Bayesian Estimation of Bessel K Form Random Vectors in AWGN

Pavel A. Khazron and Ivan W. Selesnick, *Member, IEEE*

***Abstract*—We present new Bayesian estimators for spherically-contoured Bessel K form (BKF) random vectors in additive white Gaussian noise (AWGN). The derivations are an extension of existing results for the scalar BKF and multivariate Laplace (MLAP) densities. MAP and MMSE estimators are derived. We show that the MMSE estimator can be written in exact form in terms of the generalized incomplete Gamma function. Computationally efficient approximations are given. We compare the proposed exact and approximate MMSE estimators with recent results using the BKF density, both in terms of the shrinkage rules and the associated mean-square error.**

***Index Terms*—Bayesian estimation, Bessel K form density, MAP estimator, MMSE estimator, wavelet denoising.**

I. Introduction

Wavelet coefficients of natural images and discrete Fourier transform (DFT) coefficients of short speech segments usually possess peaked, symmetric, zero-mean distributions, with heavier than Gaussian tails. The peaked behavior is roughly captured by the Laplace density. More accurate densities better model wavelet histograms near the origin, and in the tails. Some such models either require more degrees of freedom (two or more parameters), for example [1], [14], [16], or entail mixtures of densities [7], [8], [11]. A review of priors used for wavelet-domain modeling is provided by [3]. The Bessel K form (BKF) density has been used in [4], [9], [12], and [22].

Multivariate modeling offers advantages over scalar modeling, because dependencies and/or correlations between coefficients can be captured. Such models can be applied to neighborhoods of coefficients, and they are effective in applications [6], [18], [20], [21].

In this letter, we derive MAP and MMSE estimators for Bessel K form random vectors in independent additive white Gaussian noise (AWGN). The multivariate Bessel K form (MBKF) density is a two-parameter extension of the multivariate Laplace (MLAP) density we considered in [19]. We draw from and extend the results of [19]. This letter also builds upon results in [4], [9], and [17], where MAP and MMSE estimators are derived in the scalar case.

II. Multivariate Bessel K Form (MBKF) Density

The spherically-contoured zero-mean d-dimensional BKF density [10] can be written as

$$p_{\mathbf{S}}(\mathbf{s}) = \frac{2}{(2\pi c)^{d/2}\Gamma(p)} \left(\frac{\sqrt{2c}}{\|\mathbf{s}\|}\right)^{d/2-p} K_{d/2-p}\left(\sqrt{\frac{2}{c}}\|\mathbf{s}\|\right) \tag{1}$$

Manuscript received June 15, 2007; revised November 12, 2007. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Andreas Jakobsson.

The authors are with the Department of Electrical and Computer Engineering, Polytechnic University, Brooklyn, NY 11201 USA (e-mail: pkhazr01@students.poly.edu; selesi@poly.edu).

Digital Object Identifier 10.1109/LSP.2007.914927

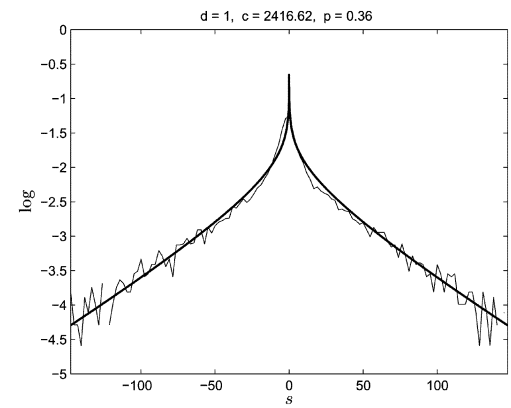

Fig. 1. Example wavelet subband histogram with a fitted scalar BKF density (1), illustrated on a $\log$ scale.

with $\mathbf{s} \in \mathbb{R}^d$, and where $K_\lambda(u)$ is the modified Bessel function of the second kind [1], and c and p are the scale and shape parameters. The marginals of (1) are scalar BKF densities. Parameter p allows a tradeoff between heaviness of the tails and sharpness of the mode. Although in general $p > 0$, wavelet data from natural images typically exhibit $0 < p \leq 1$[9], [22], as in Fig. 1. For $p = 1$, the MBKF density reduces to the MLAP density in [19, eq. (7)], while when $d = 1$, the MBKF reduces to the scalar BKF in [12, eq. (6)].

