BAYESIAN DENOISING OF GENERALIZED POISSON PROCESSES WITH FINITE RATE OF INNOVATION

Arash Amini, Ulugbek Kamilov and Michael Unser

Biomedical Imaging Group, EPFL, Lausanne, Switzerland

ABSTRACT

We investigate the problem of the optimal reconstruction of a generalized Poisson process from its noisy samples. The process is known to have a finite rate of innovation since it is generated by a random stream of Diracs with a finite average number of impulses per unit interval. We formulate the recovery problem in a Bayesian framework and explicitly derive the joint probability density function (pdf) of the sampled signal. We compare the performance of the optimal Minimum Mean Square Error (MMSE) estimator with common regularization techniques such as ℓ_1 and Log penalty functions. The simulation results indicate that, under certain conditions, the regularization techniques can achieve a performance close to the MMSE method.

Index Terms— Compound Poisson Process, Finite Rate of Innovation, MMSE, Sparsity, TV Regularization.

1. INTRODUCTION

The concept of sparsity has motivated a number of important signal-processing developments during the past decade. The resulting algorithms are based on solid variational principles (minimization of the ℓ_1 norm), but they are predominantly deterministic. Motivated by the classical equivalence between MMSE signal estimation under the Gaussian hypothesis, MAP, and least-squares regression with a quadratic ℓ_2 regularization, there is a strong incentive to formulate the recovery of sparse signals in a Bayesian framework. The recent class of signal models proposed in [1] is ideally suited for this task because it specifies sparse processes as solutions of stochastic differential equations, in complete analogy with the classical theory of Gaussian stationary processes. The main twist is to replace the traditional white Gaussian noise excitation by impulsive noise (sparse innovation). In particular, when the differential system is unstable, this allows for the generation of piecewise-smooth signals that are the natural extension of the compound Poisson process, which is sparse, but also non-stationary.

The generalized Poisson processes characterized in [2] are the stochastic counterparts of the signals with Finite Rate of Innovation (FRI) [3]. The latter family includes nonuniform splines and piecewise-polynomial functions as particular cases. A typical FRI signal has the form $\sum_{n\in\mathbb{Z}} c_n\varphi(t-t_n)$ where $\{t_n\}$ and $\{c_n\}$ are random sequences; several choices are available for the function φ, including the Dirac delta distribution and compact-support kernels [3, 4]. FRI signals can be perfectly reconstructed from their generalized samples obtained by uniform sampling of a prefiltered version of the signal. Originally, Gaussian and ideal prefilters were used for this purpose; other options such as causal filters with rational Fourier transform were introduced later [4]. In the recent framework of [1, 2], the linear operators acting on the Dirac stream (innovations) play a role similar to the continuous-domain prefilter in the FRI setup.

In this paper, we focus on estimating the sample values of a signal with finite rate of innovation. In other words, instead of estimating the FRI parameters c_n, t_n, we aim at recovering the samples of the signal directly from the noisy measurements. We use the general stochastic framework of [1] and present results for the special case of FRI signals. Thanks to the statistical framework, we derive the joint probability density functions (pdf) of both noiseless and noisy samples. The main contribution of the paper is to factorize the joint pdfs by employing the structure of the linear operators acting on the innovation. The factorization helps us to efficiently implement the MMSE estimator by benefiting from message-passing algorithms. We then use this MMSE estimator as gold-standard for evaluating the performance of other denoising methods. With simulations, we show that common sparsity-promoting regularizers such as the ℓ_1-norm approach the MMSE performance under appropriate conditions.

2. SIGNAL MODEL

We depict in Fig. 1 the general model for the FRI signals considered in this paper. The input white noise $w(x) = \sum_k c_k\delta(x-t_k)$ is a random stream of Diracs, where the amplitudes $\{c_k\}$ are identically and independently distributed (i.i.d.), and where the time instants t_k follow a point process with mean λ. In plain words, there are on average $(b-a)\lambda$ Diracs in the interval $[a,b]\subset\mathbb{R}$, and the number and location of the Diracs in two disjoint intervals are independent. This independence, in addition to the i.i.d. nature of the c_ks, imply that the innovation process $w(x)$ is white (impulsive Poisson noise). It is shown in [2] that, for symmetric probability distribution of c_ks (p_c) and for rapidly decaying test functions $\varphi(x)$, we have

$$\begin{aligned}\mathcal{Z}_w(\varphi) &\triangleq \mathscr{E}\{e^{j\langle w,\varphi\rangle}\}\\ &= \exp\Big(-\lambda\int_{\mathbb{R}}\big(1-\hat{p}_c(\varphi(x))\big)\mathrm{d}x\Big),\end{aligned}\qquad(1)$$

where the scalar product $\langle w,\varphi\rangle$ is a random variable and $\hat{p}_c$ is the characteristic function of the random variables $\{c_k\}$.

