CNR-IMATI Milano

December 14-16, 2004

WAVELETS AND SELF-SIMILARITY: THEORY AND APPLICATIONS

Lecture 2: Bayesian Modeling with Wavelets

P L A N

1. **Bayes' Paradigm**

2. **Bayesian Modeling in Wavelet Domain**

3. **BAMS**

4. **Other Bayesian Wavelet Procedures and Their Use**

The Reverend

• The *model* for a typical observation X conditional on unknown parameter θ is $f(x|\theta)$. As a function of θ, $f(x|\theta) = \ell(\theta)$ is called *likelihood.* The functional form of f is fully specified up to parameter θ.

• The parameter θ is supported by the parameter space Θ and is considered random variable. The random variable θ has distribution $\pi(\theta)$ that is called *prior.*

- If the prior for θ is specified up to parameter τ, $\pi(\theta|\tau)$, τ is called *hyperparameter*.
- The distribution $h(x, \theta) = f(x|\theta)\pi(\theta)$ is called *joint* distribution.
- Joint distribution can be factorized differently, $h(x, \theta) = \pi(\theta|x)m(x)$.
- The distribution $\pi(\theta|x)$ is called *posterior* distribution for θ, given x.
- The *marginal* distribution $m(x)$ can be obtained by integrating out the parameter θ from the joint distribution $h(x, \theta)$,

$$m(x) = \int_\Theta h(x, \theta) d\theta = \int_\Theta f(x|\theta)\pi(\theta) d\theta.$$

• Therefore, the posterior $\pi(\theta|x)$ can be expressed as

$$\pi(\theta|x) = \frac{h(x,\theta)}{m(x)} = \frac{f(x|\theta)\pi(\theta)}{m(x)} = \frac{f(x|\theta)\pi(\theta)}{\int_\Theta f(x|\theta)\pi(\theta)d\theta}.$$

• Suppose $Y \sim f(y|\theta)$ is to be observed. (Posterior) predictive distribution of Y, given observed $X = x$ is

$$f(y|x) = \int_\Theta f(y|\theta)\pi(\theta|x)d\theta.$$

The marginal distribution $m(y) = \int_\Theta f(y|\theta)\pi(\theta)d\theta$ is sometimes called prior predictive distribution.

- Wavelet Transforms $\Longrightarrow$ Wavelet Domains
- STATISTICAL MODELING in Wavelet Domains instead in the domains of original data.
  - Critically Sampled (Orthogonal) [The topic of Lecture 1]
  - Redundant (Stationary, Continuous, Frames)
  - What is Better: Parsimony or Redundancy?

Example: Minimal and Redundant Wavelet Domains.

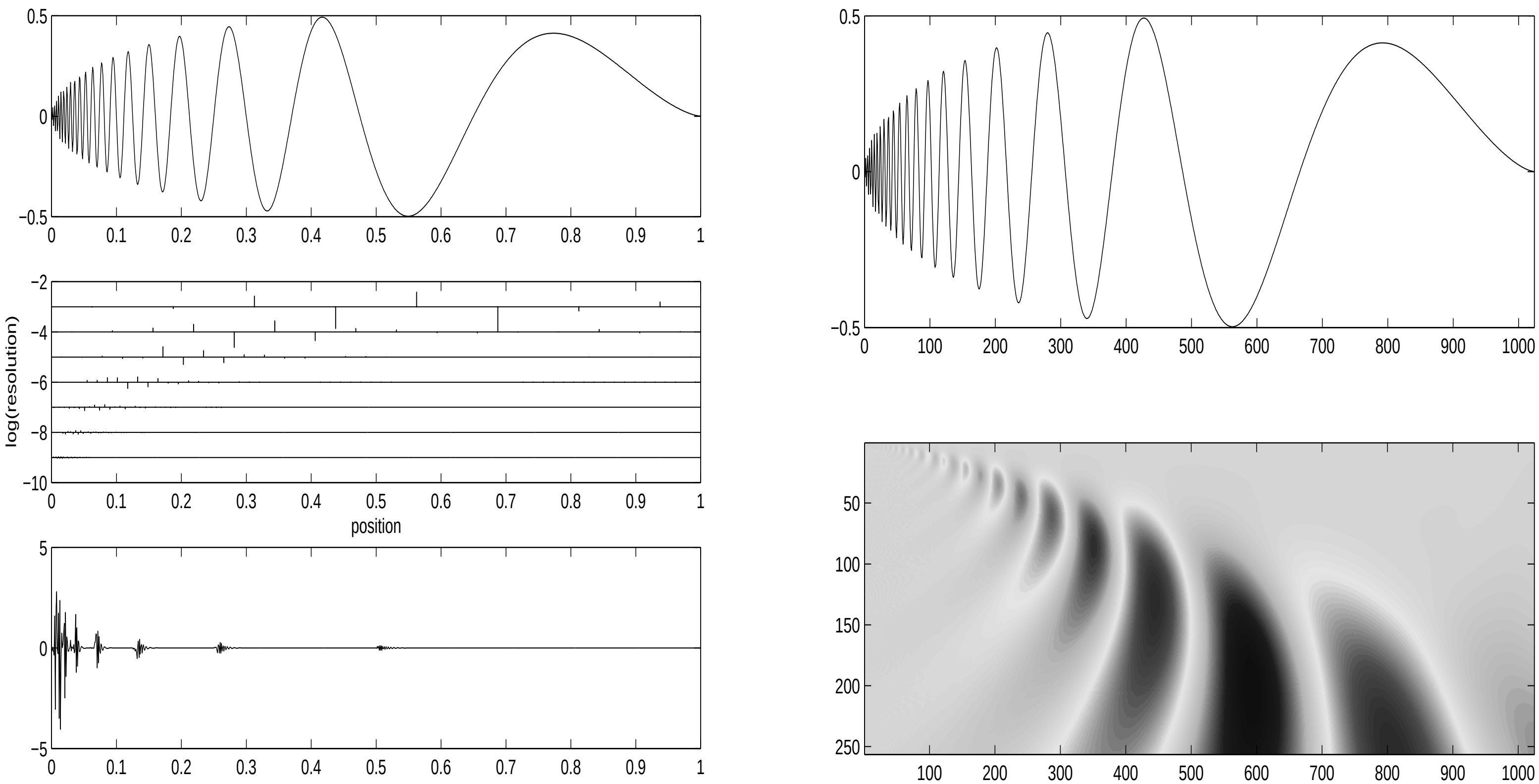


Figure 1: (left) DWT; (right) CWT of Doppler Signal

■ Wavelet-like and Related Transformations

■ Atomic Decompositions, Pursuits, General Time / Frequency Methods.

