When knockoffs fail: diagnosing and fixing non-exchangeability of knockoffs

Angel REYERO LOBO

10th Channel Network Conference
IMT & Inria Paris-Saclay
Joint work with:
Alexandre BLAIN, Julia LINHART, Bertrand THIRION & Pierre NEUVIAL.

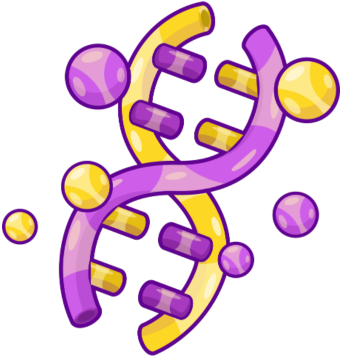
20th of May

Contents

Index

Motivation: high-dimensional feature selection

Which genes are significant in relation to the disease?



X

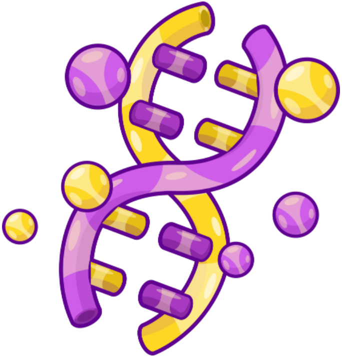
y

We need to find

$$\mathcal{H}_1 := \{j : X^j \not\perp\!\!\!\perp y\}.$$

Motivation: high-dimensional feature selection

Which genes are significant in relation to the disease?



X

y

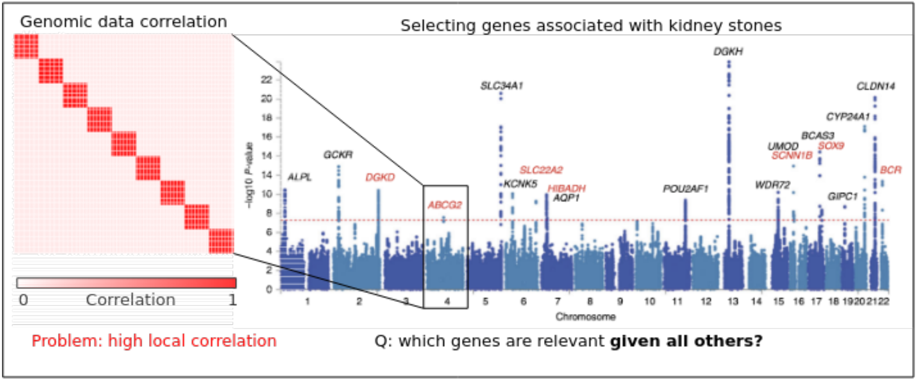
💡 We need to find

$$\mathscr{H}_1 := \{j : X^j \not\perp\!\!\!\perp y\}.$$

Conditional variable selection

$$\mathcal{H}_1 := \{ j : X^j \not\perp\!\!\!\perp y \mid X^{-j} \} .$$

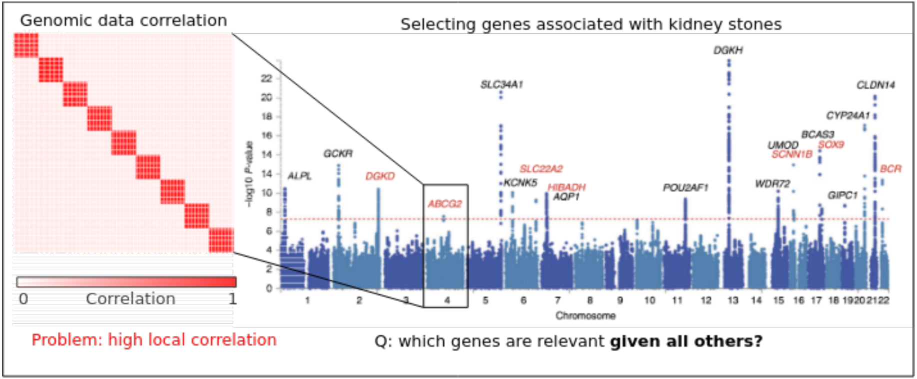
Highly correlated genomic data
$\Rightarrow$ **marginally**, many genes are important.
Are they only important because of correlation?



Conditional variable selection

$$\mathcal{H}_1 := \{ j : X^j \not\perp\!\!\!\perp y \mid \mathbf{X}^{-\mathbf{j}} \}.$$

⚠ **Highly correlated** genomic data
⇒ **marginally**, many genes are important.
Are they only important because of correlation?



