Bayesian wavelet-based Poisson intensity estimation of images using the Fisz transformation

Jalal M. Fadili, Jérôme Mathieu, Barbara Romaniuk, Michel Desvignes

Bayesian wavelet-based Poisson intensity estimation of images using the Fisz transformation

Jalal M. FADILI[a], Jérôme Mathieu[a], Barbara Romaniuk[a] and Michel Desvignes[b]

[a]*GREYC CNRS UMR 6072, ISMRA 6, Bd du Maréchal Juin, 14050 Caen, France*
[b]*LIS CNRS UMR 5083, 961 rue de la Houille Blanche, BP 46, 38402 St. Martin d'Héres*
E-mail: Jalal.Fadili@greyc.ismra.fr

Abstract

A novel wavelet-based Poisson-intensity estimator of images is presented. This method is based on the asymptotic normality of a certain function of the Haar wavelet and scaling coefficients called the Fisz *transformation. Some asymptotic results such as normality and decorrelation of the transformed image samples are extended to the 2D case. This Fisz-transformed image is then treated as if it was independent and Gaussian variables and we apply a novel Bayesian denoiser that we have recently developed. In the latter, a prior model is imposed on the wavelet coefficients designed to capture the sparseness of the wavelet expansion. Seeking probability models for the marginal densities of the wavelet coefficients, the new family of* **Bessel K forms** *densities are shown to fit very well to the observed histograms. Exploiting this prior, we designed a Bayesian nonlinear denoiser and a closed-form for its expression was derived. Our Fisz-transformation based Bayesian denoiser compares very favorably to variance stabilizing transformation methods in both smooth and piece-wise constant intensities. It clearly outperforms the other denoising methods especially in the low-count setting.*

Keywords: Wavelets, Poisson process, Fisz transformation, Bayesian denoiser, Bessel K forms.

1. Introduction

Nonparametric wavelet-based regression has been a fundamental tool in data analysis over the past two decades and is still an expanding area of ongoing research. The goal is to recover an unknown image, say g, based on sampled data that are contaminated with noise which is supposed Gaussian in most situations. Only very general assumptions about g are made such as that it belongs to a certain class of functions (e.g. Besov space). Nonparametric regression (or denoising) techniques provide a very effective and simple way of finding structure in data sets without the imposition of a parametric regression model. During the 1990s, the nonparametric regression literature was arguably dominated by nonlinear wavelet shrinkage and wavelet thresholding estimators [1–3]. These estimators are a new subset of an old class of nonparametric regression estimators, namely orthogonal series methods. Moreover, these estimators are easily implemented through fast algorithms so they are very appealing in practical situations [4].

Since the seminal papers by Donoho & Johnstone [1], the image processing literature have been inundated by hundreds of papers applying or proposing modifications of the original algorithm in image processing problems. Various data adaptive wavelet thresholding estimators have been developed, see for example the extensive reviews in [5, 6] and references therein. Various Bayesian approaches for nonlinear wavelet thresholding and nonlinear wavelet shrinkage estimators have also recently been proposed. These estimators have been shown to be effective and it is argued that they are less ad-hoc than the classical proposals discussed above. In the Bayesian approach a prior distribution is imposed on the wavelet coefficients. The prior model is designed to capture the sparseness of wavelet expansions. Then the image is estimated by applying a suitable Bayesian rule to the resulting posterior distribution of the wavelet coefficients. A detailed study involving recent classical and Bayesian wavelet methods was carried out in [7] in the development towards high-performance wavelet estimators and their finite sample properties.

Some authors have already considered the problem of estimating the intensity of a Poisson process using a wavelet-based technique. The classical approach consists in first preprocessing the data using variance stabilizing

transformations such as the Anscombe's square-root transformation (see overview in [8]), so that the noise becomes approximately Gaussian. Then the analysis proceeds as if the noise was indeed Gaussian, yielding after applying the inverse transformation an estimate of the intensity of the process. Kolaczyk [9] has also proposed a method based on the shrinkage of Haar wavelet coefficients of the original vector taking into account the particular nature of the Poisson noise and adjusts the thresholding scheme. His algorithm outperforms the standard method for a certain class of Poisson processes. However, only the Haar basis can be used in its algorithm which is restrictive when the intensity image is known to be smooth. More recently, Fryzlewicz and Nason [10] proposed an alternative wavelet-based algorithm for estimating the deterministic discretized intensity function of an inhomogeneous one-dimensional Poisson process. Their method is based on the asymptotic normality of a certain function of the Haar wavelet and scaling coefficients called the Fisz transformation. Their numerical experiments show that the Fisz transformation method tends to perform better than both the Anscombe and Kolaczyk's methods.

In this paper, we propose to extend this idea to the 2D case (images) and prove some asymptotic results of the transformed image (normality), which follow in a similar way as in [10]. The idea behind the algorithm is the following: we first preprocess the observed image $\mathbf{y}$ using the nonlinear wavelet-based Fisz transformation, and the preprocessed image is then treated as if it was Gaussian. The BKF Bayesian denoiser described above is then applied on this transformed image.

This paper is organized as follows: In Section 2 we define the nonparametric regression problem and introduce some notational aspects. The Fisz Gaussianizing transformation algorithm is then presented in Section 4. The wavelet Bayesian denoiser, using the Bessel K forms (BKF) prior, is briefly presented in Section 3. We only give the main results and the interested reader can refer to [23] for a detailed mathematical description. Section 5 compares the performance of the designed algorithm with classical denoisers on simulated and real medical images. Finally, conclusions and directions of future work are drawn.