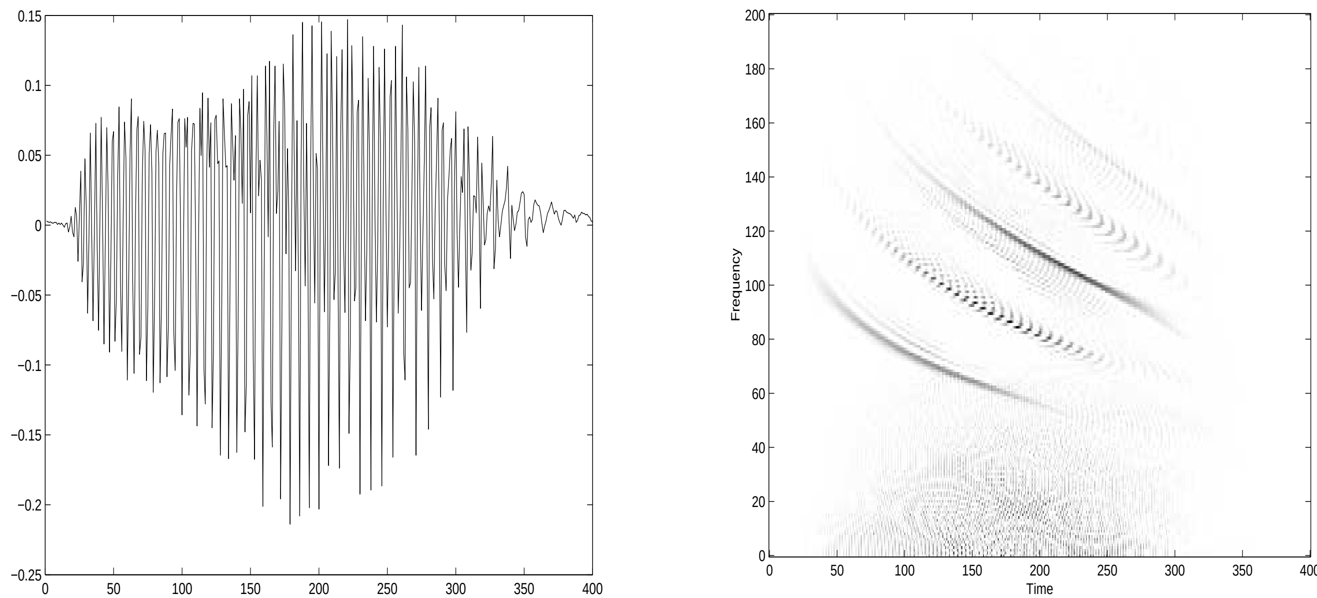


Figure 2: (left) Bat signal; (right) Its Wigner-Ville distribution

Model Based Wavelet Data Processing

$$\boxed{\textbf{DATA}} \xrightarrow{\mathbb{W}} \boxed{\textbf{Wavelet Coefficients: “Detail” \& “Smooth”}}$$

$$\|$$

$$\boxed{\textbf{Processed DATA}} \xleftarrow{\mathbb{W}^{-1}} \boxed{\textbf{Process (Detail) Coefficients}}$$

Process $\equiv$

- Shrink
- Transform
- Simulate New, Construct
- Resample
- Split

■ Various shrinkage/thresholding methods. (Start with Donoho, Johnstone,& Coauthors, the early 1990's)

■ Shrinkage induced by statistical modeling in the wavelet domain.

$$\boxed{\underset{\sim}{y} = \underset{\sim}{f} + \underset{\sim}{\epsilon} \qquad \xrightarrow{\mathbb{W}} \qquad \underset{\sim}{d} = \underset{\sim}{\theta} + \underset{\sim}{\epsilon}}$$

■ Estimate θ by $\hat{\theta}$ and set $\hat{f}$ as $\mathbb{W}^{-1}(\hat{\theta})$

Location model on d, $f(\mathbf{d} - \underset{\sim}{\theta}|\text{parameters})$

- Dimensionality of the model (Do not worry – wavelets decorrelate)
- Accounting for dependence (neighbors, parent-children), Blocking strategies (classical), Many Bayes solutions (MCMC, hidden MC's).
- **Model complexity/efficiency compromise.**
  - Simple models/Fast shrinkage ◦ Realistic?
  - Complex models ◦ Useful?

Our Focus: **Bayesianly Induced Shrinkage**

- Why Bayes? Prior information about regularity, self-similarity, periodicity, energy, and modality of the signal.
- $\hat{\theta}$ – Bayes Rule. Does it shrink? Well, not always.
- Prior Selection [Priors on θ, σ^2, hyper-priors]; What is our intuition about parameters in the wavelet-domain models?
- Need for Empirical Bayes.

- Complex Models: Often computationally prohibitive, Require MCMC and “MCMC-educated” users, Practitioners uneasy about.

- Simple Models: Efficient Shrinkage.

Normal Likelihood + Plug-in (hyper)Parameters $\rightarrow$ Observations of empirical wavelet coefficients disagree with the model!

Example: colorblue **BAMS** (**B**ayesian **A**daptive **M**ultiresolution **S**moother).

- A compromise between model reality/efficiency.

