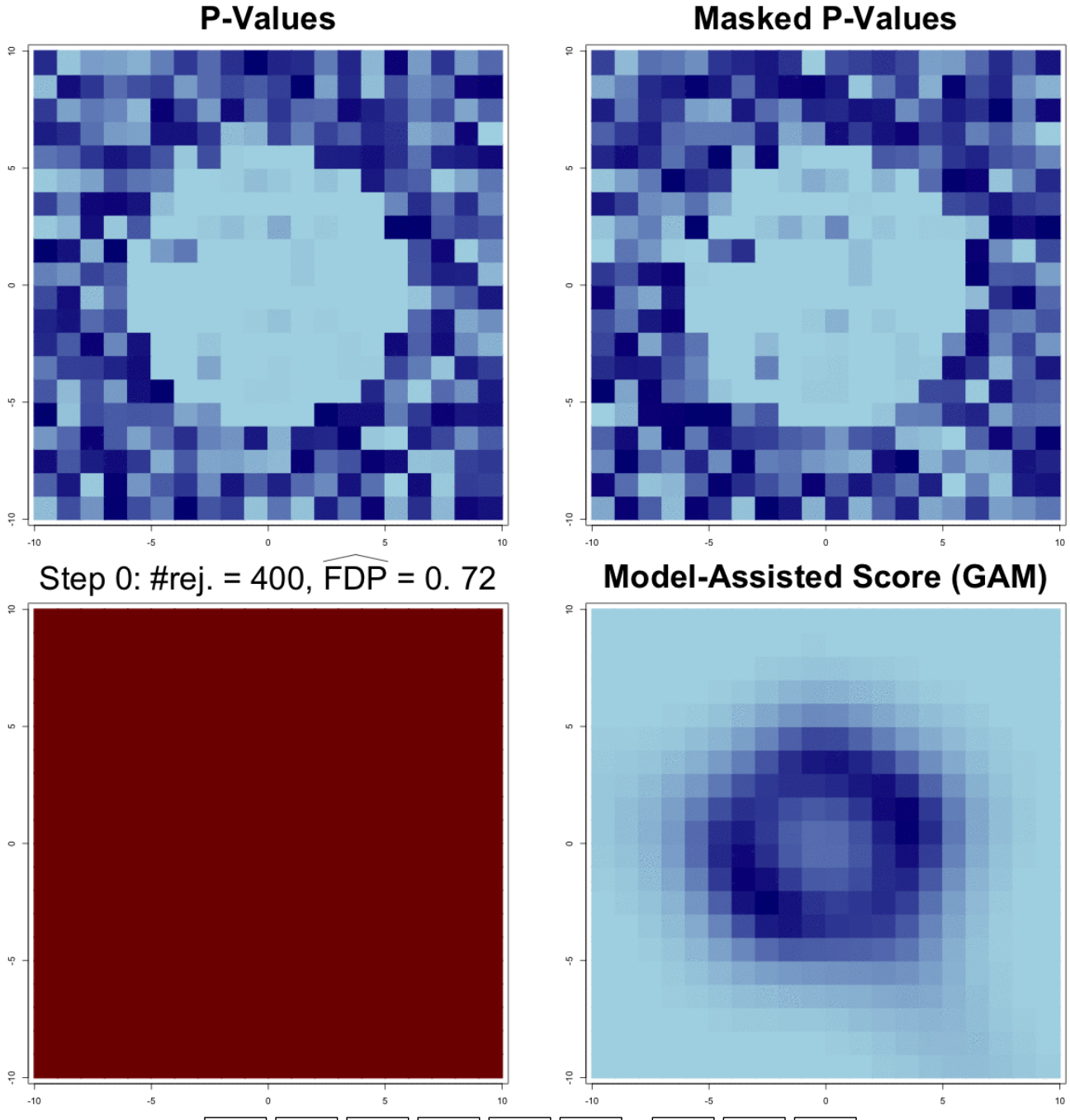
STAR: Multiple Testing With Structural Constraint

Extension with Lihua Lei and Aaditya Ramdas

Want to detect a subset of hypotheses subjected to certain structural constraint, i.e. $\{x_i : H_i \text{ rejected}\} \in \mathcal{K} \subseteq 2^{[n]}$:

- (spatial-temporal multiple testing)
  x_i: geographic location, $\mathcal{K}$: all convex sets;
- (hierarchical testing)
  x_i: node of a tree, $\mathcal{K}$: all subtrees;
- (Selection under strong/weak heredity principles)
  x_i: node of a DAG, $\mathcal{K}$: all subgraphs st. heredity principles

Animation: STAR in Convex Region Detection



Summary

AdaPT: Interactive multiple testing with side information

Estimate local fdr using partially masked data (impute p_i)

Unrestricted estimation of optimal threshold

Finite-sample FDR control

Thanks!

References

Rina Foygel Barber and Emmanuel J Candès. Controlling the false discovery rate via knockoffs. The Annals of Statistics, 43(5):2055–2085, 2015a.

Rina Foygel Barber and Emmanuel J Candès. Controlling the false discovery rate via knockoffs. The Annals of Statistics, 43(5):2055–2085, 2015b.

Yoav Benjamini and Yosef Hochberg. Controlling the false discovery rate: a practical and powerful approach to multiple testing. Journal of the Royal Statistical Society. Series B (Methodological), pages 289–300, 1995.

Kathryn R Coser, Jessica Chesnes, Jingyung Hur, Sandip Ray, Kurt J Isselbacher, and Toshi Shioda. Global analysis of ligand sensitivity of estrogen inducible and suppressible genes in mcf7/bus breast cancer cells by dna microarray. Proceedings of the National Academy of Sciences, 100(24): 13994–13999, 2003.

Sean Davis and Paul S Meltzer. Geoquery: a bridge between the gene expression omnibus (geo) and bioconductor. Bioinformatics, 23(14):1846–1847, 2007.

Bradley Efron, Robert Tibshirani, John D Storey, and Virginia Tusher. Empirical bayes analysis of a microarray experiment. Journal of the American statistical association, 96(456):1151–1160, 2001.

Egil Ferkingstad, Arnoldo Frigessi, Håvard Rue, Gudmar Thorleifsson, and Augustine Kong. Unsupervised empirical bayesian multiple testing with external covariates. The Annals of Applied Statistics, pages 714–735, 2008.

Dean P Foster and Robert A Stine. α-investing: a procedure for sequential control of expected false discoveries. Journal of the Royal Statistical Society: Series B (Statistical Methodology), 70(2): 429–444, 2008.

Max Grazier G'Sell, Stefan Wager, Alexandra Chouldechova, and Robert Tibshirani. Sequential selection procedures and false discovery rate control. Journal of the Royal Statistical Society: Series B (Statistical Methodology), 2015.

Nikolaos Ignatiadis, Bernd Klaus, Judith B Zaugg, and Wolfgang Huber. Data-driven hypothesis weighting increases detection power in genome-scale multiple testing. Nature methods, 2016.

Ang Li and Rina Foygel Barber. Accumulation tests for FDR control in ordered hypothesis testing. Journal of the American Statistical Association, (just-accepted):1–38, 2016a.

Ang Li and Rina Foygel Barber. Multiple testing with the structure adaptive benjamini-hochberg algorithm. arXiv preprint arXiv:1606.07926, 2016b.

John D Storey, Jonathan E Taylor, and David Siegmund. Strong control, conservative point estimation and simultaneous conservative consistency of false discovery rates: a unified approach. Journal of the Royal Statistical Society: Series B (Statistical Methodology), 66(1):187–205, 2004.