We need to select relevant variables **conditionally** on all the others!

Index

Knockoffs framework

Knockoffs (from Candès et al. (2018)):

1. **Knockoff variables ($\widetilde{X}$):** *imitations* of X independent of y:
   - $[X, \widetilde{X}]_{\text{swap}(s)} \stackrel{d}{=} [X, \widetilde{X}]$ for any $s \subset [p]$.

$$\left[X_1, X_2, X_3, \widetilde{X}_1, \widetilde{X}_2, \widetilde{X}_3\right]_{\text{swap}(1,3)} := \left[\widetilde{X}_1, X_2, \widetilde{X}_3, X_1, \widetilde{X}_2, X_3\right]$$

   - $\widetilde{X} \perp\!\!\!\perp y \mid X$.

Example 1 (Gaussian knockoffs)

If $X \sim \mathcal{N}(0, \Sigma)$, then $[X, \widetilde{X}] \sim \mathcal{N}(0, G)$ with
$G = \begin{bmatrix} \Sigma & \Sigma - \text{diag}(\lambda) \\ \Sigma - \text{diag}(\lambda) & \Sigma \end{bmatrix}$ for $\lambda \in \mathbb{R}^p$ such that $G \succeq 0$:

- $[X, \widetilde{X}]$ is **exchangeable**.
- $\widetilde{X} \mid X$ is known!

Knockoffs framework

Knockoffs (from Candès et al. (2018)):

1. **Knockoff variables ($\widetilde{X}$):** *imitations* of X independent of y:
2. **Knockoffs statistics (W $\in \mathbb{R}^p$):** comparison of *predictability* between the original covariate and its knockoff.

Example 1 (Lasso Coefficients Difference)

Regress $y \in \mathbb{R}$ on $[X, \widetilde{X}] \in \mathbb{R}^{2p}$ using LASSO and compute the estimated coefficient difference: $W_j = |\widehat{\beta}_j| - |\widehat{\beta}_{j+p}|$.

Knockoffs framework

Knockoffs (from Candès et al. (2018)):

1. **Knockoff variables ($\widetilde{X}$):** *imitations* of X independent of y:
2. **Knockoffs statistics ($\mathbf{W} \in \mathbb{R}^p$):** comparison of *predictability* between the original covariate and its knockoff.
3. **Threshold:** it is data-dependent and given by

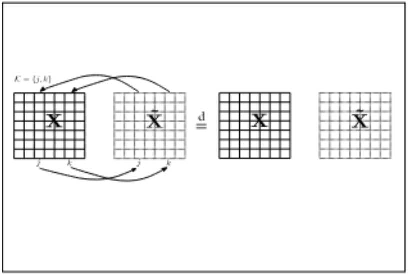
$$T_q^\star = \min\left\{ t \in \mathscr{W} : \frac{1 + \#\{j : W_j \leq -t\}}{\#\{j : W_j \geq t\} \vee 1} \leq q \right\}.$$

and $+\infty$ if empty.

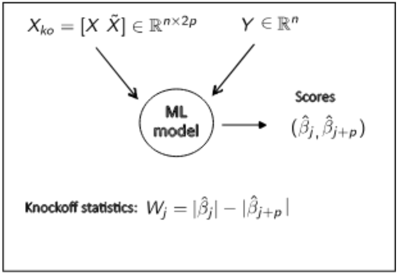
This procedure provides **FDR control**!$(\mathrm{FDR}(\{j : W_j \geq T_q^\star\}) \leq q)$

$$\mathrm{FDP}(\widehat{S}) := \frac{\left|\widehat{S} \cap \mathscr{H}_0\right|}{\left|\widehat{S}\right| \vee 1} \qquad \mathrm{FDR}(\widehat{S}) := \mathbb{E}\left[\mathrm{FDP}(\widehat{S})\right].$$