Model:

$$[d|\theta, \sigma^2] \sim \mathcal{N}(\theta, \sigma^2); \quad \sigma^2 \sim \mathcal{E}(\mu), \quad \mu > 0.$$

Marginal Model (σ^2 is out!):

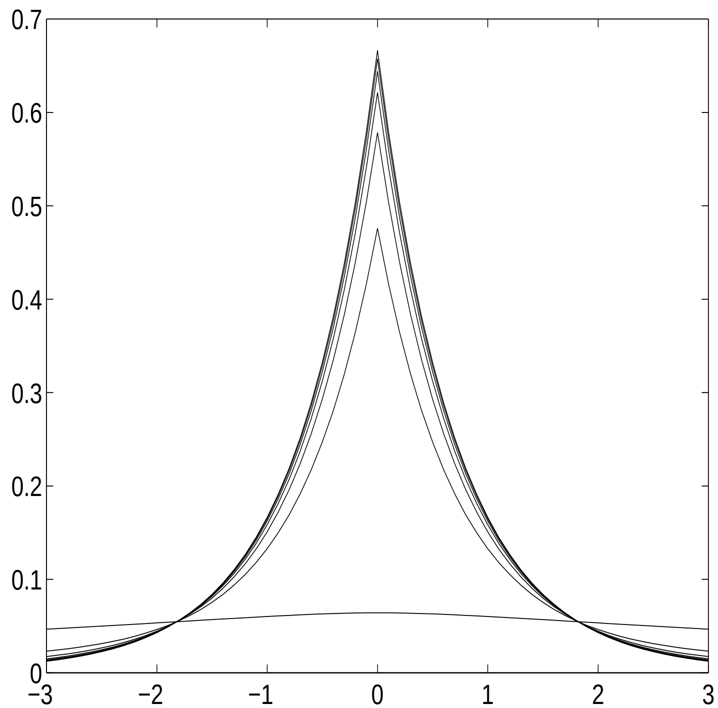
$$[d|\theta] \sim \mathcal{DE}(\theta, \frac{1}{\sqrt{2\mu}}); \quad [\frac{1}{2}\sqrt{2\mu}e^{-\sqrt{2\mu}|d-\theta|}].$$

Prior:

$$[\theta|\epsilon] \sim \epsilon\delta_0 + (1-\epsilon)\mathcal{DE}(0, \tau), \quad \epsilon = \epsilon(\text{multiresolution level}).$$

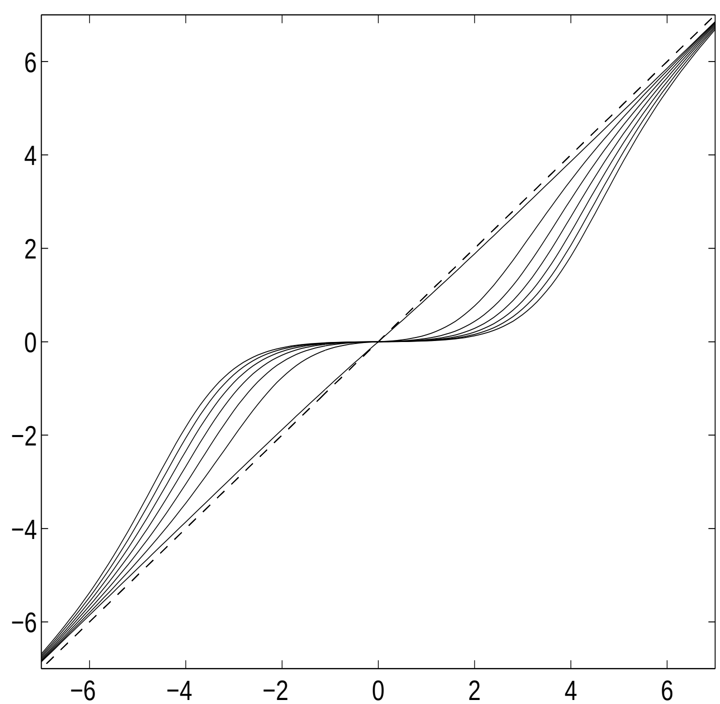
Marginal (Should agree with **d**):

$$d \sim m(d) = \epsilon \mathcal{DE}(0, \frac{1}{\sqrt{2\mu}}) + (1-\epsilon)\frac{\tau e^{-|d|/\tau} - \frac{1}{\sqrt{2\mu}} e^{-\sqrt{2\mu}|d|}}{2\tau^2 - 1/\mu}.$$



Bayes Rule:

$$\hat{\theta} = \delta(d) = \frac{\tau(\tau^2 - \sigma^2)de^{-d/\tau} + \tau^2/\mu(e^{-d\sqrt{2\mu}} - e^{-d/\tau})}{(\tau^2 - 1/(2\mu))(\tau e^{-d/\tau} - (1/\sqrt{2\mu})e^{-d\sqrt{2\mu}})}.$$



Specification of Hyperparameters:

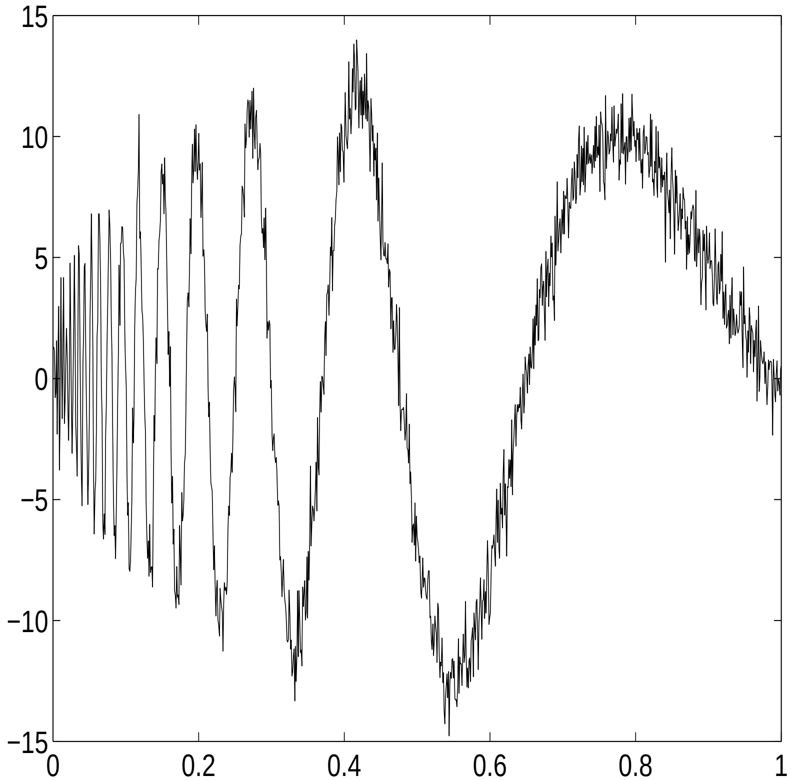
- $\mu = \frac{1}{\texttt{pseudos}}$

$\texttt{pseudos} = |Q_1 - Q_3|/C$, where Q_1 and Q_3 are the first and the third quartile of the finest level of details in the decomposition and $1.3 \leq C \leq 1.5$.

- $\epsilon(j) = 1 - \frac{1}{(j - \texttt{coarsest} + 1)^{\gamma}}, \ \gamma = \frac{3}{2}$.
- $\tau = \sqrt{\sigma_d^2 - \frac{1}{\mu}}$. (Information on selfsimilarity via τ.)

FUNCTION	VisuShrink	SureShrink
BLOCKS	0.6840 (0.0719 + 0.6122)	0.2225 (0.1369 + 0.0856)
BUMPS	1.5707 (0.1165 + 1.4543)	0.6827 (0.2660 + 0.4167)
DOPPLER	0.4850 (0.0523 + 0.4327)	0.2285 (0.0946 + 0.1340)
HEAVISINE	0.1204 (0.0339 + 0.0864)	0.0949 (0.0416 + 0.0534)

FUNCTION	ABWS (CKM Chicago '97)	BAMS
BLOCKS	0.0995 (0.0874 + 0.0121)	0.1107 (0.0965 + 0.0142)
BUMPS	0.3495 (0.2228 + 0.1267)	0.3404 (0.1976 + 0.1428)
DOPPLER	0.1646 (0.1006 + 0.0640)	0.1482 (0.0899 + 0.0584)
HEAVISINE	0.0874 (0.0442 + 0.0433)	0.0815 (0.0511 + 0.0304)



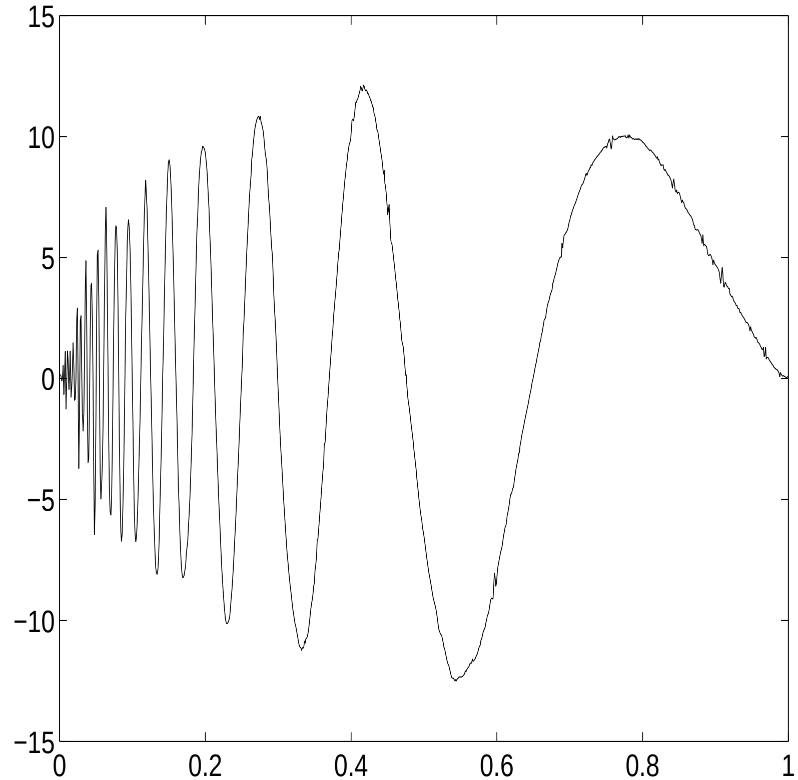


Figure 3: Heavisine signal: $n = 1024$, SNR=7.

Matlab program implementing BAMS:

`http://www.isye.gatech.edu/~brani/wavelet.html`

From: Ruggeri, F. and Vidakovic, B. (2005). Bayesian Modeling in the Wavelet Domain, To appear as Chapter in Handbook of Statistics Vol. 25, on Bayesian Statistics (C.R. Rao and Dipak Dey).