2. Nonparametric wavelet-based regression

Let's first consider the Gaussian case. For the Poisson noise, a suitable transformation will be applied to match our Gaussian noise model below. Let $g_{mn}, m, n = 1, \ldots, N$ equally-spaced samples of a real-valued image. N is considered as a power of 2. Consider the standard nonparametric regression setting :

$$y_{mn} = g_{mn} + \epsilon_{mn} \tag{1}$$

where ϵ_{mn} are iid central normal random variables with variance σ^2 independent of g_{mn}. The goal is to recover the underlying function g from the observed noisy data y_{mn}, without assuming any particular parametric structure for g. Let $\mathbf{y}$, $\mathbf{g}$ and ϵ denote the matrix representation of the corresponding samples. Let $\mathbf{D} = \mathcal{W}\mathbf{y}$, $\mathbf{S} = \mathcal{W}\mathbf{g}$ and $\mathbf{V} = \mathcal{W}\epsilon$, where $\mathcal{W}$ is the two dimensional dyadic orthonormal wavelet transform (DWT) operator [11]. In a two dimensional setting, the subbands HH_j, HL_j and LH_j, $j = 1, \ldots, J$ correspond to the detail coefficients in diagonal, horizontal and vertical orientations, and the subband LL_J is the approximation or the smooth component. J is the coarsest scale of the decomposition. Let s^{oj}_{mn} be the detail coefficient of the true image $\mathbf{g}$ at location (m, n), scale j and orientation o, and similarly for d^{oj}_{mn} and v^{oj}_{mn}. Due to the orthogonality of the basis, v^{oj}_{mn}, the DWT of white noise are also independent normal variables with the same variance. It follows from Eq.1 that :

$$d^{oj}_{mn} = s^{oj}_{mn} + v^{oj}_{mn}, j = 1, \ldots, J, m, n = 0, \ldots, 2^j - 1 \tag{2}$$

The sparseness of the wavelet expansion makes it reasonable to assume that essentially only a few large detail coefficients in $\mathbf{D}$ contain information about the underlying image $\mathbf{g}$, while small values can be attributed to the noise which uniformly contaminates all wavelet coefficients. It is also advisable to keep the approximation coefficients intact because they represent low-frequency terms that usually contain important features about the image $\mathbf{g}$. By thresholding or shrinking the detail coefficients and inverting the DWT, one can obtain an estimate of the underlying image $\mathbf{g}$. There are a variety of methods in the literature to choose the threshold level and the thresholding rule, see [7] for a review and small sample performance of these methods.

3. Bayesian denoising using the BKF prior

In the Bayesian approach a prior distribution is imposed on the wavelet coefficients in order to capture the sparseness of the wavelet expansion. The following section is intended to provide an introduction to Bessel K forms distributions family suitable to characterize the wavelet subband coefficients densities which have been already observed to be sharply peaked and heavily tailed. Many choices for this prior have been proposed in the literature: scale mixture of two normal distributions [12] or one normal distribution and a point mass at zero [13, 14], double exponential prior with a point mass at zero [15], the Generalized Gaussian Distribution (GGD) [16–18]. However, most of these priors suffer from a lack of capturing the heavy tail behavior of the observed wavelet coefficients densities. Based upon this observation, authors in [19] used α-stable distributions [20], a family of heavy tailed densities, as a prior to capture the sparseness of the wavelet coefficients at each scale. These authors showed the superiority of the α-stable distributions in fitting the mode and the tail behavior of the wavelet coefficients distributions. However, their hyperparameters estimator is very poor in the presence of contaminating noise and remains an important issue. Furthermore, in both the GGD and the α-stable priors, the derived Bayesian estimator has no closed analytical form in general situation and involves intensive numerical integration.

Using a physical model for image formation (the so called *transported generator model*), a family of two-parameter probability densities, called Bessel K forms (BKF), have been proposed in [21, 22] to model the distribution of arbitrary images that have been filtered by a variety of band-pass filters (e.g. derivative, Gabor, interpolation, steerable filters, etc). It is evident that wavelet decompositions of an image are members of such class of filters. Therefore, the BKF is a suitable model provided that the resulting wavelet coefficients marginals are: unimodal, symmetric around the mode and leptokurtic. The first two conditions are widely adopted in the literature and are common to other priors such as the α-stable or the GGD models. The last condition simply mean that the prior is a sharply peaked distribution with tails that are heavier as compared to normal density of the same variance. The BKF is then adapted to capture the heavy tail behavior of wavelet coefficients densities. Exploiting this prior in a Bayesian framework, we designed a posterior conditional mean (PCM) estimator [23]. We also derived a closed-form for its expression as a main result of our paper.

The BKF prior of [21] is given by:

$$f(x;c,p) = \frac{1}{\sqrt{\pi}\Gamma(p)} \left(\frac{c}{2}\right)^{-\frac{p}{2}-\frac{1}{4}} \left|\frac{x}{2}\right|^{p-\frac{1}{2}} K_{p-\frac{1}{2}}\left(\sqrt{\frac{2}{c}}\,|x|\right) \tag{3}$$

for $p > 0, c > 0$, where $K_\nu(x)$ is the modified Bessel function [24, 25]. p and c are respectively the shape and scale parameters. Using this prior, we have shown that the marginal pdf of of the observed wavelet coefficients d given σ^2, p and c is [23]:

$$\tilde{f}_d(d;p,c,\sigma^2) = \left(\frac{\sigma^2}{2c}\right)^{\frac{p}{2}} \phi(d;,0,\sigma^2)\,(I_+ + I_-) \tag{4}$$

where $I_\pm = e^{\left(\frac{\pm\frac{d}{\sigma}+\sigma\sqrt{\frac{2}{c}}}{2}\right)^2} D_{-p}\left(\pm\frac{d}{\sigma}+\sigma\sqrt{\frac{2}{c}}\right)$. $D_\nu(x)$ stands for the Parabolic Cylinder function of fractional order ν [24, 25]. f_d is an even function whose mode is at zero.

It is well known that L_2-based Bayes rules correspond to posterior conditional means (PCM) estimates of wavelet coefficients s (conditionally on the hyperparameters σ^2, p and c). In [23], a closed-form expression for this Bayesian denoiser has been established:

$$s_{\mathrm{PCM}}(d) = p\sigma \frac{e^{\left(\frac{-\frac{d}{\sigma}+\sigma\sqrt{\frac{2}{c}}}{2}\right)^2} D_{-p-1}\left(-\frac{d}{\sigma}+\sigma\sqrt{\frac{2}{c}}\right) - e^{\left(\frac{\frac{d}{\sigma}+\sigma\sqrt{\frac{2}{c}}}{2}\right)^2} D_{-p-1}\left(\frac{d}{\sigma}+\sigma\sqrt{\frac{2}{c}}\right)}{e^{\left(\frac{-\frac{d}{\sigma}+\sigma\sqrt{\frac{2}{c}}}{2}\right)^2} D_{-p}\left(-\frac{d}{\sigma}+\sigma\sqrt{\frac{2}{c}}\right) + e^{\left(\frac{\frac{d}{\sigma}+\sigma\sqrt{\frac{2}{c}}}{2}\right)^2} D_{-p}\left(\frac{d}{\sigma}+\sigma\sqrt{\frac{2}{c}}\right)} \tag{5}$$

for $0 < p \leq 1$ and c strictly positive. To implement the formula in Eq.5 in practice, we also proposed a cumulant-based estimator of the hyperparameters involved in this nonparametric regression problem, namely p, c and σ^2 [23].

