Controlling False Discovery Rate via Knockoffs

jinjiang zou & zeyu wang

Based on Barber & Candès (Annals of Statistics, 2015)

12 31 2025

Happy New Year!

Paper available at `http://arxiv.org/abs/1404.5609`

We gratefully acknowledge R.F.Barber for providing the materials used in this work.

About the Authors

Rina Foygel Barber

Louis Block Professor in the Department of Statistics at the University of Chicago

- Elected to National Academy of Sciences 2025
- 2020 COPSS Presidents' Award
- MacArthur Fellowship,2023 (known as the 'genius award')

Emmanuel Candès

Barnum-Simons Chair in Mathematics and Statistics at Stanford University

- Fellow of the American Academy of Arts & Sciences, 2014
- Member, National Academy of Sciences, 2014
- MacArthur Fellowship,2017

Setting

An example:
Which mutations in the reverse transcriptase (RT) of HIV-1 determine susceptibility to reverse transcriptase inhibitors (RTIs)?

$$y_i \in \mathbb{R} = \text{resistance of virus in sample } i \text{ to a RTI-type drug}$$

$$X_{ij} \in \{0, 1\} \text{ indicates if mutation } j \text{ is present in virus sample } i$$

How can we select mutations that determine drug resistance, in such a way that our answer will replicate in further trials?

Setting

Sparse linear model:

$$y = X \cdot \beta + z, \text{ where } z_i \overset{\text{iid}}{\sim} \mathcal{N}(0, \sigma^2)$$

- n observations, p features
- β is sparse

Setting

Goal: select a set of features X_j that are likely to be relevant to the response y, without too many false positives.

One way to measure performance:

$$FDR = \mathbb{E}\left[\frac{\#\text{ false positives}}{\text{total }\#\text{ of features selected}}\right] = \mathbb{E}\left[\frac{|S \cap \mathcal{H}_0|}{|S|}\right].$$

- $S =$ set of selected features
- $\mathcal{H}_0 =$ "null hypotheses" $= \{j : \beta_j^* = 0\}$

Sparse Regression

Lasso:

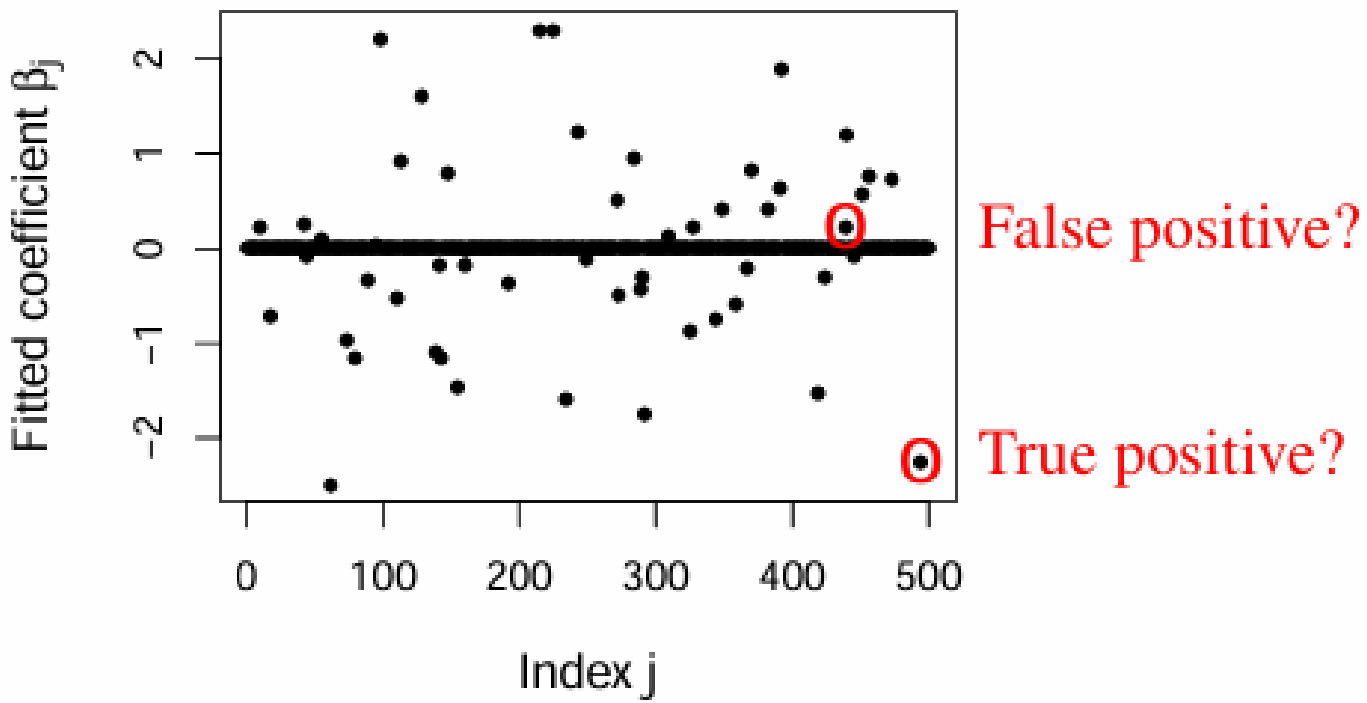
$$\beta_\lambda = \arg\min_{\beta \in \mathbb{R}^p} \left\{ \frac{1}{2} \|y - X \cdot \beta\|_2^2 + \lambda \|\beta\|_1 \right\}$$