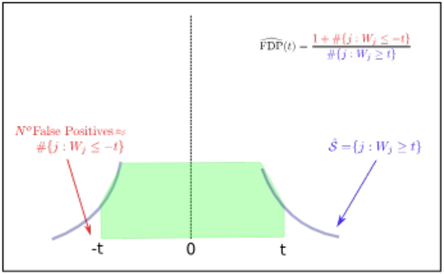
Knockoffs framework



Knockoff variables



Building Knockoff statistics



Inferring on Knockoff statistics

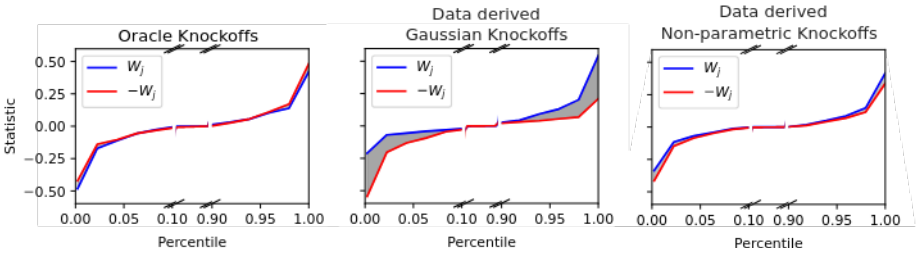
Index

Impact of exchangeability violation

- In practice, the distribution is unknown and generate $\widetilde{X}$ is complex.
- FDR control relies on null covariates being symmetric around 0:

Lemma 1 (from Candès et al. (2018))

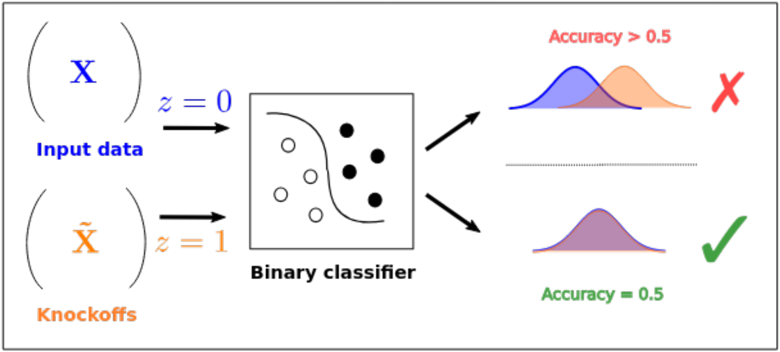
Conditionally on $(|W_1|, \ldots, |W_p|)$, the signs of the null W_js, $j \in \mathscr{H}_0$, are IID coin flips.



A necessary condition for exchangeability

- **Exchangeability:** $[X, \widetilde{X}]_{\mathrm{swap}(s)} \overset{\mathrm{d}}{=} [X, \widetilde{X}]$ for any $s \subset [p]$.
- In particular for $s = [p]$!

Diagnosing Knockoffs exchangeability with a Classifier Two-Sample Test



- **C2ST-validity:** classifier accuracy close to 0.5 (chance level).

Index

Settings

Simulated data:

- $X \in \mathbb{R}^{n \times p}$ smooth flattened 3D structure
- w smoothing kernel width
- $\beta^\star \in \{0,1\}^p$
- $s_p = \|\beta^\star\|_0 / p$
- $y = X\beta^\star + \sigma\varepsilon$
- $\sigma = \|X\beta^\star\| / (\mathrm{SNR}\|\varepsilon\|)$

Semi-simulated data:

- Functional Magnetic Resonance Imaging (fMRI)

1. Split in (X_1, y_1) and (X_2, y_2)
2. Perform a lasso on (X_1, y_1) to obtain $\beta_1^\star$
3. $y_2^{\mathrm{sim}} = X_2\beta_1^\star + \sigma\varepsilon$

Methods

Knockoffs generation methods:

- `VAE` from Zhu et al. (2021)(DeepLINK)
- `Gaussian` from Candès et al. (2018)(Graphical Lasso)
- `Deep` from Romano et al. (2020)(Deep Knockoffs)
- `Parallel` and `Auto-regressive` from Blain et al. (2025)(ours)

C2ST accuracy

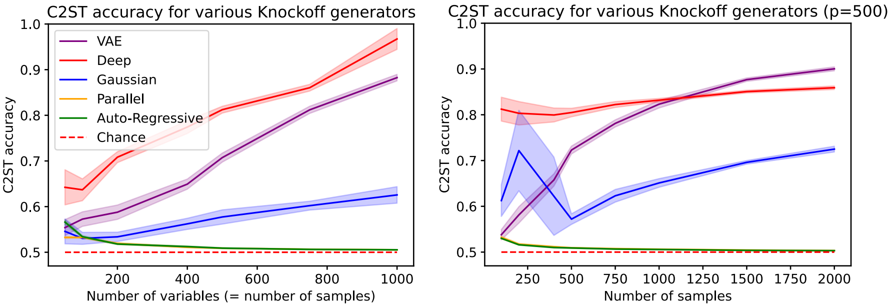


Figure 1: C2ST diagnostic metric for varying n and p.