4. Denoising with a Poisson noise

Some authors have already considered the problem of estimating the intensity of a Poisson process using a wavelet-based technique. The usual (regression) setting is as follows: the possibly inhomogeneous independent Poisson process is observed and discretized into an image $\mathbf{y}$. Each y_{mn} can be thought of as coming from a Poisson distribution with an unknown intensity λ_{mn}, which needs to be estimated. The classical approach consists in first preprocessing the data using variance stabilizing transformations such as the Anscombe's square-root transformation (see overview in [8]), so that the noise becomes approximately Gaussian. Then the analysis proceeds as if the noise was indeed Gaussian, yielding after applying the inverse transformation an estimate of the intensity of the process. Kolaczyk [9] has also proposed a method based on the shrinkage of Haar wavelet coefficients of the original vector taking into account the particular nature of the Poisson noise and adjusts the thresholding scheme. His algorithm outperforms the standard method for a certain class of Poisson processes. However, only the Haar basis can be used in its algorithm which is restrictive when the intensity image is known to be smooth. More recently, Fryzlewicz and Nason [10] proposed an alternative wavelet-based algorithm for estimating the deterministic discretized intensity function of an inhomogeneous one-dimensional Poisson process. Their method is based on the asymptotic normality of a certain function of the Haar wavelet and scaling coefficients called the Fisz transformation. Their numerical experiments show that the Fisz transformation method tends to perform better than both the Anscombe and Kolaczyk's methods.

In this paper, we propose to extend this idea to the 2D case (images) and prove some asymptotic results of the transformed image (normality), which follow in a similar way as in [10]. The idea behind the algorithm is the following: we first preprocess the observed image $\mathbf{y}$ using the nonlinear wavelet-based Fisz transformation, and the preprocessed image is then treated as if it was Gaussian. The BKF Bayesian denoiser described above is then applied on this transformed image.

4.1. The 2D Fisz transformation

Given the image of counts y_{mn}, which is the realization of independent random Poisson variables $Y_{mn} \sim \mathcal{P}(\lambda_{mn})$, the 2D Fisz transformation algorithm is as follows:

1. We apply a one step wavelet transform using the Haar wavelet. However, we use the non-normalized filters $\{1/2, -1/2\}$ and $\{1/2, 1/2\}$ instead of the usual Haar filters. The detail coefficients d^o_{mn} at each orientation o and location (m, n) and the approximation coefficient a^{LL}_{mn} are given by:

$$\begin{array}{l} a^{LL}_{mn} = \frac{(y_{2m+1,2n+1}+y_{2m,2n+1})+(y_{2m+1,2n}+y_{2m,2n})}{4} \\ d^{HH}_{mn} = \frac{(y_{2m+1,2n+1}+y_{2m,2n})-(y_{2m+1,2n}+y_{2m,2n+1})}{4} \\ d^{HL}_{mn} = \frac{(y_{2m+1,2n+1}+y_{2m,2n+1})-(y_{2m+1,2n}+y_{2m,2n})}{4} \\ d^{LH}_{mn} = \frac{(y_{2m+1,2n+1}+y_{2m+1,2n})-(y_{2m,2n+1}+y_{2m,2n})}{4} \end{array} \tag{6}$$

2. We now replace the detail coefficients in each orientation as follows:

$$z^o_{mn} = \begin{cases} 0 & \text{if } a^{LL}_{m,n} = 0 \\ \frac{d^o_{mn}}{\sqrt{a^{LL}_{mn}}} & \text{otherwise.} \end{cases} \tag{7}$$

The new detail coefficients are realizations of random variables Z^o_{mn} whose distribution is given by the following lemma:

Lemma 4.1 If Y_{mn} are independent Poisson variables, then it follows that:

$$Z^o_{m,n} \xrightarrow{d} \mathcal{N}(0, 1/2) \tag{8}$$

as

$$\text{for } (i,j) \in \{0,1\}^2 \quad \lambda_{2m+i,2n+j} \to \infty,$$
$$\frac{\lambda_{2m+1,2n}}{\lambda_{2m,2n}} \to 1, \frac{\lambda_{2m+1,2n+1}}{\lambda_{2m,2n+1}} \to 1, \frac{\lambda_{2m,2n+1}}{\lambda_{2m,2n}} \to 1$$

where $\xrightarrow{d}$ means converge in distribution. The proof of this lemma follows from Fisz theorem [26] for 1D Poisson processes. The conditions stated above has been adjusted to the 2D case. This lemma means that we can expect the modified detail coefficients to come from a normal distribution with variance 1/2 provided that the underlying Poisson intensities are large and close enough in any 2×2 neighborhood. Significant deviation from these assumptions will result in departure from normality.

3. We repeat steps 1-2 at each scale keeping the approximation coefficients intact. The modified detail coefficients at each scale can be considered as a Gaussianized version of the original detail coefficients.

4. We then reconstruct the new image $\mathbf{u}$ using the inverse Haar wavelet transform (with the non-normalized filters). We can write:

$$\mathbf{u} = \mathcal{F}\mathbf{y} \tag{9}$$

where $\mathcal{F}$ is the Fisz transformation operator.

Following the same procedure as in [10], we can establish a general explicit formula for the operator $\mathcal{F}$. One can also easily show that $\sum_{m,n} u_{mn} = \sum_{m,n} y_{mn}$ and if $\mathbf{y}$ is a constant image, then $\mathcal{F}\mathbf{y} = \mathbf{y}$. More interestingly, two properties, which are extensions of Fryzlewicz's results [10] to the 2D case, can be proved. Suppose that the Y_{mn} are iid Poisson variables with mean λ, then the coefficients of the Fisz transformed Poisson distributed image $\mathbf{u}$ are asymptotically **uncorrelated** and **normal** with unit variance and mean λ. This properties are verified provided that the sample size N and the minimum of the underlying intensity image are large enough, and that the mean of the underlying intensity image $\mathbf{\Lambda}$ is small. See [10] for a complete sketch of the proof in the 1D case. In practical situations, these assumptions hold and the Fisz transformation works well even in the case of non-constant intensities.