The signals of interest are linked to the Dirac stream (innovation) through the linear operators L and L^{-1}: By convention, the whitening operator is denoted by L while the shaping operator is represented by L^{-1}. In fact, if L^{-1} corresponds to a filter (linear shift-invariant operator), its impulse response plays the same role as the smoothing kernel in the FRI sampling. In this paper, we restrict the operator L to be a differential operator of the form $\sum_{i=0}^{n} a_i\mathrm{D}^i$, where D stands for the first-order derivative ($\frac{\mathrm{d}}{\mathrm{d}x}$) and a_is are constants. For $n\geq 1$, the filter is not uniquely invertible and we need to impose linear boundary conditions to form L^{-1}. Note that, for the

The work is supported by the FUN-SP ERC grant.

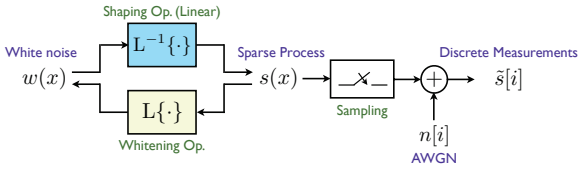

Fig. 1. The generation of the stochastic process $s(x)$ with finite rate of innovation based on the compound Poisson white noise $w(x)$.

mentioned operator L, there is a unique causal Green function $\rho_{\mathrm{L}}(x)$ such that $\mathrm{L}\rho_{\mathrm{L}}(x) = \delta(x)$. The linear shift-invariant operator $\mathrm{L}_{\mathrm{LSI}}^{-1}$ defined by the impulse response $\rho_{\mathrm{L}}(x)$ is one of several possibilities for L^{-1}, since, $\mathrm{L}\mathrm{L}_{\mathrm{LSI}}^{-1}$ is the identity operator. It is not hard to check that $\mathrm{L}_{\mathrm{LSI}}^{-1}$ corresponds to the *initially at-rest* boundary condition (at $-\infty$). It is common to refer to the stable $\mathrm{L}_{\mathrm{LSI}}^{-1}$ as the inverse of L without mentioning the boundary condition; however, when $\mathrm{L}_{\mathrm{LSI}}^{-1}$ is unstable, there is no alternative than introducing a boundary condition that destroys the property of shift-invariance.

For a given input, the difference between the outputs of $\mathrm{L}_{\mathrm{LSI}}^{-1}$ and any other L^{-1} is a signal in the null space of L that depends on the boundary conditions. Here, we assume that the linear boundary conditions are such that the mentioned additive null space part is independent of $w(x)$ for $x > 0$. In other words, $\big(\mathrm{L}^{-1} - \mathrm{L}_{\mathrm{LSI}}^{-1}\big)w(x)$ is statistically independent of $w(x)u(x)$. This means that the boundary conditions are fully determined by observing $w(x)$ or, equivalently, ${\mathrm{L}_{\mathrm{LSI}}}^{-1}w(x)$, at $x \in (-\infty, 0]$.

Results in spline theory guarantee that, for the introduced differential L operators, one can find a discretized shift-invariant operator $\mathrm{L_d}$ such that $\varphi_{\mathrm{L}}(x) = \mathrm{L_d}\mathrm{L}^{-1}\delta(\mathrm{x})$ is well-localized (compact support). More precisely, for a differential operator of order n, the impulse response of $\mathrm{L_d}$ is of the form $\sum_{k=0}^{n} d_{\mathrm{L}}[k]\delta(x-k)$ where $\{d_{\mathrm{L}}[k]\}_{k=0}^{n}$ represents an FIR filter with $n+1$ taps, and the impulse response $\varphi_{\mathrm{L}}(x)$ of $\mathrm{L_d}\mathrm{L}^{-1}$ is supported only on $[0, n)$ [5]. The impulse response $\varphi_{\mathrm{L}}(x)$ is usually known as the B-spline.