Varying smoothing

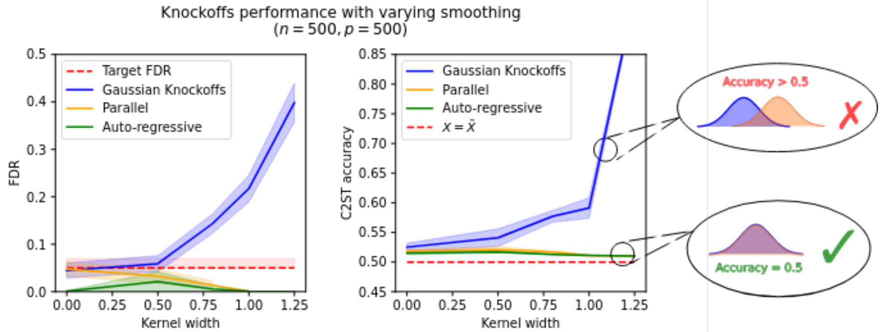


Figure 2: FDP and C2ST diagnostic metric for varying smoothing.

Semi-simulated datasets

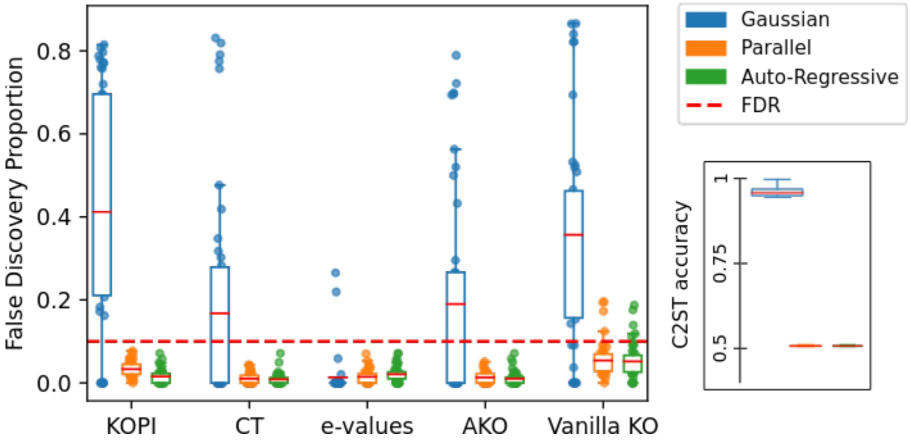


Figure 3: **Empirical FDP on semi-simulated data for 42 contrast pairs using Gaussian vs nonparametric knockoffs.** We use 7 HCP contrasts C0: "Motor Hand", C1: "Motor Foot", C2: "Gambling", C3: "Relational", C4: "Emotion", C5: "Social", C6: "Working Memory".

Index

Take-home message

$$\mathscr{H}_1 := \{ j : X^j \not\perp\!\!\!\perp y \mid \mathbf{X^{-j}} \} .$$

- Knockoffs provide an **efficient** and **statistically-valid** method.

⚠ In practice, construct $\widetilde{X}$ is difficult:
1. high-dimensionality, **2.** correlation.

$\Rightarrow$ Exchangeability violations lead to inflated FDR.

✓ **Diagnosing** non-exchangeability through 2 necessary conditions:
- C2ST between the original and knockoff distributions.
- Valid pairing between the original and knockoff individual.

✓ **Fixing** exchangeability by proposing new knockoff samplers:
- Autoregressive based on Candès et al. (2018).
- Parallel computationally-efficient approximate.

Index

References

Alexandre Blain, Angel Reyero Lobo, Julia Linhart, Bertrand Thirion, and Pierre Neuvial. When knockoffs fail: diagnosing and fixing non-exchangeability of knockoffs, 2025. URL `https://arxiv.org/abs/2407.06892`.

Emmanuel Candès, Yingying Fan, Lucas Janson, and Jinchi Lv. Panning for gold:'model-X'knockoffs for high dimensional controlled variable selection. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 80(3):551–577, 2018.

Yaniv Romano, Matteo Sesia, and Emmanuel Candès. Deep knockoffs. *Journal of the American Statistical Association*, 115(532): 1861–1872, 2020.