A BKF random vector $\mathbf{S}$ can be represented as a Gaussian scale mixture (GSM) [10], [15]

$$\mathbf{S} = \sqrt{Z}\mathbf{X} = A\mathbf{X} \tag{2}$$

where $\mathbf{X}$ is a d-dimensional zero-mean Gaussian random vector, $N_d(0, \mathbf{\Sigma})$. We take $\mathbf{\Sigma} = c\mathbf{I}$, which gives the spherically-contoured BKF. Z is a scalar Gamma random variable with parameter p. Consequently, the density of $\mathbf{S}$ can be expressed as the integral

$$p_{\mathbf{S}}(\mathbf{s}) = \int_0^\infty p_A(a) \frac{1}{a^d} p_{\mathbf{X}}\left(\frac{\mathbf{s}}{a}\right) da \tag{3}$$

with

$$p_A(a) = 2a\frac{a^{2p-2}e^{-a^2}}{\Gamma(p)}, \quad a > 0 \tag{4}$$

a form that proves useful below.

III. BKF Vectors in AWGN

We consider a d-dimensional BKF random vector $\mathbf{S}$ in independent AWGN

$$\mathbf{Y} = \mathbf{S} + \mathbf{N}. \tag{5}$$

The density of $\mathbf{Y}$ can be written as a convolution

$$p_{\mathbf{Y}}(\mathbf{y}) = \int_{\mathbb{R}^d} p_{\mathbf{S}}(\mathbf{y}-\mathbf{t}) p_{\mathbf{N}}(\mathbf{t}) d\mathbf{t} \tag{6}$$

where $p_{\mathbf{S}}(\mathbf{s})$ is given in (1), and $p_{\mathbf{N}}(\mathbf{n}) = N_d(0, \sigma_n^2 \mathbf{I})$. Using the GSM property in (3), this convolution can be written as

$$p_{\mathbf{Y}}(\mathbf{y}) = \int_0^\infty p_A(a) \left[\int_{\mathbb{R}^d} \frac{1}{a^d} p_{\mathbf{X}}\left(\frac{\mathbf{s}}{a}\right) p_{\mathbf{N}}(\mathbf{s}-\mathbf{y}) d\mathbf{s} \right] da. \tag{7}$$

The term inside brackets is a d-dimensional convolution of the Gaussians $N_d(0, \sigma_n^2 \mathbf{I})$ and $N_d(0, ca^2 \mathbf{I})$. The result is a Gaussian $N_d(0, [\sigma_n^2 + ca^2] \mathbf{I})$. Rewriting, changing the variable of integration, and simplifying gives

$$\begin{aligned} p_{\mathbf{Y}}(\mathbf{y}) = \frac{\exp\left(\frac{\sigma_n^2}{c}\right)}{(2\pi)^{d/2}\Gamma(p)} &\times \int_{\sigma_n^2/c}^\infty \left(t - \frac{\sigma_n^2}{c}\right)^{p-1} \frac{1}{(ct)^{d/2}} \\ &\times \exp\left(-t - \frac{\|\mathbf{y}\|^2}{2ct}\right) dt. \end{aligned} \tag{8}$$

The factor $(t - \sigma_n^2/c)^{p-1}$ can be expanded using the infinite form of the Binomial theorem as

$$\left(t - \frac{\sigma_n^2}{c}\right)^{p-1} = \sum_{j=0}^\infty \left(-\frac{\sigma_n^2}{c}\right)^j \frac{(p-1)_j}{j!} t^{p-1-j} \tag{9}$$

with convergence for $t > \sigma_n^2/c$, and where $(x)_k$ is defined in (27). Then (8) can be expressed using the generalized incomplete Gamma function $\Gamma(\tau, x; b)$ [5] in (26) as

$$\begin{aligned} p_{\mathbf{Y}}(\mathbf{y}) &= \frac{\exp\left(\frac{\sigma_n^2}{c}\right)}{(2\pi c)^{d/2}\Gamma(p)} \\ &\times \sum_{j=0}^\infty \left(-\frac{\sigma_n^2}{c}\right)^j \frac{(p-1)_j}{j!} \Gamma\left(p - \frac{d}{2} - j, \frac{\sigma_n^2}{c}; \frac{\|\mathbf{y}\|^2}{2c}\right). \end{aligned} \tag{10}$$