As an example, let's consider a 256×256 8 bit-grayscale Hoffman phantom as the intensity image $\mathbf{\Lambda}$ (with a maximum of 255), which is widely used in positron emission tomography (Poisson noise contaminated images). Fig.1 compares the Q-Q plots and the auto-correlation functions (ACF) of $\mathcal{F}\mathbf{y} - \mathcal{F}\mathbf{\Lambda}$ and $\mathcal{A}\mathbf{y} - \mathcal{A}\mathbf{\Lambda}$, where $\mathcal{A}$ is the Anscombe transformation operator. The horizontal planes are the 5% critical levels. The Fisz transformation does a better job than the basic Anscombe transformation in Gaussianizing the transformed image $\mathbf{y}$. The ACFs are very close and the Fisz operator does not introduce additional correlation. This normality and lack of correlation naturally yields the BKF Bayesian denoising setting described above. To estimate the original intensity Λ of a Poisson noise contaminated image, our Bayesian denoising algorithm will consist of the following steps:

1. Calculate the Fisz-transformed image $\mathbf{u}$ which is a Gaussianized version of the Poisson image.
2. Apply the nonlinear BKF Bayesian denoising algorithm described above on $\mathbf{u}$.
3. Perform the inverse Fisz transformation on the denoised image $\hat{\mathbf{u}}$ to obtain the true intensity estimate $\hat{\mathbf{\Lambda}}$.

$$\mathbf{y} \xrightarrow{\text{Fisz transformation } \mathcal{F}} \mathbf{u} \xrightarrow{\text{Wavelet-based Denoising}} \hat{\mathbf{u}} \xrightarrow{\text{Inverse Fisz transformation } \mathcal{F}^{-1}} \hat{\mathbf{\Lambda}}$$

5. Experimental results

In this section, we compare the performance of our Bayesian denoising algorithm in the context of Poisson noise to other algorithms, when the data is preprocessed using the Fisz transform. These algorithms were the PCM estimator with α-stable prior for which we have recently proposed a fast algorithm in [27], hard and soft universal thresholding[1], visu thresholding, minimax thresholding and SURE method. The coarsest level of decomposition was chosen to be $\log_2 \log N^2 + 1$ from asymptotic considerations [7]. We used the Daubechies wavelet with regularity 4.

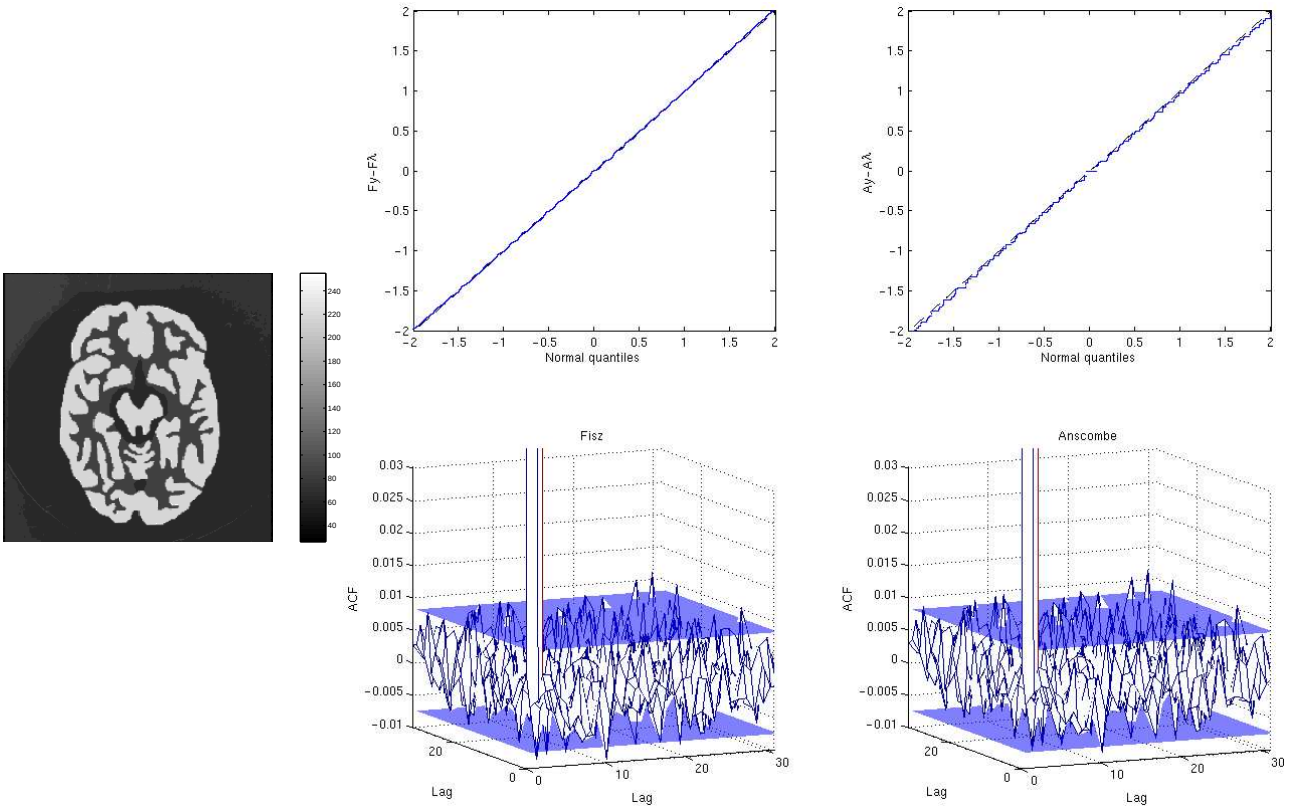

Figure 1. Comparison between Fisz and Anscombe transformations on Hoffman phantom intensity image. The top row shows the Q-Q plots and the bottom row shows the auto-correlation functions.