Zifan Zhu, Yingying Fan, Yinfei Kong, Jinchi Lv, and Fengzhu Sun. Deeplink: Deep learning inference using knockoffs with applications to genomics. *Proceedings of the National Academy of Sciences*, 118(36):e2104683118, 2021.

Thank You, Questions?

Sequential Conditional Independent Pairs

Algorithm: Sequential Conditional Independent Pairs (SCIP; Candès et al., 2018)

1: $j = 1$
2: **while** $j \leq p$ **do**
3: Sample $\widetilde{X}_j$ from $\mathscr{L}\left(X_j \mid X_{-j}, \widetilde{X}_{1:j-1}\right)$
4: $j = j + 1$
5: **end while**

Sequential generation of nonparametric knockoffs

Algorithm: Sequential generation of nonparametric knockoffs by learning to predict $\mathbf{X}_j$ from $(\mathbf{X}_{-j}, \widetilde{X}_{1:j-1})$ using a model f_j

Require: $f \in \mathscr{F}$

1: **for** $j \in [1, p]$ **do**

2: Fit a prediction model f_j on $((\mathbf{X}_{-j}, \widetilde{\mathbf{X}}_{1:j-1}), \mathbf{X}_j)$ ▷ Typically a Lasso

3: Compute the residual $\widehat{\boldsymbol{\varepsilon}}_j = \mathbf{X}_j - f_j((\mathbf{X}_{-j}, \widetilde{\mathbf{X}}_{1:j-1}))$

4: Sample $\widetilde{\mathbf{X}}_j = f_j((\mathbf{X}_{-j}, \widetilde{\mathbf{X}}_{1:j-1})) + \sigma(\widehat{\boldsymbol{\varepsilon}}_j)$

5: **end for**

6: **Return** $\widetilde{\mathbf{X}}_{\mathbf{1:p}}$

Parallel generation of nonparametric knockoffs

Algorithm: Parallel generation of nonparametric knockoffs by learning to predict X_j from X_{-j} using a model f_j

Require: $f \in \mathscr{F}$

for $j \in [1, p]$ **do**
    Fit a prediction model f_j on $(\mathbf{X}_{-j}, \mathbf{X}_j)$ ▷ Typically a Lasso
    Compute the residual $\hat{\boldsymbol{\varepsilon}}_j = \mathbf{X}_j - f_j(\mathbf{X}_{-j})$
end for
for $j \in [1, p]$ **do**
    Sample $\widetilde{\mathbf{X}}_j = f_j(\mathbf{X}_{-\mathbf{j}}) + \sigma(\hat{\boldsymbol{\varepsilon}}_j)$ ▷ σ is a permutation of $[\![n]\!]$
end for
Return $\widetilde{\mathbf{X}}_{1:p}$

Improper pairing

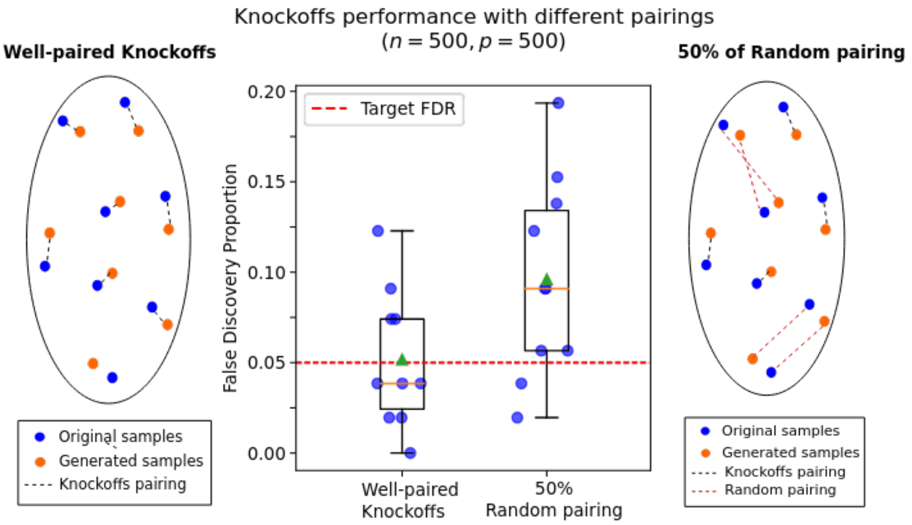


Figure 4: Exchangeability violation by improper sample pairing.