IV. MAP Estimator

With reference to (5), MAP estimation corresponds to the maximization problem

$$\hat{\mathbf{s}}(\mathbf{y}) = \arg\max_{\mathbf{s}} p_{\mathbf{S}|\mathbf{Y}}(\mathbf{s}|\mathbf{y}) \tag{11}$$

which is equivalent to

$$\hat{\mathbf{s}}(\mathbf{y}) = \arg\max_{\mathbf{s}} \left(\log[p_{\mathbf{N}}(\mathbf{y}-\mathbf{s})] + \log[p_{\mathbf{S}}(\mathbf{s})]\right). \tag{12}$$

Maximizing this expression for each component gives

$$y_i = \hat{s}_i - \sigma_n^2 \frac{d}{d\hat{s}_i} \log p_{\mathbf{S}}(\hat{\mathbf{s}}), \quad 1 \le i \le d. \tag{13}$$

Using (25), the second term can be computed as

$$\frac{d}{d\hat{s}_i} \log p_{\mathbf{S}}(\hat{\mathbf{s}}) = -\frac{\hat{s}_i}{\|\hat{\mathbf{s}}\|} \sqrt{\frac{2}{c}} \frac{K_{d/2-p+1}\left(\sqrt{\frac{2}{c}}\|\hat{\mathbf{s}}\|\right)}{K_{d/2-p}\left(\sqrt{\frac{2}{c}}\|\hat{\mathbf{s}}\|\right)}. \tag{14}$$

Therefore, the MAP estimator is

$$y_i = \hat{s}_i \left[1 + \frac{\sigma_n^2}{\|\hat{\mathbf{s}}\|} \sqrt{\frac{2}{c}} \frac{K_{d/2-p+1}\left(\sqrt{\frac{2}{c}}\|\hat{\mathbf{s}}\|\right)}{K_{d/2-p}\left(\sqrt{\frac{2}{c}}\|\hat{\mathbf{s}}\|\right)} \right]. \tag{15}$$

Rewriting this in terms of the norms gives

$$\|\hat{\mathbf{s}}\| = \left(\|\mathbf{y}\| - \sigma_n^2 \sqrt{\frac{2}{c}} \frac{K_{d/2-p+1}\left(\sqrt{\frac{2}{c}}\|\hat{\mathbf{s}}\|\right)}{K_{d/2-p}\left(\sqrt{\frac{2}{c}}\|\hat{\mathbf{s}}\|\right)} \right)_+. \tag{16}$$

This estimator can be computed by successive substitution, namely, $\|\hat{\mathbf{s}}\|^{(k+1)} = f(\|\hat{\mathbf{s}}\|^{(k)})$ with the initialization $\|\hat{\mathbf{s}}\|^{(0)} = \|\mathbf{y}\|$. For $p = 1$, the MBKF MAP estimator specializes to the MLAP case in [19, eq. (17)], while setting $p = d = 1$ results in the soft-threshold rule.

V. MMSE Estimator

In the context of (5), the MMSE estimator is given by the conditional posterior mean

$$\hat{s}_i = E[s_i|\mathbf{y}] = \frac{1}{p_{\mathbf{Y}}(\mathbf{y})} \int_{\mathbb{R}^d} s_i p_{\mathbf{N}}(\mathbf{y}-\mathbf{s}) p_{\mathbf{S}}(\mathbf{s}) d\mathbf{s}. \tag{17}$$

Applied to the MBKF density, the denominator $p_{\mathbf{Y}}(\mathbf{y})$ is given in (10), while the numerator can be manipulated similarly to (6) to give

$$\begin{aligned} \int_{\mathbb{R}^d} s_i p_{\mathbf{N}}(\mathbf{y}-\mathbf{s}) p_{\mathbf{S}}(\mathbf{s}) d\mathbf{s} &= y_i \frac{\exp\left(\frac{\sigma_n^2}{c}\right)}{(2\pi c)^{d/2}\Gamma(p)} \\ &\times \sum_{j=0}^\infty \left(-\frac{\sigma_n^2}{c}\right)^j \frac{(p)_j}{j!} \Gamma\left(p - \frac{d}{2} - j, \frac{\sigma_n^2}{c}; \frac{\|\mathbf{y}\|^2}{2c}\right). \end{aligned} \tag{18}$$