Some other Bayesian Procedures + Annotated Bibliography

Chipman, Kolaczyk, and McCulloch (1997) based on the stochastic search variable selection (SSVS)

$$[d|\theta] \sim \mathcal{N}(\theta, \sigma^2).$$

$$[\theta|\gamma_j] \sim \gamma_j \mathcal{N}(0, (c_j\tau_j)^2) + (1-\gamma_j)\mathcal{N}(0, \tau_j^2),$$

where

$$[\gamma_j] \sim Ber(p_j).$$

$$\delta(d) = \left[P(\gamma_j = 1|d) \frac{(c_j \tau_j)^2}{\sigma^2 + (c_j \tau_j)^2} + P(\gamma_j = 0|d) \frac{\tau_j^2}{\sigma^2 + \tau_j^2} \right] \, d,$$

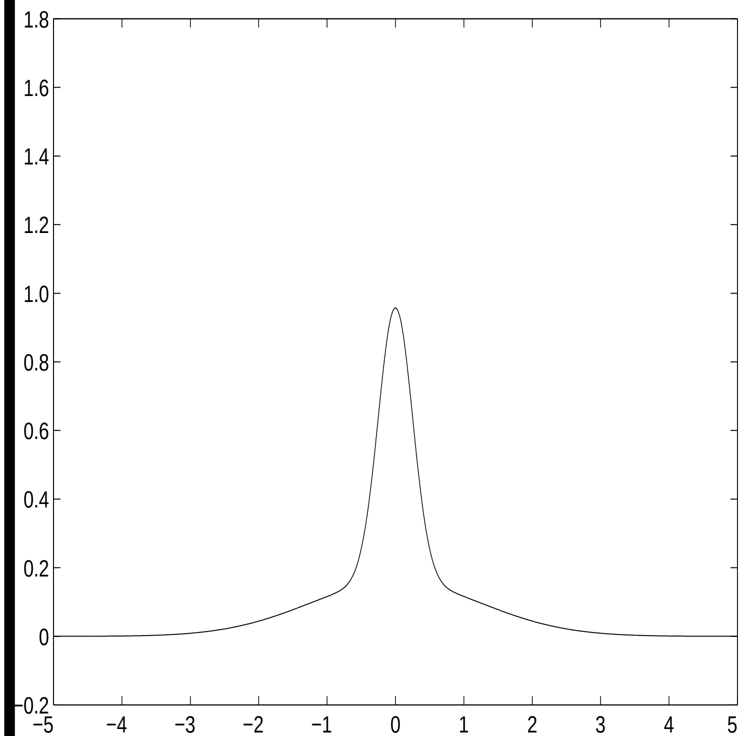
where

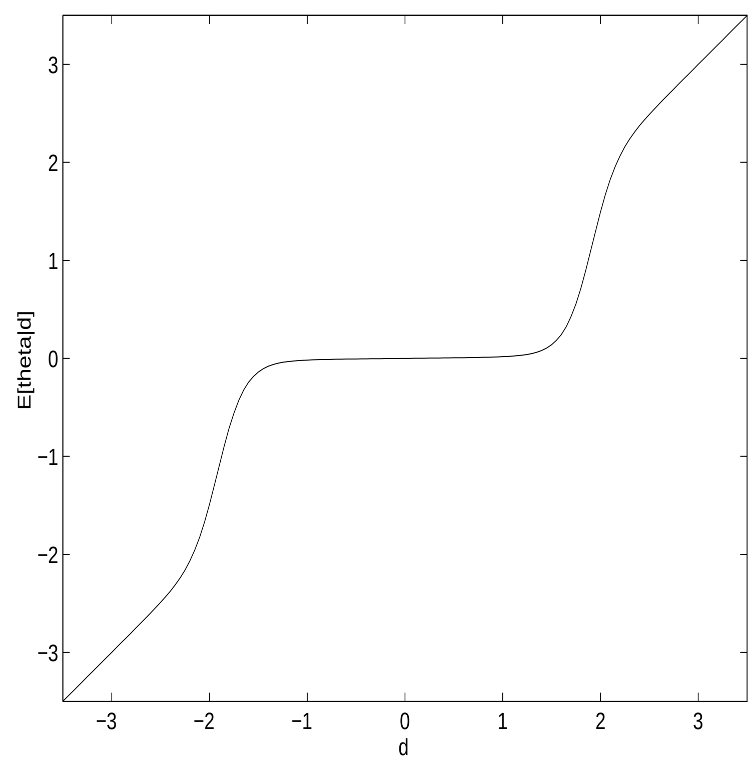
$$P(\gamma_j = 1|d) = \frac{p_j \pi(d|\gamma_j = 1)}{(1 - p_j)\pi(d|\gamma_j = 0)}$$

and

$$\pi(d|\gamma_j = 1) = \phi_{\sigma^2 + (c_j \tau_j)^2}(d) \;\; \text{and} \;\; \pi(d|\gamma_j = 0)\phi_{\sigma^2 + \tau_j^2}(d).$$

De Canditiis and Vidakovic (2004) extend the ABWS method to multivariate case (block shrinkage) and unknown σ^2 using a mixture of normal-inverse Gamma priors.


1.8
1.6
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0
-0.2
−5
−4
−3
−2
−1
0
1
2
3
4
5


E[theta|d]
3
2
1
0
−1
−2
−3
d

Clyde, Parmigiani, and Vidakovic (1998): *Biometrika*

$$\begin{aligned}[\theta|\gamma_j, \sigma^2] &\sim \mathcal{N}(0, (1-\gamma_j) + \gamma_j c_j \sigma^2), \\ \left[\lambda\nu/\sigma^2\right] &\sim \chi^2_\nu,\end{aligned}$$

where λ and ν are fixed hyperparameters.

$$\gamma_j \sim \mathcal{B}er(p_j)$$

The posterior mean of $\boldsymbol{\theta}|\boldsymbol{\gamma}$ is

$$E(\boldsymbol{\theta}|\boldsymbol{d}, \boldsymbol{\gamma}) = \Gamma(I_n + C^{-1})^{-1}\boldsymbol{d},$$

where Γ and C are diagonal matrices with γ_{jk} and c_{jk}, respectively, on the diagonal and 0 elsewhere.