5.1. Simulated data

We first give some simulations, using the Hoffman phantom image as the true intensity image, to assess the performance of the different algorithms. An example of the visual quality of some denoised images using the Bayesian BKF and the classical universal thresholding methods is shown in Fig.2 for the Hoffman intensity image, scaled with a factor of 0.01 (low counts). The details of the brain are not easily visible in the noisy image. The Fisz (first column) and the Anscombe (second column) transformations are compared for each denoising method. It can be clearly seen that the visual quality as well as the output SMR are higher for the Fisz transformation. The Bayesian BKF is clearly the best denoising method among the one tested. The result of this experiment are confirmed by the Monte Carlo simulation results shown in Fig.3. This figure shows the output mean signal-to-mean square error ratio (SMR) in dB in estimating the Hoffman phantom intensity image as a function of the logarithm of a scaling factor, by which the intensity image was multiplied in order to assess the performance of the methods in both the low- and high-count setting. The output SMR was averaged for each scaling factor on 100 simulated trials. The first 7 panels compare the Fisz transformation (solid line) to the Anscombe transformation (dashed line) for each denoising method. The last two panels respectively compare the various denoising techniques for each transformation. Our Bayesian denoiser clearly outperforms the other methods especially in the low-count setting. Only the α-stable Bayesian denoiser competes with it in the high-count configuration. In the first 7 panel, one clearly see that the Fisz transformation does a better job than the Anscombe transformation for all the denoising procedures used. The difference in performances between the two transformations becomes higher (in favor of the Fisz transformation) as the scaling factor decreases, i.e. low-count setting. As pointed out in [10] in the 1D case, the Fisz transform appears to perform better with hard than with soft thresholding.

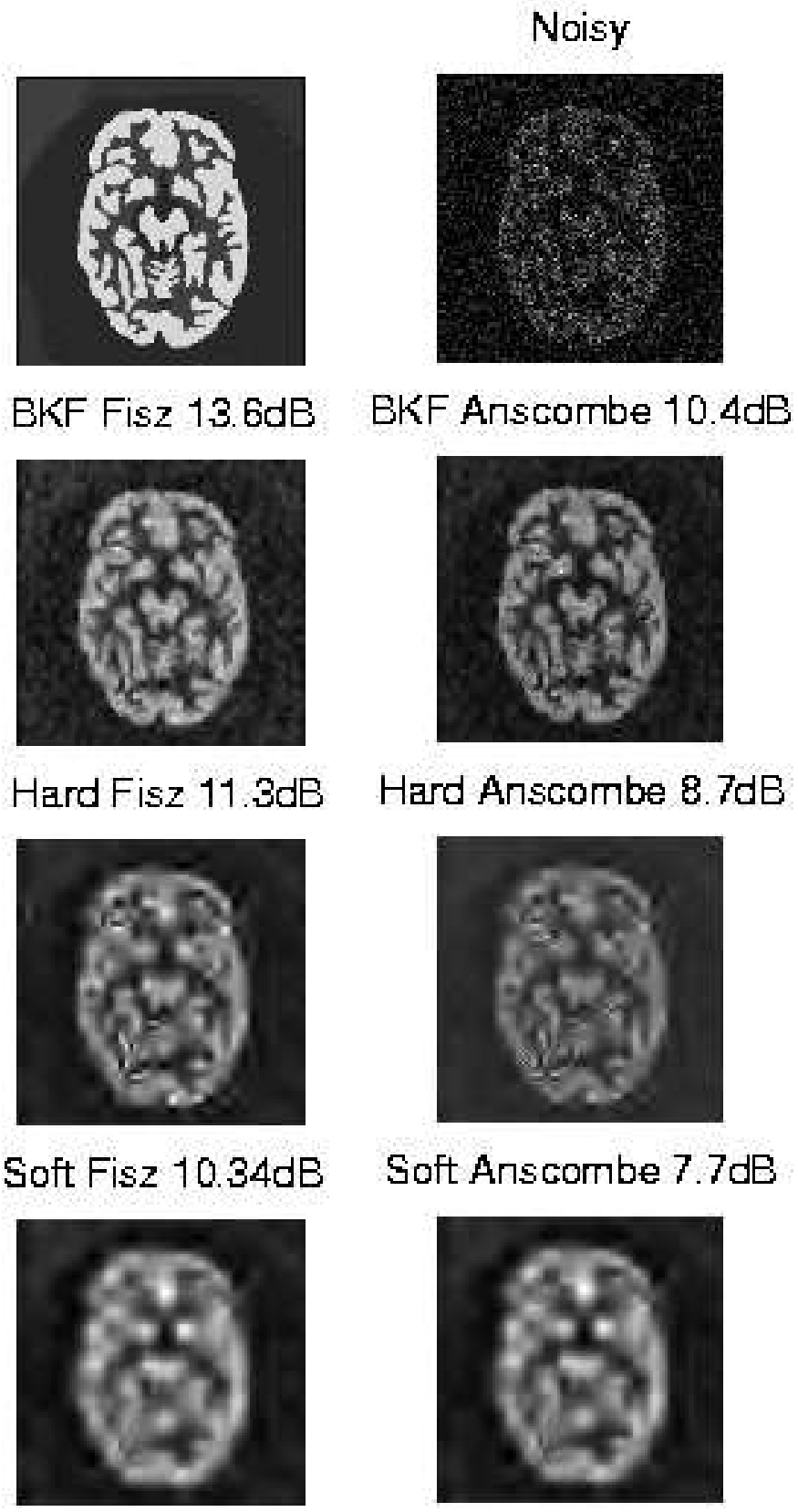

Figure 2. Example of applying the Fisz (first column) and the Anscombe variance-stabilizing (second column) transformations with various denoising methods for the Hoffman phantom image corrupted by Poisson noise with a scaling factor of 0.01 (low counts). The first panel depicts the original (scaled) intensity image and its noisy version (second panel). Results of the Bayesian BKF denoiser are in the second row, universal hard and soft thresholding are respectively in the third and fourth rows.