Taking the ratio of (18) and (10) gives the MBKF MMSE estimator in exact form as

$$\hat{s}_i = y_i \frac{\sum_{j=0}^\infty \left(-\frac{\sigma_n^2}{c}\right)^j \frac{(p)_j}{j!} \Gamma\left(p - \frac{d}{2} - j, \frac{\sigma_n^2}{c}; \frac{\|\mathbf{y}\|^2}{2c}\right)}{\sum_{j=0}^\infty \left(-\frac{\sigma_n^2}{c}\right)^j \frac{(p-1)_j}{j!} \Gamma\left(p - \frac{d}{2} - j, \frac{\sigma_n^2}{c}; \frac{\|\mathbf{y}\|^2}{2c}\right)}. \tag{19}$$

For $p = 1$, the infinite sums become finite and (19) simplifies to the MLAP case in [19, eq. (24)]. For $d = 1$, the MMSE estimator is given in [17]. For $p = d = 1$, we get the MMSE estimator of a scalar Laplace random variable in AWGN which has been given in [13].

VI. Approximations

The MBKF MAP and MMSE estimators can be approximated in ways that are computationally efficient, and yet remain close to exact expressions in mean-square error sense.

The MAP estimator in (16) is implicit; replacing $\|\hat{\mathbf{s}}\|$ on the right-hand side by $\|\mathbf{y}\|$ gives an approximate MAP estimator

$$\|\hat{\mathbf{s}}\| = \left(\|\mathbf{y}\| - \sigma_n^2 \sqrt{\frac{2}{c}} \frac{K_{d/2-p+1}\left(\sqrt{\frac{2}{c}}\|\mathbf{y}\|\right)}{K_{d/2-p}\left(\sqrt{\frac{2}{c}}\|\mathbf{y}\|\right)} \right)_+. \tag{20}$$

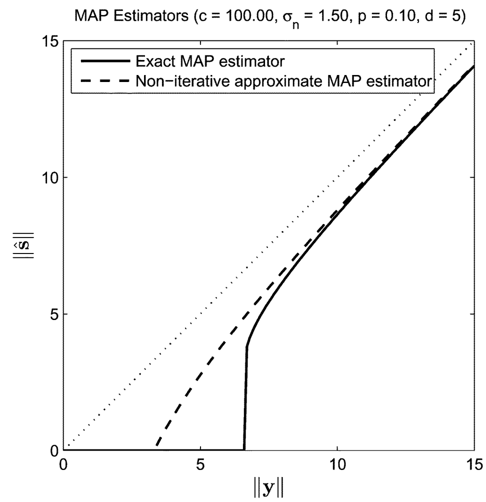

Fig. 2. Exact and approximate estimators (16) and(20) of BKF random vectors in AWGN.

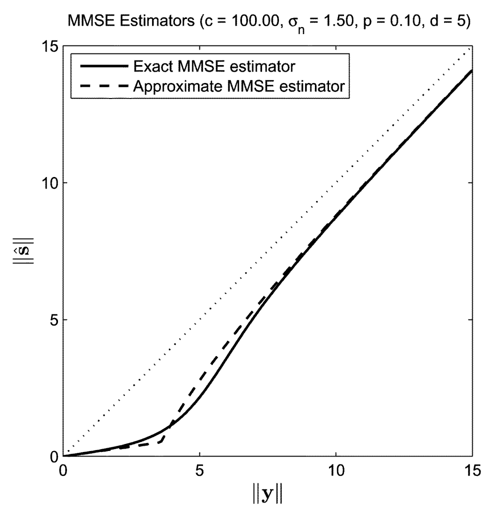

Fig. 3. Exact and approximate estimators (19) and (22) of BKF random vectors in AWGN.

Fig. 2 illustrates the exact and approximate MAP estimators in (16) and (20).