Asymptotically, Lasso will select the correct model (at a good λ).
In practice for a finite sample,

- True positives & false positives intermixed along the Lasso path
- How to pick λ to balance FDR vs power?
- Need to account for correlations between X_j & weak signals

Sparse Regression

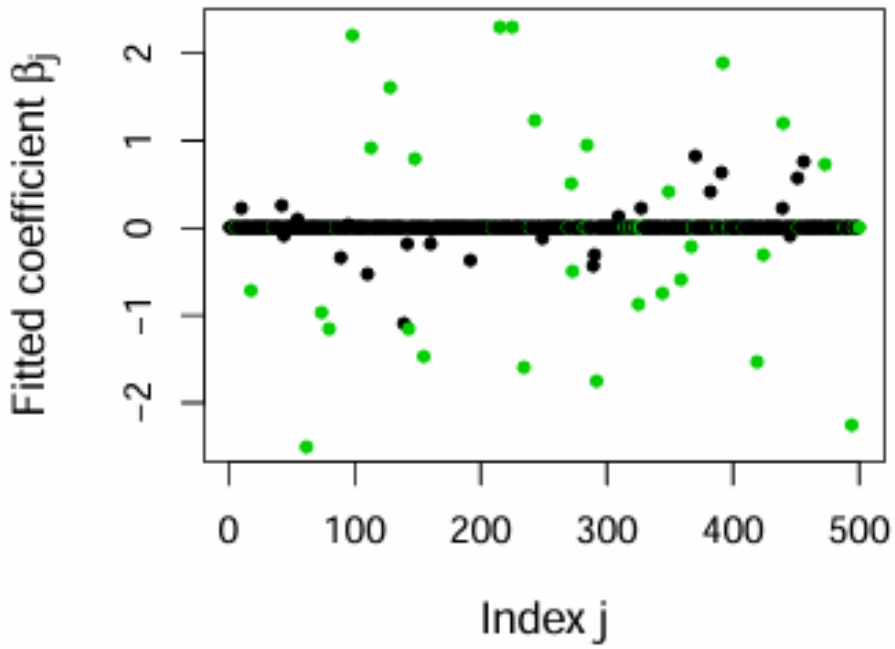
Simulated data with $n = 1500, p = 500$.
Lasso fitted model for $\lambda = 1.75$:



Sparse Regression

Simulated data with $n = 1500, p = 500$.
Lasso fitted model for $\lambda = 1.75$:

- **FDP** $= \frac{26}{55} = 47\%$



To estimate FDP, would need to calculate distribution of β_j^λ for null j (would need to know σ^2, β^*, ...). (Donoho et al 2009)

Construct Knockoffs

Main idea:
For each feature X_j, construct a knockoff version $\tilde{X}_j$.
The knockoffs serve as a "control group" $\Rightarrow$ can estimate FDP.

Setting:

- Require $n > p$
- Don't need to know σ^2
- Don't need any information about β^*
- Will get an exact, finite-sample guarantee for FDR

Construct Knockoffs

Construction:

- The knockoffs replicate the correlation structure of X:

$$\widetilde{X}_j^T \widetilde{X}_k = X_j^T X_k \text{ for all } j, k$$

- Also preserve correlations between knockoffs & originals:

$$\widetilde{X}_j^T X_k = X_j^T X_k \text{ for all } j \neq k$$

Augmented design matrix

$$[X\widetilde{X}] = (X_1, X_2, \ldots, X_p, \widetilde{X}_1, \widetilde{X}_2, \ldots, \widetilde{X}_p) \in \mathbb{R}^{n \times 2p}$$

Construct Knockoffs

How?

Define $\tilde{X} = X \cdot (I_p - 2\xi\Sigma^{-1}) + U \cdot C$, where:

$$\Sigma = X^\top X \succeq \xi I_p$$

$$U = n \times p \text{ orthonormal matrix orthogonal to } X$$

$$C^\top C = 4(\xi I_p - \xi^2 \Sigma^{-1}) \quad \text{(Cholesky decomposition)}$$