The posterior mean is obtained by averaging over all models. Model averaging leads to a multiple shrinkage estimator of θ:

$$E(\boldsymbol{\theta}|\boldsymbol{d}) = \sum_{\boldsymbol{\gamma}} \pi(\boldsymbol{\gamma}|\boldsymbol{d})\Gamma\left(I_n + C^{-1}\right)^{-1}\boldsymbol{d},$$

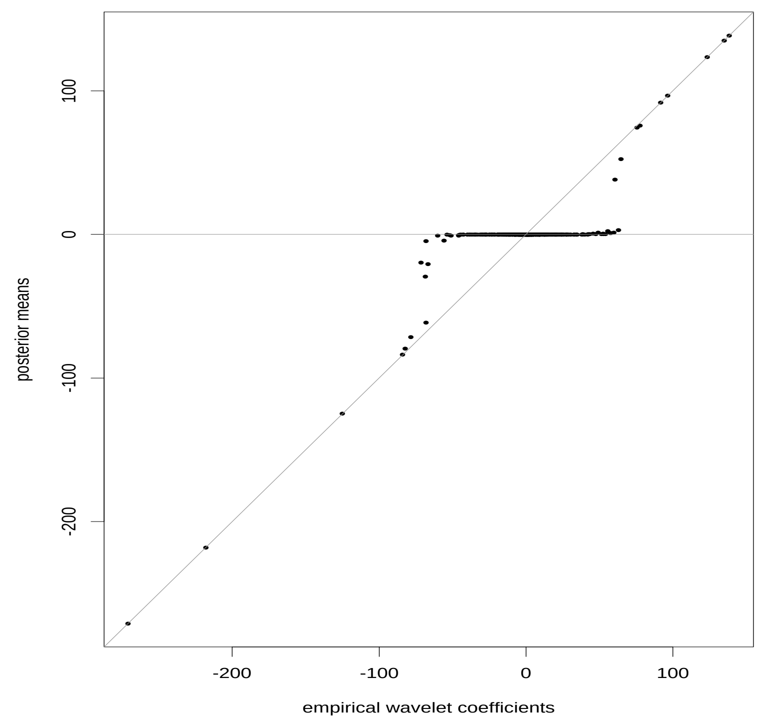
where $\pi(\boldsymbol{\gamma}|\boldsymbol{d})$ is the posterior probability of a particular subset $\boldsymbol{\gamma}$.

The approximation can be achieved by either conditioning on σ (plug-in approach) or by assuming independence of the elements in $\boldsymbol{\gamma}$.

$$\begin{aligned}
\pi(\boldsymbol{\gamma}|\boldsymbol{d}) &\approx \prod_{j,k} \rho_{jk}^{\gamma_{jk}}(1-\rho_{jk})^{1-\gamma_{jk}} \\
\rho_{jk}(\boldsymbol{d},\sigma) &= \frac{a_{jk}(\boldsymbol{d},\sigma)}{1+a_{jk}(\boldsymbol{d},\sigma)},
\end{aligned}$$

where

$$\begin{aligned}
a_{jk}(\boldsymbol{d},\sigma) &= \frac{p_{jk}}{1-p_{jk}}(1+c_{jk})^{-1/2}\cdot\exp\left\{\frac{1}{2}\frac{S_{jk}^2}{\sigma^2}\right\} \\
S_{jk}^2 &= d_{jk}^2/(1+c_{jk}^{-1}).
\end{aligned}$$



posterior means
100
0
-100
-200
-200
-100
0
100
empirical wavelet coefficients

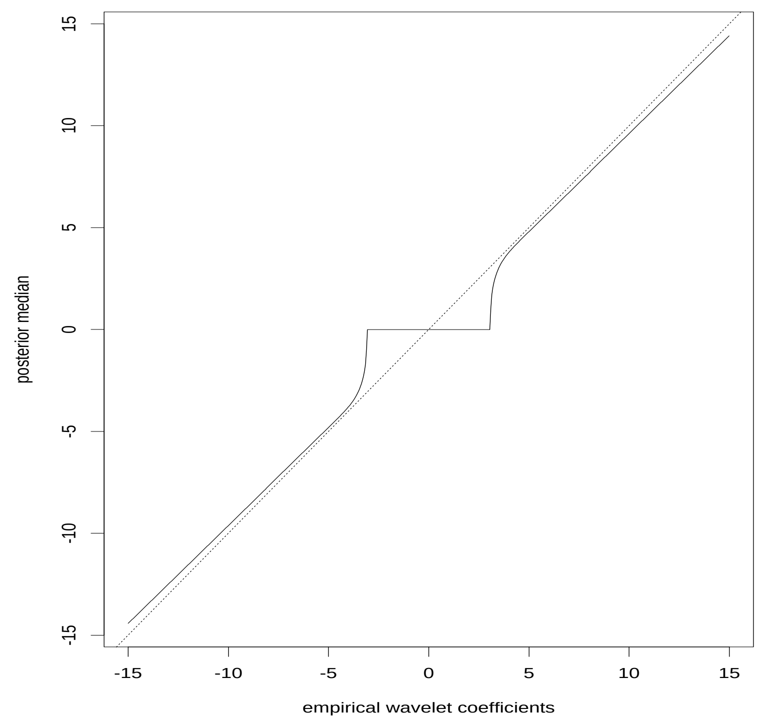
Abramovich, Sapatinas, and Silverman (1998) use weighted absolute error loss and show that for a prior on θ

$$[\theta] \sim \pi_j \mathcal{N}(0, \tau_j^2) + (1 - \pi_j)\delta(0)$$

and normal $\mathcal{N}(\theta, \sigma^2)$ likelihood, the posterior median is

$$Med(\theta|d) = \text{sign}(d) \max(0, \zeta). \tag{1}$$

$$\begin{aligned} \zeta &= \frac{\tau_j^2}{\sigma^2 + \tau_j^2}|d| - \frac{\tau_j \sigma}{\sqrt{\sigma^2 + \tau_j^2}} \Phi^{-1}\left(\frac{1 + \min(\omega, 1)}{2}\right), \text{ and} \\ \omega &= \frac{1 - \pi_j}{\pi_j} \frac{\sqrt{\tau_j^2 + \sigma^2}}{\sigma} \exp\left\{-\frac{\tau_j^2 d^2}{2\sigma^2(\tau_j^2 + \sigma^2)}\right\}. \end{aligned}$$


posterior median
empirical wavelet coefficients
-15
-10
-5
0
5
10
15

REFERENCES

Abramovich F., Amato U., and Angelini C. (2004). On optimality of Bayesian wavelet estimators. *Scandinavian Journal of Statistics*, 31, 217–234.

