# Wavelet shrinkage using Bayesian FDR Methods

**Rodney V. Fonseca[a] (presenter) and Debashis Mondal[b]**

a - University of Campinas, Campinas, SP 13083, Brazil (contact: ra192588@ime.unicamp.br)
b - Oregon State University, Corvallis, OR, 97331, USA

## 1. Introduction

- In this work we study statistical signal denoising using wavelet methods.
- Consider that we have a set of observed values given by

$$y_i = g_i + \epsilon_i, \quad i = 1, \ldots, n,$$

where $g_i$ is the function we want to estimate and $\epsilon_1, \ldots, \epsilon_n$ are independent and identically distributed (iid) errors with distribution $N(0, \sigma^2)$, $\sigma^2 > 0$.
- The wavelet transform of these observations gives

$$(W\mathbf{y})_{jk} = d_{jk} = \theta_{jk} + \epsilon_{jk}, \quad j = 0, \ldots, J, \quad k = 0, \ldots, 2^j - 1,$$

where $\theta_{jk}$ are true wavelet coefficients and $\epsilon_{jk}$ are iid normal random variables.
- The denoising process can be regarded as tests over $\theta_{jk}$, but since there are multiple coefficients being evaluated at the same time, the tests must be carried out in an appropriate way to avoid a large number of null hypothesis being wrongly rejected (false discoveries).
- A way to deal with this is to use the False Discovery Rate (FDR) method [1], where p-values are ordered and then compared with a critical value such that the expected proportion of null hypothesis wrongly rejected is lower than some pre-specified value $\alpha$.
- In this work we extend the Bayesian FDR methods of [3] and [4] to perform wavelet shrinkage by considering the empirical Bayes method of [2].

## 2. Bayesian FDRs for wavelet shrinkage

- Considering the hypothesis $H_0 : \theta_{jk} = 0$ and $H_1 : \theta_{jk} \neq 0$, the main difference of the Bayesian FDRs compared to the original version is that in the former, posterior probabilities $p_0(d_{jk}) = P(H_0|d_{jk})$ are considered instead of p-values.
- The Bayesian FDR approaches of [4] (TIVG) and [3] (LJRV) were proposed with different frameworks, but overall they work with the same idea:
  1. Find the posterior probabilities of the null hypothesis $H_0 : \theta_{jk} = 0$, and order them such that $p_{0(1)} \leq p_{0(2)} \leq \ldots$.
  2. Fix a maximum expected FDR $\alpha$ and set the number of erroneously rejected hypothesis as $R = 1$.
  3. Increase $R$ by one and then compute the proportion of wrongly rejected $H_0$'s, say $bFDR$. Different ways of computing $bFDR$ are the approaches TIVG and LRVJ.
  4. If $bFDR > \alpha$, then the maximum posterior probability of rejection is $p_{0(R-1)}$. Otherwise, if $bFDR \leq \alpha$, return to step (3).
  5. Reject the null hypothesis of all coefficients such that $p_0(d_{jk}) < p_{0(R)}$.

## 3. Extensions of the FDR shrinkage methods

We combined the empirical Bayes of [2] with the FDR methods TIVG and LJRV, estimating the hyperparameters in a frequentist-like manner. The prior distribution of $\theta_{jk}$ we consider is

$$\begin{aligned} \theta_{jk}|\lambda_{jk}, \gamma_{jk} &\sim N(0, \sigma^2 c_j \gamma_{jk}/\lambda_{jk}), \quad c_j > 0, \sigma^2 > 0 \\ \lambda_{jk} &\sim h, \\ \gamma_{jk} &\sim Bernoulli(\omega_j), \quad \omega_j \in [0, 1], \end{aligned}$$

where $h$ is a scale distribution on $(0, \infty)$. The main difference of this method is that we estimate the values of $\sigma^2$, $c_j$ and $\omega_j$ by maximizing the log-likelihood given by

$$\mathcal{L}(\boldsymbol{\omega}, \mathbf{c}, \sigma^2) = \sum_j \sum_k \log\{\omega_{jk} f(d_{jk}|\gamma_{jk} = 1) + (1 - \omega_{jk}) f(d_{jk}|\gamma_{jk} = 0)\},$$

where $\boldsymbol{\omega} = (\omega_1, \ldots, \omega_J)^T$ and $\mathbf{c} = (c_1, \ldots, c_J)^T$. We shall consider that $\lambda_{jk} \sim Gamma(\nu/2, \nu/2)$, where $\nu > 0$ is some fixed value. Table 1 contains numerical evidence in favor of the empirical Bayes version of these two methods.