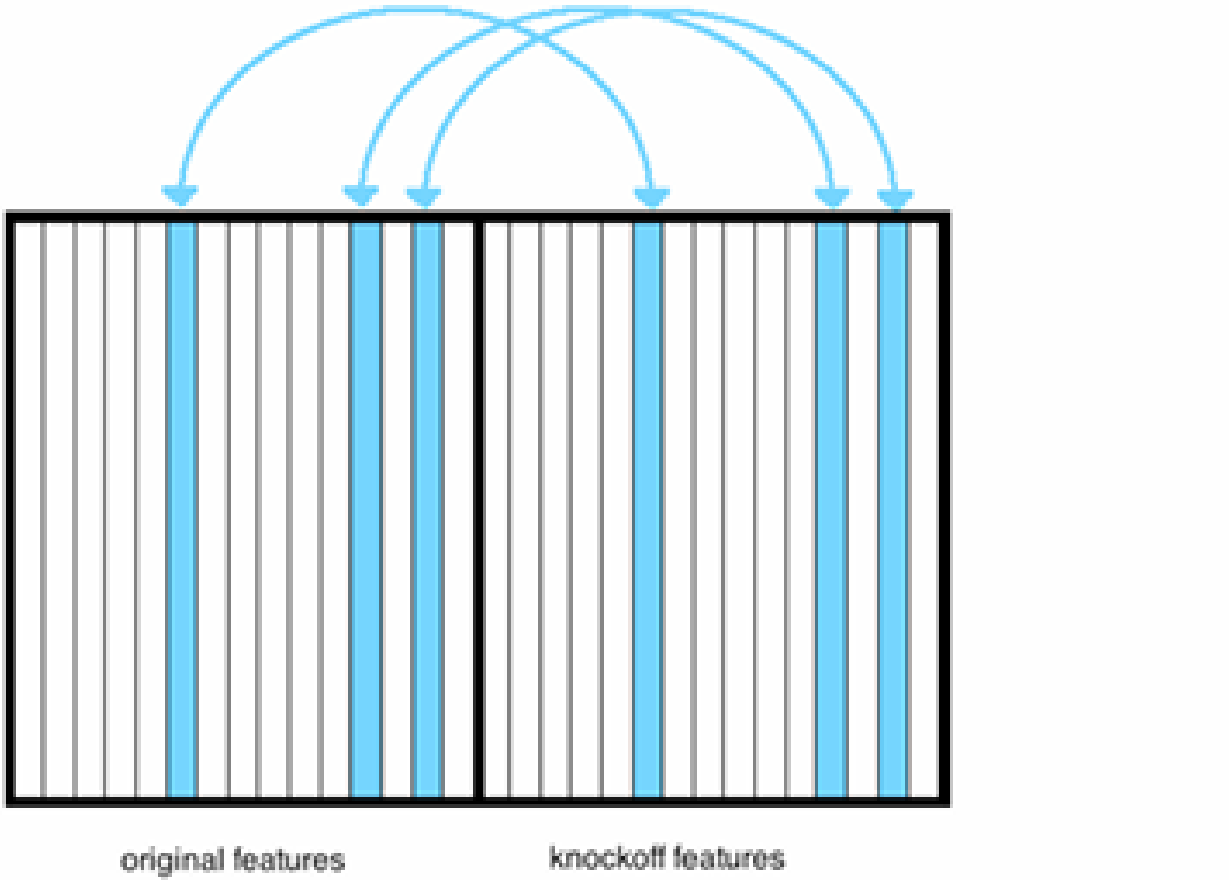
$$\Rightarrow [X\tilde{X}]^\top [X\tilde{X}] = \begin{pmatrix} \Sigma & \Sigma - 2\xi I_p \\ \Sigma - 2\xi I_p & \Sigma \end{pmatrix}$$

Construct Knockoffs

Why?

For a null feature X_j,

$$X_j^T y = X_j^T X\beta^* + X_j^T z \overset{\mathcal{D}}{=} \tilde{X}_j^T X\beta^* + \tilde{X}_j^T z = \tilde{X}_j^T y$$



Construct Knockoffs

Lemma (Pairwise exchangeability property)

For any $N \subset \mathcal{H}_0$,

$$\left(\begin{bmatrix} X & \tilde{X} \end{bmatrix}_{swap(N)} \right)^T y \overset{\mathcal{D}}{=} \begin{bmatrix} X & \tilde{X} \end{bmatrix}^T y$$

$\Longrightarrow$ the knockoffs are a "control group" for the nulls

Knockoff Method

Steps:

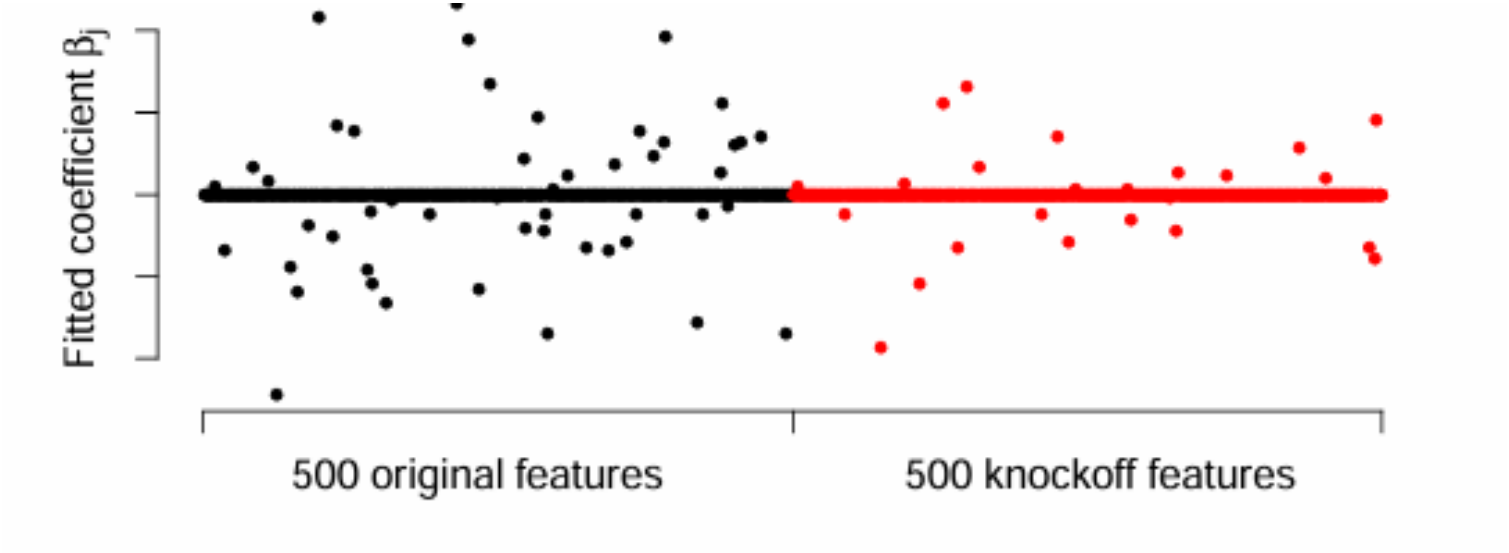
1. Construct knockoffs
2. Compute Lasso with augmented matrix:

$$\beta_\lambda = \arg\min_{\beta \in \mathbb{R}^{2p}} \left\{ \frac{1}{2} \left\| y - [X\widetilde{X}] \cdot \beta \right\|_2^2 + \lambda \|\beta\|_1 \right\}$$

3. Use $\widetilde{X}_j$ as a "control group" for X_j

Knockoff Method

Fitted model for $\lambda = 1.75$ on the simulated dataset:



- **Lasso selects 49 original features & 24 knockoff features**
- Pairwise exchangeability of the nulls $\Longrightarrow$ probably ≈ 24 false positives among the 49 original features

Knockoff Method

Compute Lasso on the entire path $\lambda \in [0, \infty)$.

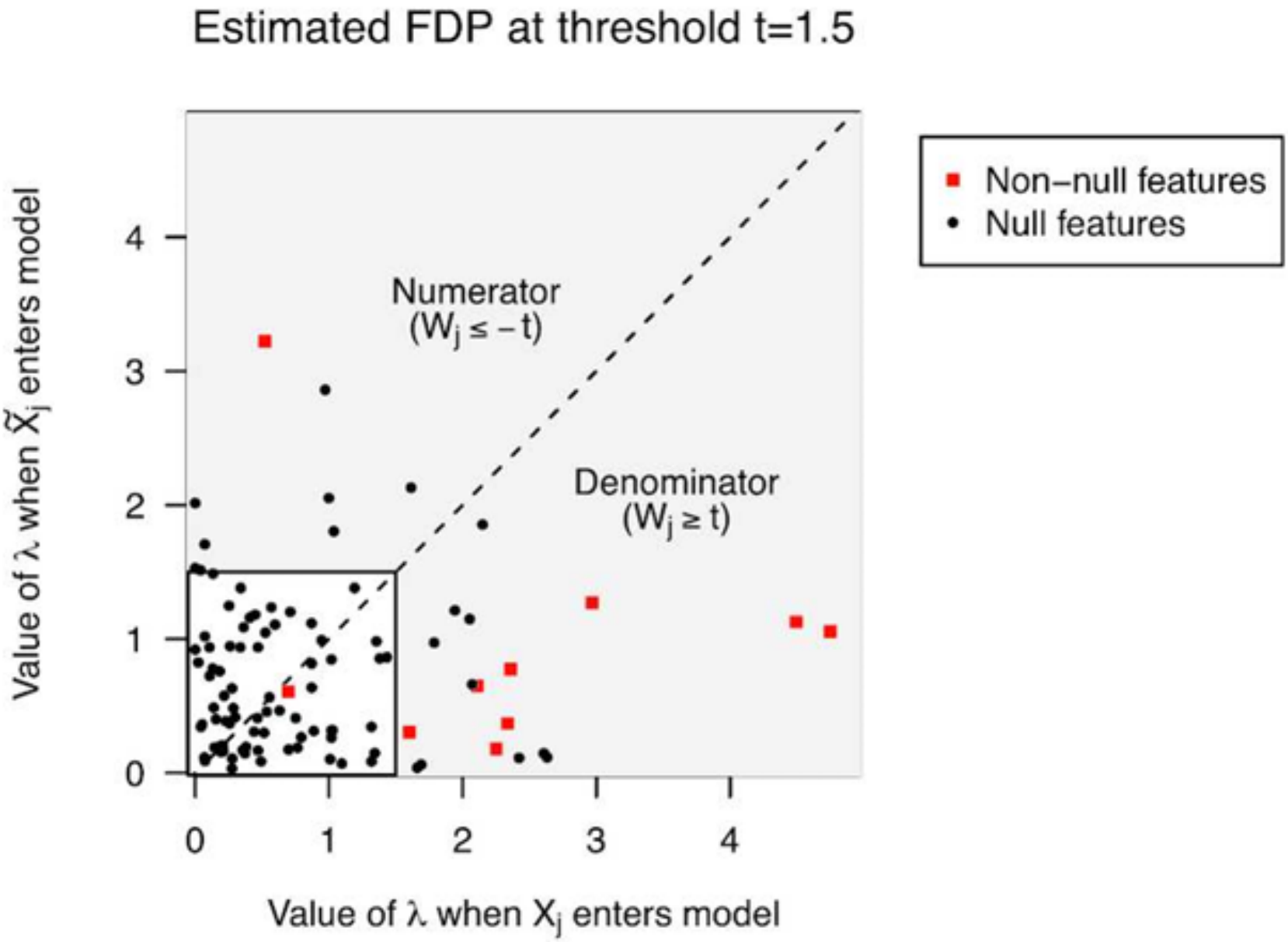
$$\lambda_j = \sup\left\{\lambda : \beta_j^\lambda \neq 0\right\} = \text{first time } X_j \text{ enters Lasso path}$$