The last thing to mention about the model is the measurement procedure. We assume that the stochastic process $s(x)$ is obtained by "integrating" the Dirac stream. It is sampled uniformly at the integers and corrupted in the discrete domain by a Gaussian noise that is independent of the signal. Now, the denoising problem studied in this paper is to estimate the noise-free samples $\{s(x=i)\}_{i=0}^{m}$ from the noisy measurements $\{\tilde{s}[i]\}_{i=0}^{m}$.

3. PROBABILITY FACTORIZATION

In order to employ the statistics of the model for estimating the noise-free samples, we need to obtain the joint distribution of $s[i]$s (the samples of $s(x)$ at the integers). For this purpose, we define the generalized increment process of s associated to the operator L to be:

$$\begin{aligned} u[i] &= \sum_{k=0}^{n} d_{\mathrm{L}}[k]s[i-k] = \mathrm{L_d}\, s(x)\big|_{x=i} \\ &= \mathrm{L_d}\mathrm{L}^{-1}w(x)\big|_{x=i} = (w * \varphi_L)(x)\big|_{x=i}. \end{aligned} \tag{2}$$

Lemma 1 *$u[i+N]$ is statistically independent of $u[i]$ and $s[i]$ for $N \geq n$ and $i \geq 0$.*

Proof From (2), we have

$$u[i] = (w * \varphi_{\mathrm{L}})(x)\big|_{x=i} = \langle w, \varphi_{\mathrm{L}}(i-\cdot)\rangle. \tag{3}$$

Thus, $u[i+N]$ and $u[i]$ are random variables obtained through the inner product of the white noise with the B-spline functions $\varphi_{\mathrm{L}}(i+N-\cdot)$ and $\varphi_{\mathrm{L}}(i-\cdot)$, respectively. Note that φ_{L} is supported on an interval of length n, which shows that $\varphi_{\mathrm{L}}(i+N-\cdot)$ and $\varphi_{\mathrm{L}}(i-\cdot)$ have no support in common. Therefore, $u[i+N]$ and $u[i]$ are independent due to the whiteness of the innovation.

For the second part, recall that $s(x)$ can be obtained by filtering the innovation by the causal filter $\mathrm{L}_{\mathrm{LSI}}^{-1}$, and adding a term in the null space of L which is fully determined by $w(x)$ for $x \leq 0$. This shows that, for $i \geq 0$, the statistics of $s[i]$ are fully determined by $w(x)$ for $x \leq i$. On the other hand, as (3) indicates, the statistics of $u[i+N]$ depend only on $w(x)$ for $i+N-n < x \leq i+N$. Since the two intervals are disjoint for $N \geq n$, $s[i]$ and $u[i+N]$ are statistically independent. ■

The following theorem shows the main result for factorizing the joint distribution of $\{s[k]\}_{k=0}^{m}$.

Theorem 1 *For the mentioned model, we have that*

$$\begin{aligned} &p_s\big(s[m], \ldots, s[0]\big) = \\ &\qquad p_s\big(s[2n-2], \ldots, s[0]\big) \\ &\qquad \times \prod_{k=2n-1}^{m} \big|d_{\mathrm{L}}[0]\big|\, p_u\Big(u[k] \;\Big|\; \{u[k-i]\}_{i=1}^{n-1}\Big). \end{aligned} \tag{4}$$

Proof For $k \geq n$, let us first define

$$\mathbf{v}[k] = \Big[u[k] \;\; \ldots \;\; u[n] \;\; s[n-1] \;\; \ldots \;\; s[0]\Big]^T. \tag{5}$$

It is not hard to check that

$$\mathbf{v}[k] = \mathbf{L}_{(k+1)\times(k+1)}\Big[s[k] \;\; \ldots \;\; s[0]\Big]^T, \tag{6}$$

where $\mathbf{L}_{(k+1)\times(k+1)}$ is an upper-triangular matrix that is formed by the taps of the FIR filter as

$$L_{i,j} = \begin{cases} d_{\mathrm{L}}[j-i] & 1 \leq i \leq j \leq n+i \leq k+1 \\ 1 & k+2-n \leq i = j \leq k+1 \\ 0 & \text{otherwise.} \end{cases} \tag{7}$$