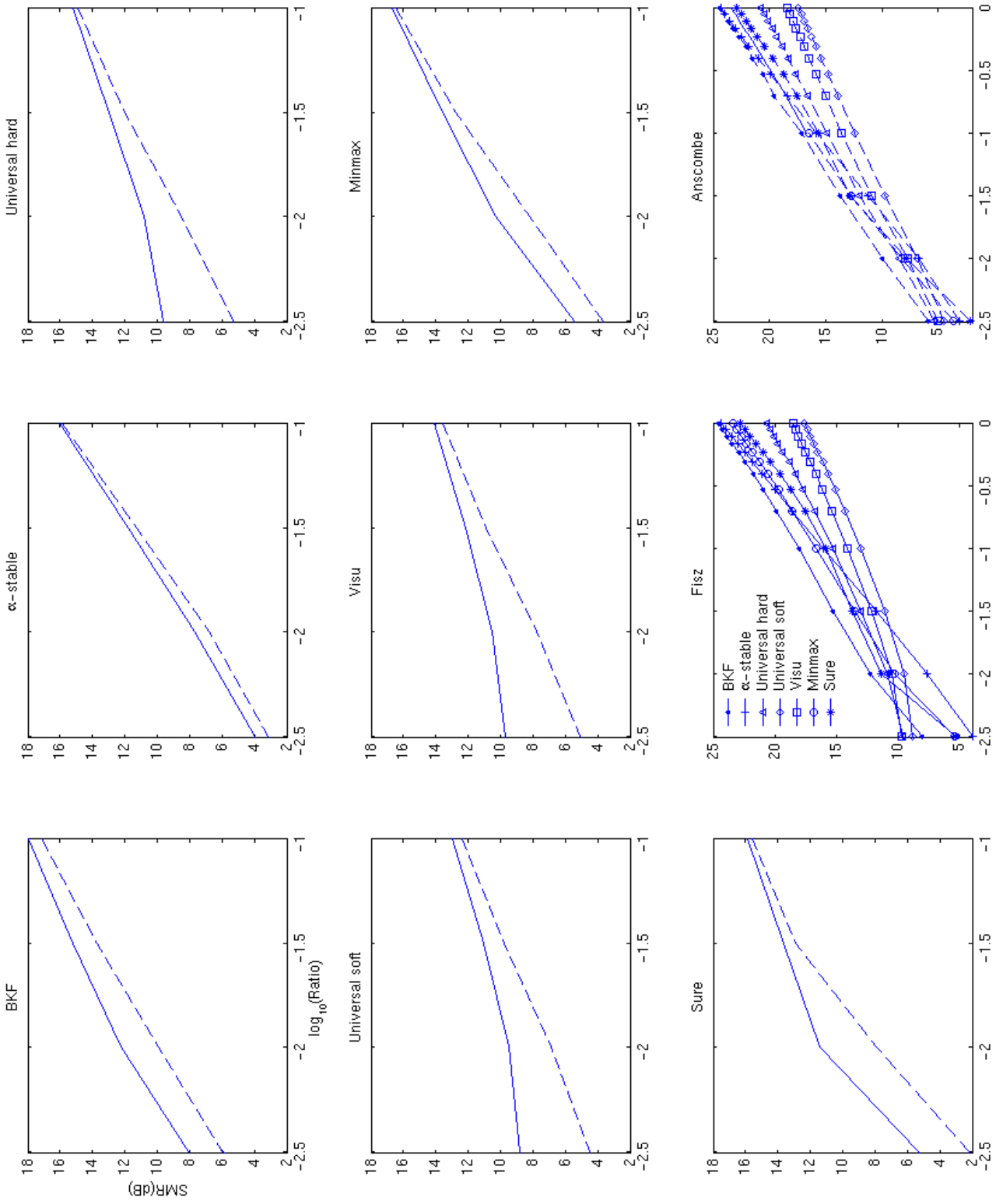

Figure 3. The average SMR in dB over 100 trial simulation in estimating the Hoffman phantom intensity image as a function of the logarithm of a scaling factor, by which the intensity image was multiplied in order to assess the performance of the methods in both the low- and high-count setting. The first 7 panels compare the Fisz transformation (solid line) to the Anscombe transformation (dashed line) for each denoising method. The last two panels respectively compare the various denoising techniques for each transformation.

5.2. X-ray medical angiography

In this section, the proposed method is applied to X-ray medical angiography images. The value of each pixel represents the number of X photons counted at the pixel location on the sensor. It is well known that such a counting process typically has a Poisson distribution. the goal is then to estimate the unknown intensity image $\mathbf{\Lambda}$ using the proposed wavelet-based method.

Fig.4 compares the Fisz and Anscombe transformations with various denoising algorithms on an X-ray medical angiography of the breast. Only a zoom on a specific region of interest is shown. The first image is the original zoomed region of interest. The first column (a,c,e,g images) shows the result of denoising using the Fisz transformation and the second column (b,d,f,h images) corresponds to the Anscombe transformation with the same denoising algorithms. One can clearly see that the denoised image obtained by the BKF Bayesian method is visually better than the other methods (better than the $\alpha - stable$ Bayesian denoise). The contrast in the denoised image is higher yielding a better visibility of anatomical structures such as bones. As expected, the soft thresholding gives an oversmooth estimate while the hard thresholding approach results in a high ringing effect which can be observed at many locations of the estimated image. There is only a slight improvement using the Fisz transformation instead of the Anscombe transformation. The differences are not clearly visible.

We have also plotted the quantiles of the difference between the noisy transformed image and its denoised version (before inverting the transform) for each denoising algorithms. In the perfect case, as predicted by theoretical results (Section 4.1), this residual image should be Gaussian for the Fisz transformation and tends to Gaussian for the Anscombe transformation. The results are depicted in Fig.5. The first Q-Q plot is the one for which the BKF denoising algorithm was applied without any Gaussianizing transformation applied on the noisy image. As expected, the residuals clearly depart from normality even at low scores. When combining the Fisz transformation and the BKF denoising algorithm, the best performance is reached as the Q-Q plot follows the identity line. The slight departure from normality is due to low denoising errors. The same behavior is observed with the Anscombe transformation and only slight differences are observed between the two transformations for any denoising algorithm. This can be explained by the fact that the treated image has high counts pixels. This observation confirms the previous simulations results where the saliency of the Fisz transformation is higher in the low-count setting. When compared to the other denoising methods, the BKF is clearly better. Serious departure from Gaussianity are observed for example for the soft universal thresholding which comes from the higher estimation bias of this method. Its residual image contains not only noise but also many missed true signal pixels.

6. Conclusion

In this paper, we have introduced a new wavelet-based Gaussianizing transformation for Poisson intensity image estimation. Its statistical as well as empirical properties are investigated using simulated and real medical data sets. Combining this transformation with a previously developed Bayesian denoising algorithm has proven very powerful in estimating the true Poisson intensity image.