The MMSE estimator may be approximated for small $\|\mathbf{y}\|$ by a line with slope equal to the slope of (19) at $\|\mathbf{y}\| = 0$

$$\|\hat{\mathbf{s}}\| \approx \|\mathbf{y}\| \frac{\sum_{j=0}^{\infty} \left(-\frac{\sigma_n^2}{c}\right)^j \frac{(p)_j}{j!} \Gamma\left(p - \frac{d}{2} - j, \frac{\sigma_n^2}{c}\right)}{\sum_{j=0}^{\infty} \left(-\frac{\sigma_n^2}{c}\right)^j \frac{(p-1)_j}{j!} \Gamma\left(p - \frac{d}{2} - j, \frac{\sigma_n^2}{c}\right)} \tag{21}$$

where $\Gamma(\tau, x)$ is the incomplete Gamma function [5]. For large $\|\mathbf{y}\|$, the approximate MAP estimator (20) is accurate. Therefore, an approximate MMSE estimator is given by

$$\|\hat{\mathbf{s}}\| \approx \max(S, L) \tag{22}$$

where S is the linear approximation in (21), and L denotes the estimator in (20). Fig. 3 compares the exact and approximate MMSE rules in the vector case.

Fig. 4 compares the exact MMSE estimator (19) and the approximate MMSE estimator in [9, eq. (17)]. The exact BKF MMSE nonlinearity shrinks less than the estimator in [9, eq. (17)]. The estimators agree for large $|y|$ and are also almost identical for $p = 1$.

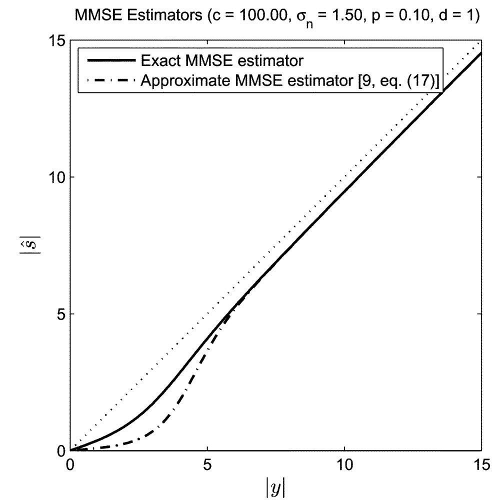

Fig. 4. Two estimators of a BKF random variable in AWGN: exact MMSE estimator (19) and approximate MMSE estimator [9, eq. (17)].

As a further comparison, we derive the conditional mean-square error (MSE)

$$E\left[e_i^2|\mathbf{y}\right] = E\left[(s_i - \hat{s}_i)^2|\mathbf{y}\right] = E\left[s_i^2|\mathbf{y}\right] - 2\hat{s}_i E\left[s_i|\mathbf{y}\right] + \hat{s}_i^2 \tag{23}$$

with $1 \le i \le d$. The term $E[s_i|\mathbf{y}]$ is given in (19), and $E\left[s_i^2|\mathbf{y}\right]$ can be derived as

$$E\left[s_i^2|\mathbf{y}\right] = \frac{\sum_{j=0}^{\infty} \left(-\frac{\sigma_n^2}{c}\right)^j \Gamma_{(j)}^g \left[y_i^2 (p+1)_j + \sigma_n^2 (p)_j\right]}{\sum_{j=0}^{\infty} \left(-\frac{\sigma_n^2}{c}\right)^j \Gamma_{(j)}^g (p-1)_j} \tag{24}$$

where

$$\Gamma_{(j)}^g := \frac{1}{j!} \Gamma\left(p - \frac{d}{2} - j, \frac{\sigma_n^2}{c}; \frac{\|\mathbf{y}\|^2}{2c}\right).$$

With $d = i = 1$, Fig. 5 illustrates the conditional MSE where we have taken the estimators $\hat{s}_i$ to be (19), (22), and [9, eq. (17)]. The proposed approximate MMSE estimator (22) achieves near-optimum MSE performance; the approximation in [9, eq. (17)] differs from the optimum for small to midrange values of $|y|$. The approximation in (22) behaves similarly in the vector case as well.

VII. Image Denoising Experiments

We apply the exact MAP (16), approximate MAP (20), and approximate MMSE (22) estimators to two images corrupted by AWGN, using a critically sampled DWT with three decomposition levels. We have used 80 terms to compute (21) (the terms decay rapidly when $\sigma_n^2 \ll c$). The parameters c and p have been estimated from noisy wavelet data as in [9]. It is assumed that the noise variance σ_n^2 is known. The results (averaged over five realizations), shown in Table I, show the improvement obtained by multivariate probability modeling, compared to univariate probability modeling. In terms of PSNR, the exact MAP estimator (16) is surpassed by its approximate version (20), and the approximate MMSE estimator (22) gives the best overall performance.