Since $\mathbf{L}_{(k+1)\times(k+1)}$ is invertible (upper-triangular matrix with non-zero elements on the main diagonal), the linear equation in (6) suggests that

$$p_{s,u}\big(\mathbf{v}[k]\big) = \frac{p_s\big(s[k], \ldots, s[0]\big)}{\big|\det\ \mathbf{L}_{(k+1)\times(k+1)}\big|}. \tag{8}$$

By using Lemma 1, for $k \geq 2n-1$ we have

$$\begin{aligned} \frac{p_{s,u}\big(\mathbf{v}[k]\big)}{p_{s,u}\big(\mathbf{v}[k-1]\big)} &= p_{s,u}\big(u[k] \,\big|\, \mathbf{v}[k-1]\big) \\ &= p_u\Big(u[k] \;\Big|\; \{u[k-i]\}_{i=1}^{n-1}\Big). \end{aligned} \tag{9}$$

To conclude the theorem, it is now sufficient to multiply equations of the form (9) for $k = 2n-1, \ldots, m$, and replace $p_{s,u}\big(\mathbf{v}[m]\big)$ and $p_{s,u}\big(\mathbf{v}[2n-2]\big)$ by their equivalent forms as in (8). Note that, due to the upper-triangular structure, the determinant of $\mathbf{L}_{(k+1)\times(k+1)}$ is equal to $d_{\mathrm{L}}[0]^{k+1-n}$. ■

Theorem 1 provides the factorization of the general a priori m-dimensional joint distribution into n-dimensional conditional distributions of $u[k]$s where n reflects the order of the operator, and a single $(2n-2)$-dimensional term that includes the boundary conditions. The joint probability density of $u[k], \ldots, u[k+i]$ can be derived from (1) as explained in Proposition 1.

Proposition 1

$$p_u\big(\{u[k+\alpha]\}_{\alpha=0}^{i}\big) = \mathcal{F}^{-1}_{\{\omega_\alpha\}}\left\{\mathrm{e}^{-\lambda\int_{\mathbb{R}}\left[1-\hat{p}_c\left(\sum_{\alpha=0}^{i}\omega_\alpha\varphi_{\mathrm{L}}(\alpha-x)\right)\right]\mathrm{d}x}\right\}\Big|_{u[\alpha]_{\alpha=k}^{k+i}}, \quad (10)$$

where $\mathcal{F}^{-1}_{\{\omega_\alpha\}}\{\cdot\}\big|_{u[\alpha]_{\alpha=k}^{k+i}}$ *denotes the* $(i+1)$*-dimensional inverse Fourier transform with respect to* ω_α*s at* $\big(u[k],\ldots,u[k+i]\big)$.

Proof We show the result by using $\sum_{\alpha=0}^{i}\omega_\alpha\varphi_{\mathrm{L}}(k+\alpha-x)$ as $\varphi(x)$ in (1):

$$\begin{aligned}
&\mathcal{Z}_w\Big(\sum_{\alpha=0}^{i}\omega_\alpha\varphi_{\mathrm{L}}(k+\alpha-x)\Big)\\
&= \mathscr{E}\Big\{\mathrm{e}^{\mathrm{j}\sum_{\alpha=0}^{i}\omega_\alpha\langle w,\varphi_{\mathrm{L}}(k+\alpha-\cdot)\rangle}\Big\} = \mathscr{E}\Big\{\mathrm{e}^{\mathrm{j}\sum_{\alpha=0}^{i}\omega_\alpha u[k+\alpha]}\Big\}\\
&= \int_{u[\alpha]_{\alpha=k}^{k+i}} p_u\big(\{u[\alpha]\}_{\alpha=k}^{k+i}\big)\mathrm{e}^{\mathrm{j}\sum_{\alpha=0}^{i}\omega_\alpha u[k+\alpha]}\\
&= \mathcal{F}_{u[\alpha]}\Big\{p_u\big(\{u[k+\alpha]\}_{\alpha=0}^{i}\big)\Big\}\Big|_{\{\omega_\alpha\}_{\alpha=0}^{i}}. \qquad (11)
\end{aligned}$$

Thus, the desired joint pdf is achieved by taking the inverse Fourier transform of $\mathcal{Z}_w\Big(\sum_{\alpha=0}^{i}\omega_\alpha\varphi_{\mathrm{L}}(k+\alpha-x)\Big)$ with respect to w_αs. By employing the explicit form of $\mathcal{Z}_w$ from (1), we obtain (10). ■

Now that we can compute the joint pdf of $u[i]$s, we can simply compute the conditional pdfs involved in Theorem 1 by writing them as a fraction of two joint pdfs.