$$\tilde{\lambda}_j = \sup\left\{\lambda : \tilde{\beta}_j^\lambda \neq 0\right\} = \text{first time } \tilde{X}_j \text{ enters Lasso path}$$

Then define statistics

$$W_j = \max\{\lambda_j, \tilde{\lambda}_j\} \cdot \text{sign}(\lambda_j - \tilde{\lambda}_j)$$

Knockoff Method



Knockoff Method

Lemma (Pairwise exchangeability of the nulls)

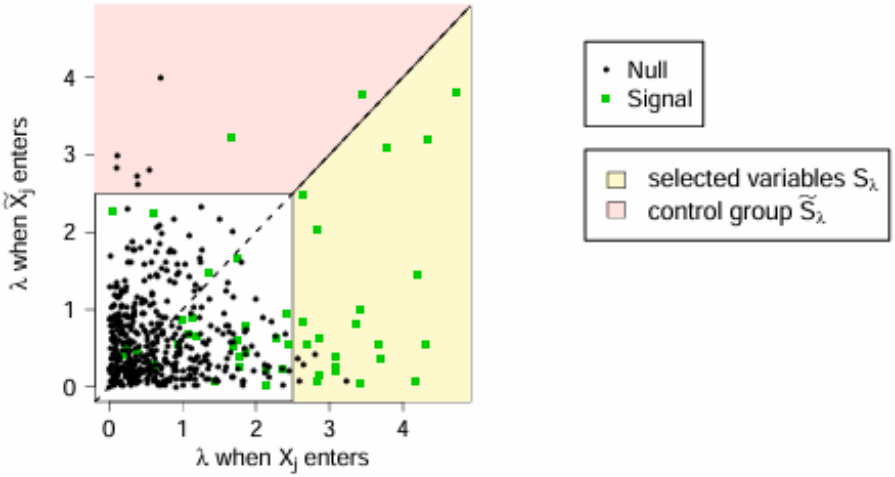
$$(W_1, W_2, \ldots, W_p) \overset{\mathcal{D}}{=} (|W_1| \cdot \epsilon_1, |W_2| \cdot \epsilon_2, \ldots, |W_p| \cdot \epsilon_p)$$

where $\epsilon_j = sign(W_j)$ *for non-nulls and* $\epsilon_j \overset{iid}{\sim} \{\pm 1\}$ *for nulls.*

Knockoff Method

Selected variables: $S_\lambda = \{j : W_j \geq +\lambda\}$,

Control group: $\widetilde{S}_\lambda = \{j : W_j \leq -\lambda\}$, $\qquad \widehat{\mathrm{FDP}}(S_\lambda) := \dfrac{|\widetilde{S}_\lambda|}{|S_\lambda|}$.



$$\mathrm{FDP}(S_\lambda) = \frac{|S_\lambda \cap \mathcal{H}_0|}{|S_\lambda|} \approx \frac{|\widetilde{S}_\lambda \cap \mathcal{H}_0|}{|S_\lambda|} \leq \widehat{\mathrm{FDP}}(S_\lambda).$$

Knockoff Method

The knockoff filter: define

$$FDP(S_\lambda) := \frac{|\widetilde{S}_\lambda|}{|S_\lambda|} = \frac{\#\{j : W_j \leq -\lambda\}}{\#\{j : W_j \geq +\lambda\}},$$

then choose

$$\Lambda = \min\{\lambda : FDP(S_\lambda) \leq q\} \quad \text{(or } \lambda = \infty \text{ if empty set)}$$

and select the variable set

$$S_\Lambda = \{j : W_j \geq \Lambda\}.$$

Theoretical Guarantees

Theorem 1: For S_Λ chosen by the knockoff filter,

$$\mathbb{E}\left[mFDP(S_\Lambda)\right] \leq q$$

where the modified FDP is given by

$$mFDP(S) = \frac{|S \cap \mathcal{H}_0|}{|S| + q^{-1}}.$$

Theoretical Guarantees

The knockoff+ filter: define

$$FDP_+(S_\lambda) := \frac{|\widetilde{S}_\lambda| + 1}{|S_\lambda|} = \frac{\#\{j : W_j \le -\lambda\} + 1}{\#\{j : W_j \ge +\lambda\}},$$

then choose

$$\Lambda_+ = \min\{\lambda : FDP_+(S_\lambda) \le q\} \quad (\text{or } \lambda = \infty \text{ if empty set})$$

and select the variable set

$$S_{\Lambda_+} = \{j : W_j \ge \Lambda_+\}.$$

Theoretical Guarantee: Knockoff

Theorem (Knockoff FDR Control)