VIII. Conclusion

This letter presents new Bayesian estimators for spherically-contoured BKF random vectors in AWGN derived

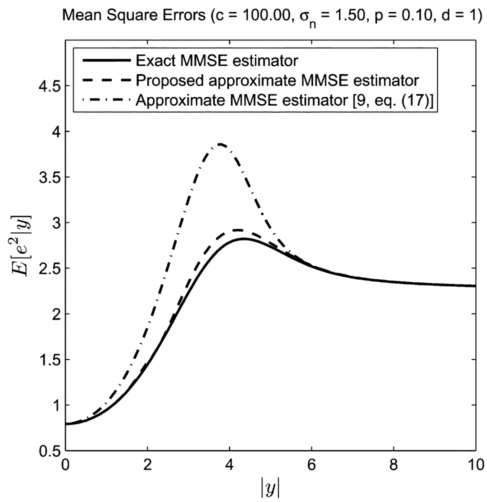

Fig. 5. Comparison of several estimators of BKF random vectors in AWGN. The estimators are (19), (22), and [9, eq. (17)].

TABLE I
PSNR RESULTS FOR WAVELET-DOMAIN IMAGE DENOISING USING THE EXACT MAP ESTIMATOR (16), THE APPROXIMATE MAP ESTIMATOR (20), AND THE APPROXIMATE MMSE ESTIMATOR (22)

	MAP (16)			A-MAP (20)			A-MMSE (22)		
σ_n	1×1	3×3	5×5	1×1	3×3	5×5	1×1	3×3	5×5
Barbara (512×512)									
10	31.00	31.67	31.43	31.36	32.63	32.65	31.50	32.68	32.68
15	28.54	29.14	28.89	28.90	30.20	30.27	29.05	30.25	30.30
20	26.90	27.49	27.27	27.28	28.55	28.66	27.44	28.61	28.69
25	25.67	26.23	25.95	26.09	27.33	27.46	26.27	27.39	27.49
House (256×256)									
10	32.70	32.90	32.40	33.06	33.92	33.71	33.25	34.00	33.74
15	30.71	30.94	30.44	30.99	31.82	31.61	31.15	31.90	31.65
20	29.39	29.64	29.04	29.57	30.39	30.21	29.69	30.48	30.26
25	28.46	28.59	27.96	28.66	29.34	29.13	28.79	29.42	29.19

using the generalized incomplete Gamma function. The proposed MMSE estimator (19) builds upon [9] and [17] by addressing the multivariate, albeit spherically-contoured, case. This letter also proposes a computationally efficient approximate MMSE estimator (22). The presented estimators are limited because in many cases, especially denoising using overcomplete transforms, the multivariate prior should not be spherically-contoured and the noise is not white. However, the MMSE estimator in the more general case, currently under investigation, cannot be expressed in terms of the special functions used here.

APPENDIX

A useful property of the modified Bessel function of the second kind $K_\lambda(u)$ is

$$\frac{d}{du}\log K_\lambda(u) = \frac{\lambda}{u} - \frac{K_{\lambda+1}(u)}{K_\lambda(u)}. \qquad (25)$$

The generalized incomplete Gamma function [5] is given by

$$\Gamma(\tau, x; b) = \int_x^\infty t^{\tau-1}\exp\left(-t - \frac{b}{t}\right) dt. \qquad (26)$$

The Pochhammer symbol $(x)_k$ is defined as

$$(x)_k = x(x-1)\cdots(x-k+1). \qquad (27)$$

ACKNOWLEDGMENT

The authors would like to thank Prof. M. Leung of Polytechnic University for providing a MATLAB implementation of the generalized incomplete Gamma function.