For the posterior distribution, by using Bayes' rule and the fact that the additive noise is Gaussian, we obtain

$$p_s\big(\{s[k]\}_{k=0}^{m}\,\big|\,\{\tilde{s}[k]\}_{k=0}^{m}\big) = \frac{\mathrm{e}^{\frac{-1}{2\sigma_n^2}\sum_{k=0}^{m}(\tilde{s}[k]-s[k])^2}\,p_s\big(s[m],\ldots,s[0]\big)}{(2\pi\sigma_n^2)^{(m+1)/2}\,p_{\tilde{s}}\big(\tilde{s}[m],\ldots,\tilde{s}[0]\big)}. \quad (12)$$

The conditional distribution (12) specifies our signal recovery problem completely, since the MMSE estimation of the noise-free values $\{s[k]\}$ is the same as their expected value conditioned to the noisy observations $\{\tilde{s}[k]\}$.

4. NUMERICAL RESULTS

For simulations, we consider the simple case of $n=1$. The simplest scenario for $n=1$ is when $\mathrm{L}=\mathrm{D}=\frac{\mathrm{d}}{\mathrm{d}x}$ and the inverse is $\mathrm{L}^{-1}w(x)=\int_0^x w(\tau)\mathrm{d}\tau$. This choice of L^{-1} implies $s(0)=0$ as the boundary condition; clearly, this boundary condition is independent of $w(x)$ for $x>0$ (required condition). The processes obtained in this fashion are called Lévy processes in general; for the special case of impulsive Poisson innovation, the process consists of piecewise-constant signals and is called compound Poisson. Figure 2 shows a part of a realization for this process when $\lambda=0.3$ (average number of Diracs per unit interval) and the jump amplitudes follow a standard Gaussian distribution.

It is well-known that the discrete counterpart of the continuous-time differentiation is the finite difference. More precisely, $d_{\mathrm{L}}[0]=1$ and $d_{\mathrm{L}}[1]=-1$. Here, the B-spline $\varphi_{\mathrm{L}}(x)$ here is the rectangular function supported on $[0,1)$ with constant unit amplitude. For this special case, Theorem 1 can be rewritten as

$$p_s\big(s[m],\ldots,s[0]\big) = \prod_{k=1}^{m} p_u\big(u[k]\big)\times p_s\big(s[0]\big), \quad (13)$$

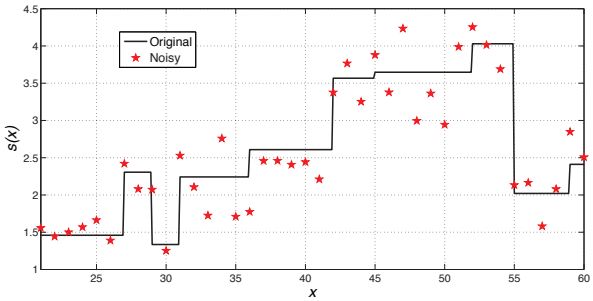

Fig. 2. A realization of the piecewise constant signal and the noisy measurements. The average number of Diracs per unit interval (λ) of the impulsive innovation process is 0.3 and the jumps follow a standard Gaussian distribution.

which, in turn, results in

$$p_s\big(\{s[k]\}_{k=0}^{m}\,\big|\,\{\tilde{s}[k]\}_{k=0}^{m}\big) = \frac{\mathrm{e}^{-\frac{\sum_{k=0}^{m}(\tilde{s}[k]-s[k])^2}{2\sigma_n^2}}\prod_{k=1}^{m}p_u\big(s[k]-s[k-1]\big)\times p_s\big(s[0]\big)}{(2\pi\sigma_n^2)^{(m+1)/2}\,p_{\tilde{s}}\big(\tilde{s}[m],\ldots,\tilde{s}[0]\big)}, \quad (14)$$