For any $q \in [0, 1]$, the knockoff filter satisfies

$$\mathbb{E}\left[\frac{\#\{j : \beta_j = 0 \text{ and } j \in \widehat{S}\}}{\#\{j : j \in \widehat{S}\} + q^{-1}}\right] \leq q,$$

where the expectation is taken over the Gaussian noise z in the linear model

$$y = X\beta + z,$$

while treating the design matrix X and its knockoff $\widetilde{X}$ as fixed.

Theoretical Guarantees

Theorem 2: For S_{Λ_+} chosen by the knockoff+ filter,

$$\mathbb{E}\left[\mathrm{FDP}(S_{\Lambda_+})\right] \leq q.$$

Proof sketch:

$$\mathrm{FDP}(S_{\Lambda_+}) = \frac{|S_{\Lambda_+} \cap \mathcal{H}_0|}{|S_{\Lambda_+}|} = \frac{|\widetilde{S}_{\Lambda_+} \cap \mathcal{H}_0| + 1}{|S_{\Lambda_+}|} \cdot \frac{|S_{\Lambda_+} \cap \mathcal{H}_0|}{|\widetilde{S}_{\Lambda_+} \cap \mathcal{H}_0| + 1}$$

$$M(\lambda) = \frac{|S_\lambda \cap \mathcal{H}_0|}{|S_\lambda \cap \mathcal{H}_0| + 1}$$

is a supermartingale w.r.t. increasing λ, and Λ_+ is a stopping time.

Theoretical Guarantees

Theorem (Knockoff+ FDR Control)

For any $q \in [0, 1]$, the knockoff+ filter satisfies

$$\mathrm{FDR} = \mathbb{E}\left[\frac{\#\{j : \beta_j = 0 \text{ and } j \in \widehat{S}\}}{\#\{j : j \in \widehat{S}\} \vee 1}\right] \leq q,$$

where the expectation is taken over the Gaussian noise z, while treating X and $\widetilde{X}$ as fixed.

HIV Data

Which mutations in the RT or protease of HIV-1 determine susceptibility to RT inhibitors or protease inhibitors?
Data:
Genotypic predictors of HIV type 1 drug resistance, Rhee et al (2006)
Available at `hivdb.stanford.edu` (Stanford HIV Drug Resistance Database)

- Each drug analysed separately
- Response y = resistance to the drug
- Features X = which mutations are present in the RT or in the protease

HIV Data

The data set:

Drug type	# drugs	Sample size	# protease or RT positions genotyped
PI	6	848	99
NRTI	6	639	240
NNRTI	3	747	240

To validate results:

- Treatment-selected mutation (TSM) panel: A separate study identifies mutations frequently present in patients who have been treated with each type of drug

Benjamini–Hochberg (BHq) Procedure

Given p hypotheses with p-values $P_1, \ldots, P_p$.

1. Sort p-values:
$$P_{(1)} \leq \cdots \leq P_{(p)}.$$
2. Find
$$k = \max \left\{ i : P_{(i)} \leq \frac{i}{p} q \right\}.$$
3. Reject hypotheses $H_{(1)}, \ldots, H_{(k)}$.