7. References

[1] D. L. Donoho and I. M. Johnstone, "Ideal spatial adaptation by wavelet shrinkage," *Biometrika* **81**(3), pp. 425–455, 1994.

[2] D. L. Donoho and I. M. Johnstone, "Adapting to unknown smoothness via wavelet shrinkage," *Journal of the American Statistical Association* **90**(432), pp. 1200–1224, 1995.

[3] D. L. Donoho, I. M. Johnstone, G. Kerkyacharian, and D. Picard, "Wavelet shrinkage: Asymptopia?," *J. R. Statist. Soc. B.* **57**(2), pp. 301–337, 1995.

[4] S. G. Mallat, "A theory for multiresolution signal decomposition: the wavelet representation," *IEEE Trans. PAMI* **11**(7), pp. 674–693, 1989.

[5] B. Vidakovic, *Statistical Modeling by Wavelets*, John Wiley & Sons, New York, 1999.

[6] D. B. Percival and A. T. Walden, *Wavelet Methods for Time Series Analysis*, Cambridge press, 2000.

[7] A. Antoniadis, J. Bigot, and T. Sapatinas, "Wavelet estimators in nonparametric regression: A comparative simulation study," *Journal of Statistical Software* **6**(6), 2001.

[8] J.-L. Starck and F. Murtagh, "Astronomical image and signal processing looking at noise, information and scale," *IEEE Sig. Proc. Magazine* **18**, pp. 30–40, 2001.

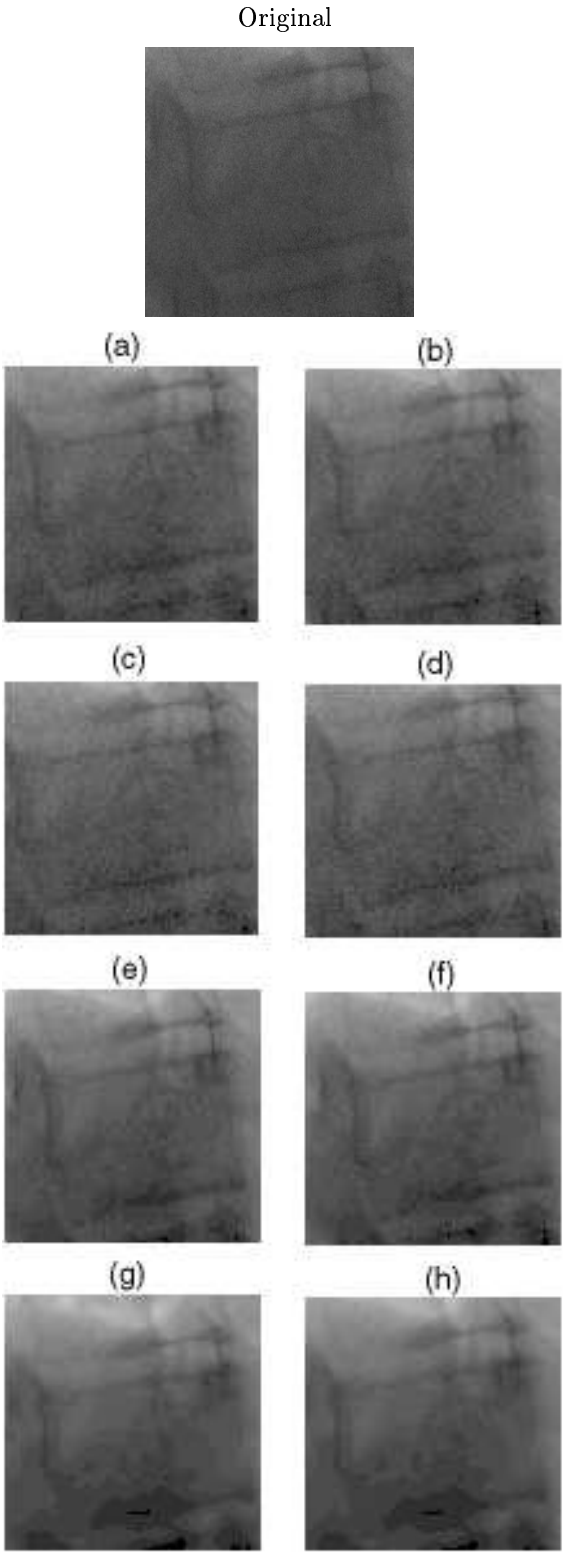

Figure 4. Comparing the Fisz and Anscombe transformations with various denoising algorithms on an X-ray medical angiography of the breast. Only a zoom on a specific region is shown. The first image is the original zoomed region of interest. The first column (a,c,e,g images) shows the result of denoising using the Fisz transformation and the second column (b,d,f,h images) corresponds to the Anscombe transformation with the same denoising algorithms. (a)-(b) BKF. (c)-(d) α-stable. (e)-(f) hard universal thresholding. (g)-(h) soft universal thresholding.