where $p_s\big(s[0]\big)=\delta\big(s[0]\big)$. Also, for the pdf of the increment process, the general form in Proposition 1 simplifies to

$$\begin{aligned}
p_u(u) &= \mathcal{F}_{\omega}^{-1}\Big\{\mathrm{e}^{-\lambda\int_{\mathbb{R}}\left[1-\hat{p}_c\left(\omega\varphi_{\mathrm{L}}(-x)\right)\right]\mathrm{d}x}\Big\}\Big|_u\\
&= \mathcal{F}_{\omega}^{-1}\Big\{\mathrm{e}^{-\lambda\left(1-\hat{p}_c(\omega)\right)}\Big\}\Big|_u\\
&= \mathrm{e}^{-\lambda}\Big(\delta(u)+\sum_{k=1}^{\infty}\frac{\lambda^k}{n!}\big(\underbrace{p_c*\cdots*p_c}_{\text{k times}}\big)(u)\Big). \quad (15)
\end{aligned}$$

Due to the simple factorization of p_u in (15), we are able to apply the message-passing algorithm which is an efficient tool for computing the marginal distributions $p\big(s[i]\,\big|\,\{\tilde{s}[k]\}_k\big)$ from the joint conditional pdf. In plain words, the message-passing algorithm is an elegant iterative way of integrating the joint pdf to obtain all the marginals simultaneously with an almost-minimal computational complexity. The general concepts of the message-passing algorithm can be found in [6] while the implementation details for the specific case of Lévy processes are available in the companion paper [7]. The associated graph for $n=1$ (Lévy) is a tree, which guarantees the convergence to the desired marginal pdfs in a finite number of iterations [6].

The performance of various denoising techniques including the MMSE with the message-passing implementation is shown in Fig. 3 and 4. They plot $10\log_{10}\frac{\sum_k\left|s[k]-\tilde{s}[k]\right|^2}{\sum_k\left|s[k]-\hat{s}[k]\right|^2}$ under various additive noise powers which is the improvement of the SNR in dB; the curves are averaged over 500 different realizations and the innovation (w) has, on average, 0.5 Diracs per unit interval (λ).

For alternative denoising approaches, we have considered regularization methods that promote sparsity. The piecewise constant signals can be viewed as constant signals with sparse (infrequent) jumps. Thus, the finite differences $\{u[k]\}$ are sparse, in the sense that, with some non-zero probability $\mathrm{e}^{-\lambda}$, where λ is the average number of Diracs per unit interval, the differences vanish. We use

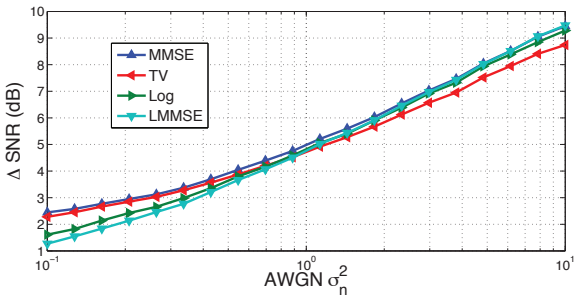

Fig. 3. SNR improvement of various denoising methods vs. the variance of the additive white Gaussian noise. The impulsive Poisson innovation contains $\lambda = 0.5$ Diracs on average per unit interval whose amplitudes follow a standard Gaussian distribution.

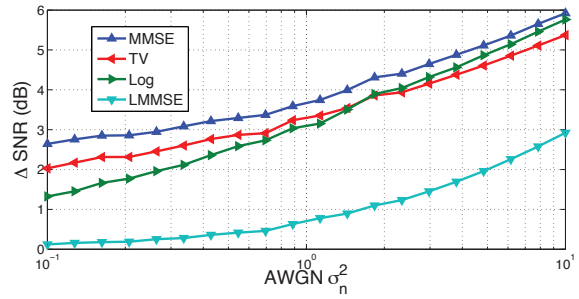

Fig. 4. SNR improvement of various denoising methods vs. the variance of the additive white Gaussian noise. The impulsive Poisson noise contains $\lambda = 0.5$ Diracs on average per unit interval whose amplitudes follow a standard Cauchy distribution.

the following form for the regularization methods:

$$\{\hat{s}[k]\}_{k=0}^{m} = \arg\min_{\{s[k]\}} \|\mathbf{s} - \tilde{\mathbf{s}}\|_2^2 + \tau \sum_{k=1}^{m} \Psi\big(s[k] - s[k-1]\big). \quad (16)$$