Abramovich, F., Besbeas, P., and Sapatinas, T. (2002). Empirical Bayes approach to block wavelet function estimation. *Comput. Statist. Data Anal.*, 39, 435–451.

Abramovich, F., Sapatinas, T. and Silverman, B. W. (1998). Wavelet thresholding via Bayesian approach. *Journal of the Royal Statistical Society, Ser. B*, 60, 725–749.

Ambler, G. K. and Silverman, B. W. (2004a). Perfect simulation of spatial point processes using dominated coupling from the past with application to a multiscale area-interaction point process. Manuscript, University of Bristol.

Ambler, G. K. and Silverman, B. W. (2004b). Perfect simulation for wavelet thresholding with correlated coefficients. Technical Report 04:01, University of Bristol.

Angelini, C. and Sapatinas, T. (2004). Empirical Bayes approach to wavelet regression using ϵ-contaminated priors. *Journal of Statistical Computation and Simulation*, 74, 741–764.

Angelini, C. and Vidakovic, B. (2004). Γ-Minimax Wavelet Shrinkage: A Robust Incorporation of Information about Energy of a Signal in Denoising Applications. *Statistica Sinica,* 14, 103–125.

Antoniadis, A., Bigot, J., and Sapatinas, T. (2001). Wavelet estimators in nonparametric regression: A comparative simulation study. *Journal of Statistical Software*, 6, 1–83.

Berliner, L. M., Wikle, C. K., and Milliff, R. F. (1999). Multiresolution wavelet analyses in hierarchical Bayesian turbulence models. Bayesian Inference in Wavelet Based Models, In *Bayesian Inference in Wavelet Based Models*, eds. P. Müller and B. Vidakovic, Lecture Notes in Statistics, vol. 141, 341–359. Springer-Verlag, New York.

Brown, P. J., Fearn, T., and Vannucci, M. (2001). Bayesian wavelet regression on curves with application to a spectroscopic calibration problem. *Journal of the American Statistical Association*, 96, 398–408.

Chambolle, A., DeVore, R. A., Lee, N-Y, and Lucier, B. J. (1998). Nonlinear wavelet image processing: variational problems, compression and noise removal through wavelet shrinkage, *IEEE Trans. Image Processing*, 7, 319–335.

Chang, W. and Vidakovic, B. (2002). Wavelet estimation of a baseline signal from repeated noisy measurements by vertical block shrinkage. *Computational Statistics and Data Analysis*, 40, 317–328.

Chipman, H., McCulloch, R., and Kolaczyk, E. (1997). Adaptive Bayesian Wavelet Shrinkage. *Journal of the American Statistical Association*, 92, 1413–1421.

Clyde, M. and George, E. (1999). Empirical Bayes Estimation in Wavelet Nonparametric Regression. In *Bayesian Inference in Wavelet Based Models*, eds. P. Müller and B. Vidakovic, Lecture Notes in Statistics, vol. 141, 309–322. Springer-Verlag, New York.

Clyde, M. and George, E. (2000). Flexible Empirical Bayes Estimation for Wavelets, *Journal of the Royal Statistical Society, Ser. B*, 62, 681–698.

Clyde, M., Parmigiani, G., and Vidakovic, B. (1998). Multiple shrinkage and subset selection in wavelets, *Biometrika* , 85, 391–402.

Crouse, M., Nowak, R., and Baraniuk, R. (1998). Wavelet-based statistical signal processing using hidden Markov models. *IEEE Trans. Signal Processing*, 46, 886-902.

De Canditiis, D. and Vidakovic, B. (2004). Wavelet Bayesian Block Shrinkage via Mixtures of Normal-Inverse Gamma Priors. *Journal of Computational and Graphical Statistics*, 13, 383–398.

Fadili, M. J. and E.T. Bullmore, E. T. (2004). A comparative evaluation of wavelet-based methods for hypothesis testing of brain activation maps. *NeuroImage*, 23, 1112-1128.

Figueiredo, M. and Robert Nowak, R. (2001). Wavelet-Based Image Estimation: An empirical Bayes approach using Jeffreys' noninformative prior. *IEEE Transactions on Image Processing*, 10, 1322–1331.

Huang, H.-C. and Cressie, N. (1999). Empirical Bayesian spatial prediction using wavelets. In *Bayesian Inference in Wavelet Based Models*, eds. P. Müller and B. Vidakovic, Lecture Notes in

Statistics, vol. 141, 203–222. Springer-Verlag, New York.

Huang, H.-C. and Cressie, N. (2000). Deterministic/stochastic wavelet decomposition for recovery of signal from noisy data. *Technometrics*, 42, 262–276.

Huang, S. Y. (2002). On a Bayesian aspect for soft wavelet shrinkage estimation under an asymmetric linex loss. *Statistics and Probability Letters*, 56, 171–175.

Huang, S. Y. and Lu, H. S. (2000). Bayesian wavelet shrinkage for nonparametric mixed-effects models. *Statistica Sinica*, 10, 1021–1040.

Kohn, R., Marron, J.S., and Yau, P. (2000). Wavelet estimation using Bayesian basis selection and basis averaging, *Statistica Sinica*, 10, 109 – 128.

Kolaczyk, E. D. (1999). Bayesian multi-scale models for Poisson processes. *Journal of the American Statistical Association*, 94, 920–933.

Leporini, D., and Pesquet, J.-C. (1998). Wavelet thresholding for a wide class of noise distributions, EUSIPCO'98, Rhodes, Greece, 993–996.

Leporini, D., and Pesquet, J.-C. (1999). Bayesian wavelet denoising: Besov priors and non-gaussian noises, *Signal Processing*, 81, 55–67.