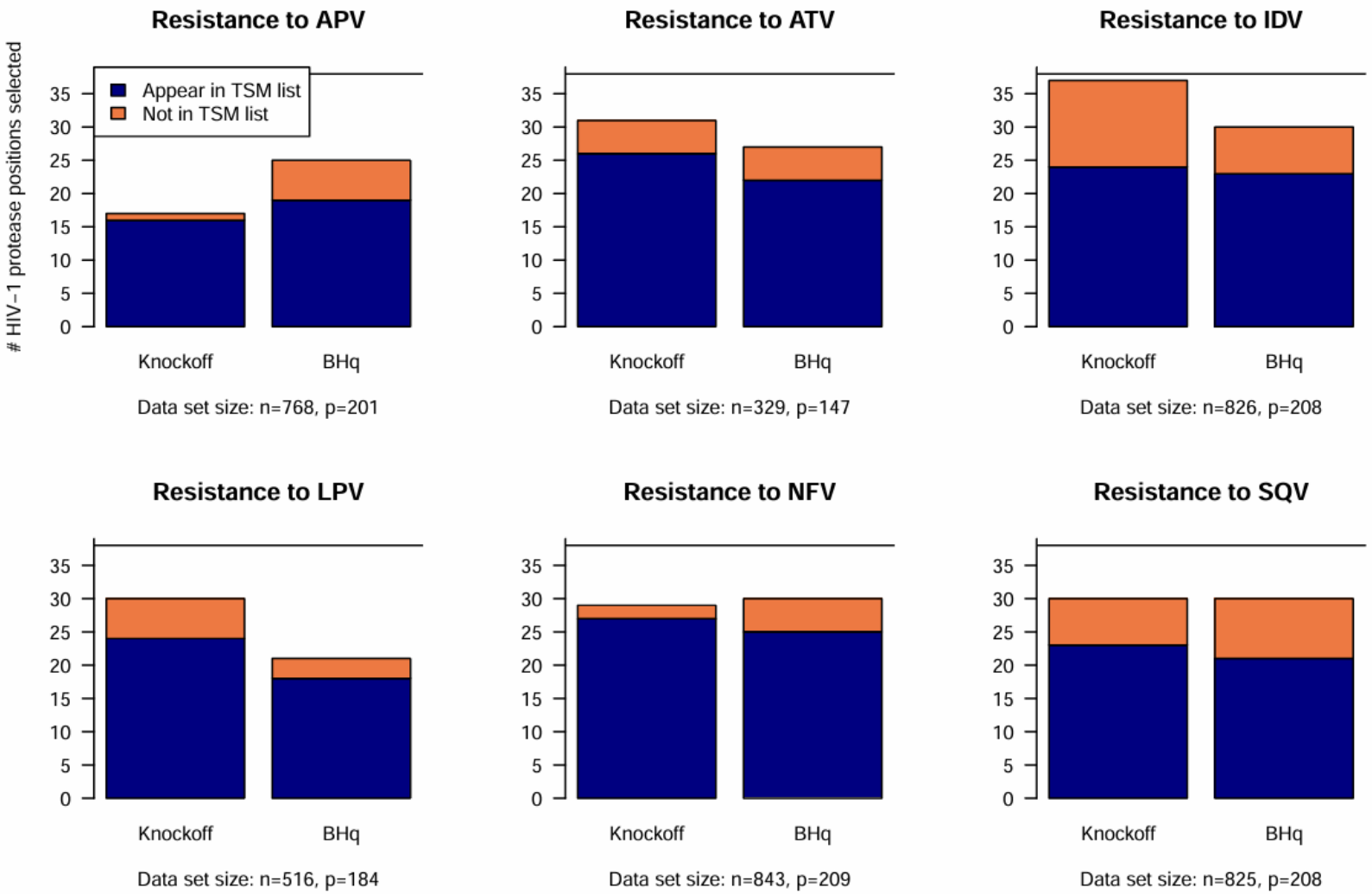
Guarantee: If null p-values are independent and uniform,

$$\mathrm{FDR} \leq q.$$

BHq relies on valid p-values, which are difficult to obtain in high-dimensional regression with correlated features. Knockoffs avoid this by replacing p-values with a competition between original variables and their knockoffs.