Here, we restrict $\Psi(x)$ to be one of the $|x|$, x^2, and $\log(1 + x^2)$ functions, which result in Total Variation (TV), Linear Minimum Mean-Square Error (LMMSE), and Logarithmic (Log) regularizers, respectively. In the above formulation, the optimum τ depends on the additive noise power. For setting τ for a given additive noise power, in the simulations we tuned τ by having oracle knowledge about the noise-free values for a number of realizations and then, we used the average τ for the rest of the realizations. The Log regularizer is known to be an ℓ_1-ℓ_0 relaxation for recovering sparse vectors [8]. This regularizer is not convex; thus, since we use gradient descent methods, we might get trapped into the local minima of the cost function.

We show in Fig. 3 the results for a Gaussian distribution of the Dirac amplitudes. We see that, at small noise powers, the TV regularizer almost achieves the MMSE performance. On the other hand, at large noise powers the statistics of the noisy samples are mainly governed by the Gaussian noise. There, the curve of the LMMSE method (which is optimal for Gaussian processes) coincides with that of MMSE. The results in Fig. 4 are based on the Cauchy distribution of the amplitudes. This distribution is heavy-tailed and, even at large Gaussian noise powers, the dominant distribution is still Cauchy. This, in fact, explains the poor performance of the LMMSE method. Finally, we see that the Log regularizer, which is somehow matched to the tail of the Cauchy distribution, better follows the MMSE performance at large noise powers.

5. CONCLUSION

In this paper, we presented a stochastic framework for studying FIR signals and we concentrated on the denoising problem. Unlike the usual approach of estimating the innovation parameters, we directly focused on the estimation problem by factorizing the joint probability distributions. Thanks to the factorized pdf, we were able to implement the MMSE estimator by using the message-passing algorithm, thereby obtaining a gold standard against which to compare other algorithms. In particular, we used our model to compare the performance of regularization techniques based on sparsity constraints (such as ℓ_1 norm) through simulations. Results indicate that, under certain conditions, the regularization techniques can almost achieve the MMSE performance.

Acknowledgment: The authors would like to thank A. Bourquard, N. Chenouard, H. Kirshner, S. Lefkimmiatis, P. Pad, and P. Tafti for their support regarding the simulation results. The authors are also indepted to P. Thévenaz for his fruitful comments.

6. REFERENCES

[1] M. Unser, P. Tafti, and Q. Sun, "A unified formulation of Gaussian vs. sparse stochastic processes: Part I—Continuous-domain theory," *arXiv:1108.6150v1*.

[2] M. Unser and P. Tafti, "Stochastic models for sparse and piecewise-smooth signals," *IEEE Transaction on Signal Processing*, vol. 59, no. 3, pp. 989–1006, March 2011.

[3] M. Vetterli, P. Marziliano, and T. Blu, "Sampling signals with finite rate of innovation," *IEEE Transaction on Signal Processing*, vol. 50, no. 6, pp. 1417–1428, June 2002.

[4] P. L. Dragotti, M. Vetterli, and T. Blu, "Sampling moments and reconstructing signals of finite rate of innovation: Shannon meets Strang-Fix," *IEEE Transaction on Signal Processing*, vol. 55, no. 5, pp. 1741–1757, May 2007.

[5] M. Unser and T. Blu, "Cardinal exponential splines: Part I—Theory and filtering algorithms," *IEEE Transaction on Signal Processing*, vol. 53, no. 4, pp. 1425–1438, April 2005.

[6] H. A. Loeliger, J. Dauwels, J. Hu, S. Korl, L. Ping, and F. R. Kschischang, "The factor graph approach to model-based signal processing," *Proceedings of the IEEE*, vol. 95, no. 6, pp. 1295–1322, June 2007.

[7] U. Kamilov, A. Amini, and M. Unser, "MMSE denoising of sparse Lévy processes via message passing," *Submitted to ICASSP2012*, 2012.

[8] E. J. Candés, M. B. Wakin, and S. P. Boyd, "Enhancing sparsity by reweighted ℓ_1 minimization," *Journal of Fourier Analysis and Applications*, vol. 14, no. 5–6, pp. 877–905, 2008.