**Table 1:** *Average mean square error of estimators obtained applying wavelet shrinkage with the Bayesian FDR methods and their versions using empirical Bayes (EB) estimates.* $\alpha = 0.05$.

| Function | EB-LJRV | LJRV | EB-TIVG | TIVG |
|---|---|---|---|---|
| Blocks | 0.1888 | 0.2659 | 0.1866 | 0.2323 |
| Bumps | 0.0252 | 0.0329 | 0.0253 | 0.0292 |
| Doppler | 0.0028 | 0.0041 | 0.0028 | 0.0047 |
| Heavisine | 0.2350 | 0.2101 | 0.1828 | 0.2712 |

We also applied the FDR methods to perform wavelet shrinkage in a real data set, the atomic force microscopy (AFM) data analyzed by [3].

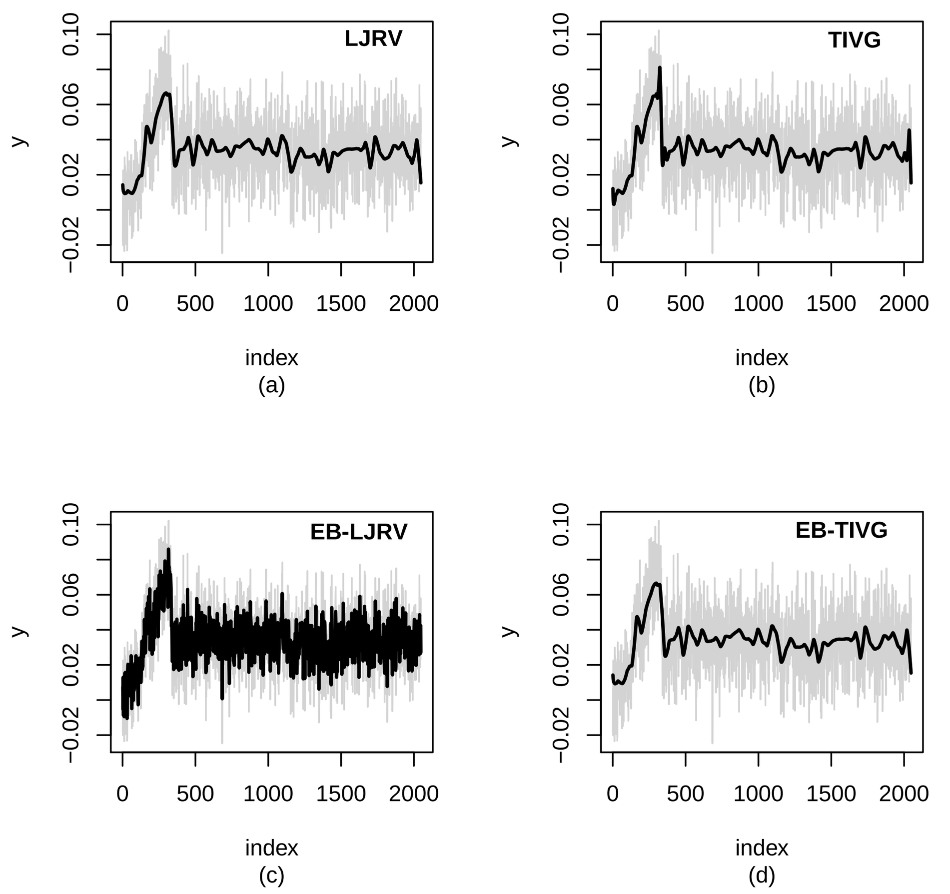


**Figure 1:** *Estimated functions (black line) using Bayesian FDR on wavelet shrinkage for the AFM data (gray line).* $\alpha = 0.05$.

### Future works

- Consider some correlation structure between the wavelet coefficients.
- Compound this method with a generalized linear model.

**Acknowledgments** We thank Brani Vidakovic for providing the AFM data. The research of Debashis Mondal was supported by National Science Foundation grant No. DMS 1519890 and Rodney V. Fonseca was supported by São Paulo Research Foundation (FAPESP) grant No. 2018/06874-6.

## References

[1] Benjamini, Y. and Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B*, 57:289–300.

[2] Clyde, M. and George, E. I. (2000). Flexible empirical Bayes estimation for wavelets. *Journal of the Royal Statistical Society. Series B*, 62(4):681–698.

[3] Lavrik, I., Jung, Y., Ruggeri, F. and Vidakovic, B. (2008). Bayesian false discovery rate wavelet shrinkage: theory and applications. *Communications in Statistics - Simulation and Computation*, 37(6):1086–1100. (FDR approach LJRV)

[4] Tadesse,M. G., Ibrahim, J. G., Vannucci, M., and Gentleman, R. (2005). Wavelet thresholding with Bayesian false discovery rate control. *Biometrics*, 61(1):25–35. (FDR approach TIVG)