Leporini, D., Pesquet, J.-C., Krim, H. (1999). Best basis representations with prior statistical models. In *Bayesian Inference in Wavelet Based Models*, eds. P. Müller and B. Vidakovic, Lecture Notes in Statistics, vol. 141, 109–113. Springer-Verlag, New York.

Lina, J-M., and MacGibbon, B. (1997). Non-Linear shrinkage estimation with complex Daubechies wavelets. Proceedings of SPIE, Wavelet Applications in Signal and Image Processing V, vol. 3169, 67–79.

Morris, J.S., Vannucci, M., Brown, P.J. and Carroll, R.J. (2003). Wavelet-based nonparametric modeling of hierarchical functions in colon carcinogenesis (with discussion). *Journal of the American Statistical Association*, 98, 573–597.

Müller, P. and Vidakovic, B. (1999a). Bayesian inference with wavelets: Density estimation. *Journal of Computational and Graphical Statistics,* **7**, 456–468.

Müller, P. and Vidakovic, B. (1999b). MCMC methods in wavelet shrinkage: Non-equally spaced regression, density and spectral density estimation. In *Bayesian Inference in Wavelet Based Models,* eds. P. Müller and B. Vidakovic, Lecture Notes in Statistics, vol. 141, 187–202. Springer-Verlag, New York.

Müller, P. and Vidakovic, B. (Editors) (1999c). *Bayesian Inference in Wavelet Based Models.* Lecture Notes in Statistics **141**, Springer-Verlag, New York.

Nowak, R. and Kolaczyk, E. (2000). A Bayesian multiscale framework for Poisson inverse problems. *IEEE Transactions on Information Theory, Special Issue on Information-Theoretic Imaging*, 46, 1811–1825.

Ogden, R. T. and Lynch, J. D. (1999). Bayesian analysis of change-point models. In *Bayesian Inference in Wavelet Based Models,* eds. P. Müller and B. Vidakovic, Lecture Notes in Statistics, vol. 141, 67–82. Springer-Verlag, New York.

Pensky, M. and Sapatinas, T. (2004). Frequentist optimality of Bayes factor thresholding estimators in wavelet regression models. Technical Report at University of Cyprus.

Pesquet, J., Krim, H., Leporini, D., and Hamman, E. (1996). Bayesian approach to best basis selection. *IEEE International Conference on Acoustics, Speech, and Signal Processing*, 5, 2634–2637.

Portilla, J., Strela, V., Wainwright, M., and Simoncelli, E. (2003). Image denoising using scale mixtures of Gaussians in the wavelet domain. *IEEE Transactions on Image Processing*, 12, 1338–1351.

Ruggeri, F. (1999). Robust Bayesian and Bayesian decision theoretic wavelet shrinkage. In *Bayesian Inference in Wavelet Based Models,* eds. P. Müller and B. Vidakovic, Lecture Notes in Statistics, vol. 141, 139–154. Springer-Verlag, New York.

Ruggeri, F. and Vidakovic, B. (1999). A Bayesian decision theoretic approach to the choice of thresholding parameter. *Statistica Sinica*, 9, 183–197.

Simoncelli, E. (1999). Bayesian denoising of visual images in the wavelet domain. In *Bayesian Inference in Wavelet Based Models,* eds. P. Müller and B. Vidakovic, Lecture Notes in Statistics, vol. 141, 291–308. Springer-Verlag, New York.

Vannucci, M., Brown, P.J., and Fearn, T. (2001). Predictor selection for model averaging. Bayesian methods with applications to science, policy and official statistics. (Eds E. I. George and P. Nanopoulos), Eurostat: Luxemburg, 553–562.

Vannucci, M., Brown, P.J., and Fearn, T. (2003). A decision theoretical approach to wavelet regression on curves with a high number of regressors. *Journal of Statistical Planning and Inference*, 112, 195–212.

Vannucci, M. and Corradi, F. (1999). Covariance structure of wavelet coefficients: Theory and models in a Bayesian perspective. *Journal of the Royal Statistical Society, Ser. B*, 61, 971–986.

Vidakovic, B. (1998a). Nonlinear wavelet shrinkage with Bayes rules and Bayes factors. *Journal of the American Statistical Association*, 93, 173–179.

Vidakovic, B. (1998b). Wavelet-based nonparametric Bayes methods. In *Practical Nonparametric and Semiparametric Bayesian Statistics,* eds. D. Dey, P. Müller and D. Sinha, Lecture Notes in Statistics, vol. 133, 133–155. Springer-Verlag, New York.

Vidakovic, B. (1999). *Statistical Modeling by Wavelets.* John Wiley & Sons, Inc., New York, 384 pp.

Vidakovic, B. and Bielza Lozoya, C. (1998). Time-adaptive wavelet denoising. *IEEE Transactions on Signal Processing*, **46**, 2549–2554.

Vidakovic, B. and Müller, P. (1999). An introduction to wavelets. In *Bayesian Inference in Wavelet Based Models,* eds. P. Müller and B. Vidakovic, Lecture Notes in Statistics, vol. 141, 1–18. Springer-Verlag, New York.

Vidakovic, B. and Ruggeri, F. (1999). Expansion estimation by

Bayes rules. *Journal of Statistical Planning and Inference*, 79, 223–235.

Vidakovic, B. and Ruggeri, F. (2001). BAMS method: Theory and simulations. *Sankhyā, Series B*, 63, 234–249. (Matlab code is available from the Wavelet Denoising software written by Antoniadis, Bigot, and Sapatinas, 2001)

Wikle, C. K., Milliff, R. F., and Large, W. G. (1999). Observed wavenumber spectra from 1 to 10,000 km of near-surface winds during the TOGA COARE IOP. *Journal of Atmospheric Science*, 56, 2222–2231.

Wikle, C. K., Milliff, R. F., Nychka, D., and Berliner, L. M. (2001). Spatiotemporal hierarchical Bayesian modeling: Tropical ocean surface winds. *Journal of the American Statistical Association*, 96, 382–397.