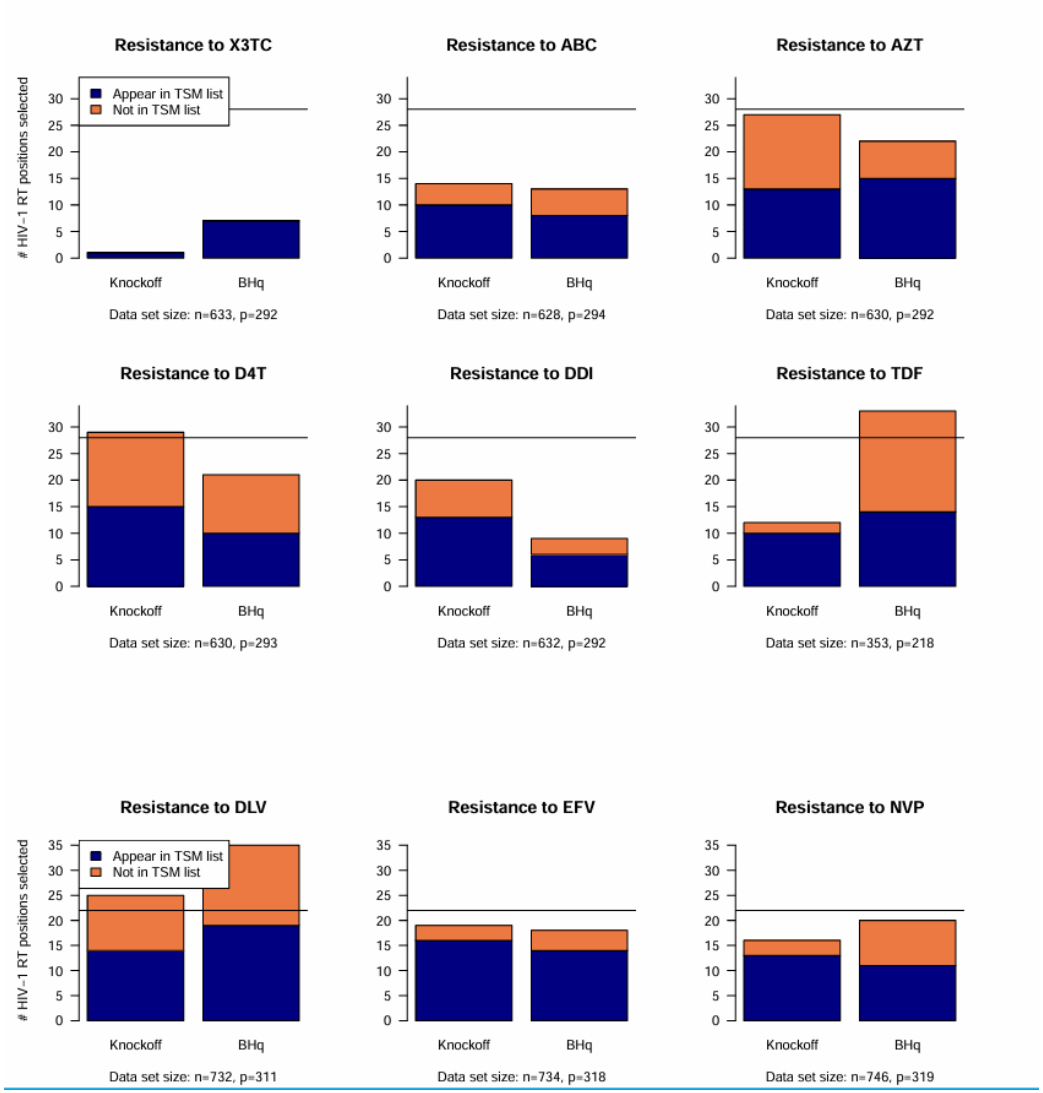
HIV Data

Results for PI type drugs



HIV Data

Results for NRTI and NNRTI type drugs

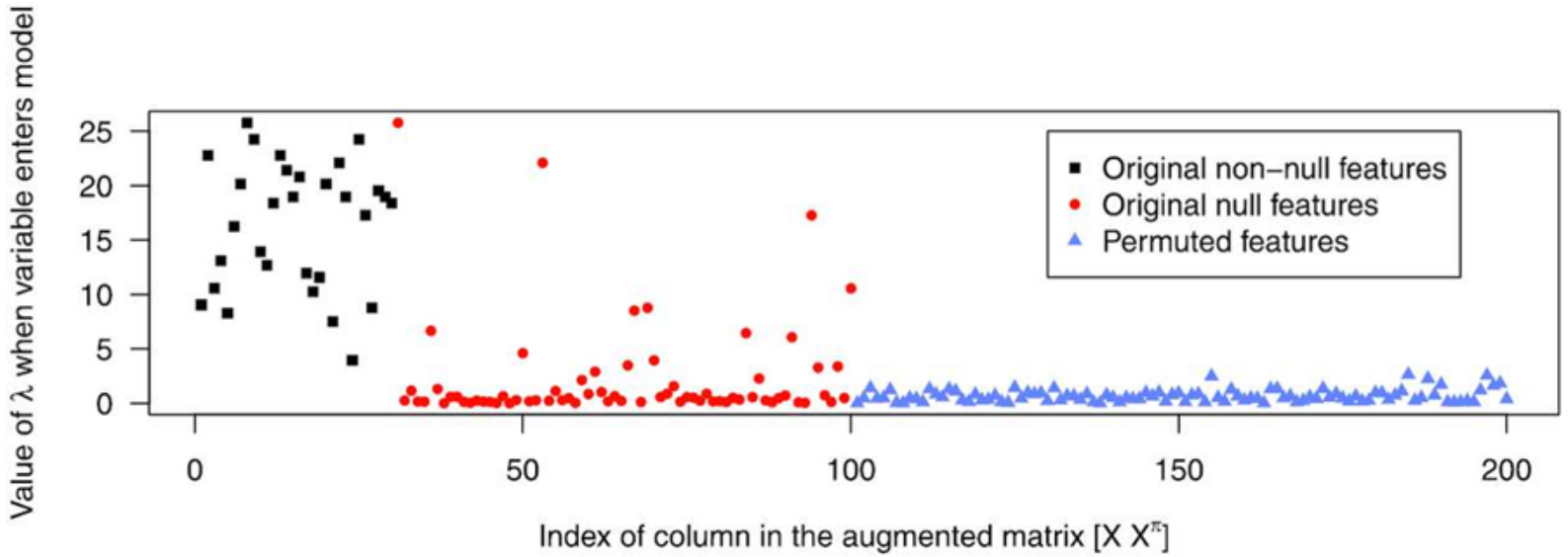


Can Knockoffs be Replaced by Permutations?

Let $X^{\pi} = X$ with rows randomly permuted. Then

$$[X\ X^{\pi}]^{T}[X\ X^{\pi}] \approx \begin{pmatrix} \Sigma & 0 \\ 0 & \Sigma \end{pmatrix}$$

Method	FDR(target level $q = 20\%$)
Knockoff method	12.29%
Permutation method	45.61%



Summary

The knockoff filter for inference in a sparse linear model:

- Creates a "control group" for any type of statistic
- Handles any type of feature correlation
- Unknown noise level & sparsity level
- Finite-sample FDR guarantees

Summary

Future work:

1. Extend to GLMs or other regression models?
2. Similar principles for other problems, e.g. graphical models?
3. Absence of distributional assumptions.

Questions?

1. If we do not select any features, our FDR is 0. But it is not a useful way. We know knockoff depends on W_j, but I think this method could not guarantee your selection is enough.
2. We know we can make independent copies of samples. But we can make two sets of samples using methods like the methods introduced in section 2.
3. The differences of using different f satisfying our assumption.

Thank you!