REFERENCES

[1] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions*. New York: Dover, 1972.
[2] A. Achim, A. Bezerianos, and P. Tsakalides, "Novel Bayesian multiscale method for speckle removal in medical ultrasound images," *IEEE Trans. Med. Imag.*, vol. 20, no. 8, pp. 772–783, Aug. 2001.
[3] A. Antoniadis, J. Bigot, and T. Sapatinas, "Wavelet estimators in nonparametric regression: A comparative simulation study," *J. Statist. Softw.*, vol. 6, no. 6, pp. 1–83, 2001.
[4] L. Boubchir and J. M. Fadili, "Bayesian denoising based on the MAP estimation in wavelet-domain using Bessel K form prior," in *Proc. IEEE Int. Conf. Image Processing (ICIP)*, Sep. 2005.
[5] M. A. Chaudhry and S. M. Zubair, *On a Class of Incomplete Gamma Functions with Applications*. London, U.K.: Chapman & Hall/CRC, 2001.
[6] D. Cho and T. D. Bui, "Multivariate statistical modeling for image denoising using wavelet transforms," *Signal Process.: Image Commun.*, vol. 20, no. 1, pp. 77–89, Jan. 2005.
[7] M. A. Clyde and E. I. George, "Empirical bayes estimation in wavelet nonparametric regression," in *Bayesian Inference in Wavelet Based Models*, P. Muller and B. Vidakovic, Eds. New York: Springer-Verlag, 1999, pp. 309–322.
[8] M. S. Crouse, R. D. Nowak, and R. G. Baraniuk, "Wavelet-based signal processing using hidden Markov models," *IEEE Trans. Signal Process.*, vol. 46, no. 4, pp. 886–902, Apr. 1998.
[9] J. M. Fadili and L. Boubchir, "Analytical form for a Bayesian wavelet estimator of images using the Bessel K form densities," *IEEE Trans. Image Process.*, vol. 14, no. 2, pp. 1370–1382, Feb. 2005.
[10] K. Fang, S. Kotz, and K. Ng, *Symmetric Multivariate and Related Distributions*. London, U.K.: Chapman & Hall, 1990.
[11] S. Gazor and W. Zhang, "Speech enhancement employing Laplacian-Gaussian mixture," *IEEE Trans. Signal Process.*, vol. 13, no. 5, pt. 2, pp. 896–904, Sep. 2005.
[12] U. Grenander and A. Srivastava, "Probability models for clutter in natural images," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 23, no. 4, pp. 424–429, Apr. 2001.
[13] M. Hansen and B. Yu, "Wavelet thresholding via MDL for natural images," *IEEE Trans. Inf. Theory*, vol. 46, no. 8, pp. 1778–1788, Aug. 2000.
[14] A. Hanssen and T. A. Øigård, "The normal inverse Gaussian distribution: A versatile model for heavy-tailed stochastic processes," in *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing (ICASSP)*, 2001.
[15] S. Kotz, T. Kozubowski, and K. Podgorski, *The Laplace Distribution and Generalizations*. Cambridge, MA: Birkhaüser, 2001.
[16] R. Martin, "Speech enhancement based on minimum mean-square error estimation and supergaussian priors," *IEEE Trans. Speech Audio Process.*, vol. 13, no. 5, pp. 845–856, Sep. 2005.
[17] S. Nadarajah and S. Kotz, "The BKF Bayesian wavelet estimator," *Signal Process.*, vol. 87, no. 9, pp. 2268–2271, Sep. 2007.
[18] J. Portilla, V. Strela, M. J. Wainwright, and E. P. Simoncelli, "Image denoising using scale mixtures of Gaussians in the wavelet domain," *IEEE Trans. Image Process.*, vol. 12, no. 11, pp. 1338–1351, Nov. 2003.
[19] I. W. Selesnick, "Laplace random vectors, Gaussian noise, and the generalized incomplete gamma function," in *Proc. IEEE Int. Conf. Image Processing (ICIP)*, Atlanta, GA, Oct. 2006.
[20] L. Sendur and I. W. Selesnick, "Bivariate shrinkage with local variance estimation," *IEEE Signal Process. Lett.*, vol. 9, no. 12, pp. 438–441, Dec. 2002.
[21] E. P. Simoncelli, "Modeling the joint statistics of images in the wavelet domain," *Proc. SPIE*, vol. 313, no. 1, pp. 188–195, 1999.
[22] A. Srivastava, X. Liu, and U. Grenander, "Universal analytical forms for modeling image probabilities," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 24, no. 9, pp. 1200–1214, Sep. 2002.