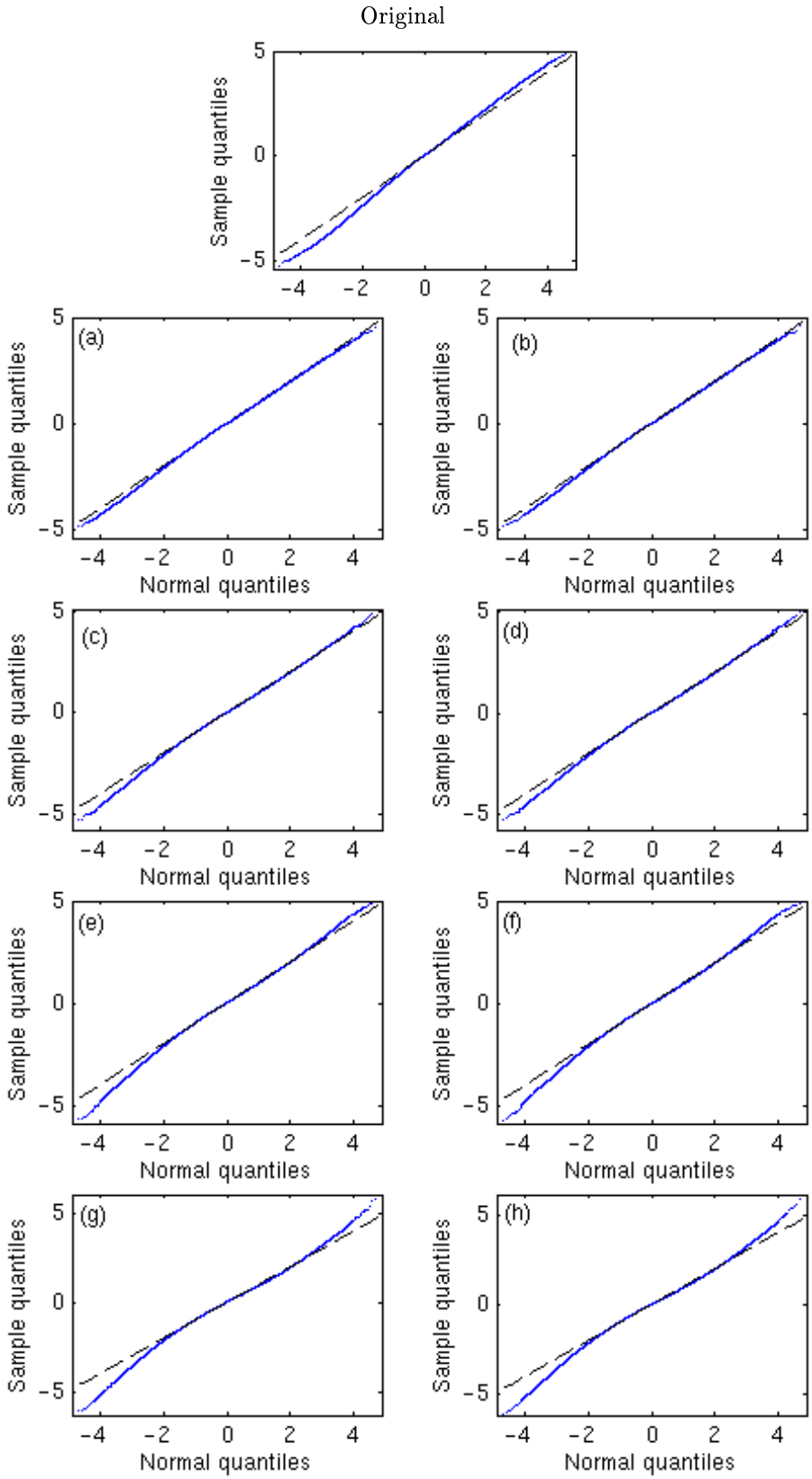

Figure 5. Data quantiles vs Gaussian quantiles for each residual image before the inverse transformation (Fisz or Anscombe) is applied. The first image shows the Q-Q plot of the residual image without any transformation applied on the noisy original image. The first column (a,c,e,g images) shows the Q-Q plots using the Fisz transformation and the second column (b,d,f,h images) corresponds to the Anscombe transformation with the same denoising algorithms. (a)-(b) BKF. (c)-(d) α-stable. (e)-(f) hard universal thresholding. (g)-(h) soft universal thresholding.

[9] E. Kolaczyk, "Wavelet shrinkage estimation of certain poisson intensity signals using corrected thresholds," *Statistica Sinica* **9**, pp. 119–135, 1999.
[10] P. Fryzlewicz and G. P. Nason, "A wavelet-fisz algorithm for poisson intensity estimation," tech. rep., Department of Mathematics, University of Bristol, 2002.
[11] S. G. Mallat, *A Wavelet tour of signal processing*, Academic Press, New York, 2nd ed., 1999.
[12] H. Chipman, E. Kolaczyk, and R. McCulloch, "Adaptive bayesian wavelet shrinkage," *J. Am. Statist. Ass.* **92**, pp. 1413–1421, 1997.
[13] M. Clyde, G. Parmigiani, and B. Vidakovic, "Multiple shrinkage and subset selection in wavelets," *Biometrika* **85**(2), pp. 391–401, 1998.
[14] F. Abramovich, T. Sapatinas, and B. Silverman, "Wavelet thresholding via a bayesian approach," *J. R. Statist. Soc. B* **60**, pp. 725–749, 1998.
[15] B. Vidakovic and F. Ruggeri, "Bams method: Theory and simulations," tech. rep., Institute of Statistics and Decision Sciences, Duke University, 2000.
[16] E. P. Simoncelli and E. H. Adelson, "Noise removal via bayesian wavelet coring," in *Third Int'l Conf on Image Proc*, **1**, pp. 379–382, IEEE Sig Proc Society, (Lausanne), 1996.
[17] S. Chang, B. Yu, and M. Vetterli, "Adaptive wavelet thresholding for image denoising and compression," *IEEE Transactions on Image Processing* **9**(9), pp. 1522–1531, 2000.
[18] S. Chang, B. Yu, and M. Vetterli, "Spatially adaptive wavelet thresholding with context modeling for image denoising," *IEEE Trans. Image Processing* **9**(9), pp. 1532–1546, 2000.
[19] A. Achim, A. Bezerianos, and P. Tsakalides, "Novel bayesian multiscale method for speckle removal in medical ultrasound images," *IEEE Trans. Med. Imag.* **20**, pp. 772–783, 2001.
[20] C. L. Nikias and M. Shao, *Signal Processing with Alpha-Stable Distributions and Applications*, Wiley-Interscience, 1995.
[21] U. Grenander and A. Srivastava, "Probability models for clutter in natural images," *IEEE PAMI* **23**(4), pp. 424–429, 2001.
[22] A. Srivastava, X. Liu, and U. Grenander, "Universal analytical forms for modeling image probabilities," *IEEE PAMI* **24**(9), pp. 1200–1214, 2002.
[23] J. Fadili, "Analytical form for a bayesian wavelet estimator of images using the bessel k forms densities," *submitted for publication* , 2003.
[24] M. Abramowitz and I. A. Stegun, *Handbook of mathematical functions*, Dover Publications, New York, 1972.
[25] I. Gradshteyn and I. Ryzhik, *Table of integrals, series and products*, Academic Press, a. jeffrey ed., 1980.
[26] M. Fisz, "The limiting distribution function of two independent random variables and its statistical application," *Colloquium Mathematicum* **3**, pp. 138–146, 1955.
[27] J. Mathieu, "Transformée en ondelettes et régression non-paramétrique dans un contexte bayesien," Master's thesis, Ecole Nationale Supérieure d'Ingénieur, Caen, 2002.