A BERNOULLI-GAUSSIAN APPROACH TO THE RECONSTRUCTION OF NOISY SIGNALS WITH FINITE RATE OF INNOVATION

Andrea Ridolfi †, *Irena Maravić* †, *Julius Kusuma*‡, *Martin Vetterli* †,∗

† Ecole Polytechnique Fédérale de Lausanne, I&C - LCAV, Lausanne, Switzerland;
{andrea.ridolfi,irena.maravic,martin.vetterli}@epfl.ch
‡ Massachusetts Institute of Technology, LIDS , Cambridge, MA, USA;
kusuma@mit.edu
∗ University of California at Berkeley, EECS, Berkeley, CA, USA.

ABSTRACT

Recently, it was shown that a large class of non-bandlimited signals that have a finite rate of innovation, such as streams of Diracs, non-uniform splines and piecewise polynomials, can be perfectly reconstructed from a uniform set of samples taken at the rate of innovation [1]. While this is true in the noiseless case, in the presence of noise the finite rate of innovation property is lost and exact reconstruction is no longer possible.

In this paper we consider the problem of reconstructing such signals when noise is present. We focus on the case when a discrete-time signal is made up of a sum of weighted Diracs and we propose a stochastic reconstruction method based on the Bernoulli-Gauss model and on a maximum *a posteriori* optimization. Our approach is numerically stable and yields precise reconstruction by sampling the signal way below the Nyquist rate, significantly outperforming commonly used subspace methods [2, 3].

Applications of our method can be found in acquisition and processing of signals in wideband communication systems, such as ultra-wideband (UWB) systems.

1. INTRODUCTION

Sampling theory has undergone a strong research revival over the past decade, and more advanced formulations of Shannon's original theory have been developed, with immediate relevance to signal processing and communications. It was recently shown [1] that it is possible to develop sampling schemes for a large class of non-bandlimited signals that have a parametric representation with a finite number of degrees of freedom, the so-called signals with a finite rate of innovation. Such signals are uniquely represented by a set of uniform samples taken at (or above) the rate of innovation, typically much lower than the traditional Nyquist rate [4], leading to very efficient, critical sampling schemes for a large class of signals.

In the presence of noise, the previous developed theory is no longer applicable. However, it seems intuitive that by sampling above the rate of innovation and using an appropriate estimation procedure, it is possible to recover the underlying signal with high precision. In this work we focus on the retrieval of discrete-time streams of K weighted Diracs over a period of L samples, which has rate of innovation $(2K + 1)/L$. As in [1], the data we use for the estimation process consists of sub-samples of the low pass approximation of the streams of Diracs. We precise that in [1] authors adopt a continuous-time model. However, our discrete approach can approximate it at a very fine level. Reconstruction for noiseless, discrete-time, finite rate of innovation signals have been studied in [5].

Our contribution is twofold: we propose a stochastic approach that models the stream of Diracs as a Bernoulli-Gaussian process [6] and derive an optimization criterion in the frequency domain that has an *algebraic structure* with an *explicit relation* between bandwidth of the low pass filter and number of samples to be used and that assures *numerical stability* in the optimization process.

Numerical results show that it is still possible to precisely reconstruct the stream of Diracs by sampling near the rate of innovation but still well below the Nyquist rate. A comparison with subspace methods, commonly used in the context of high-resolution harmonic retrieval [2, 3], shows that the latter are outperformed by the Bernoulli-Gaussian approach. As far as we know,this method provides the best results, significantly better than the mentioned commonly used approaches. However, such a performance is achieved at a price of a high numerical complexity. It is then a natural benchmark candidate in the analysis of real-time algorithms.

Applications of our method can be found in the analysis of signal processing methods for wideband communication systems, such as ultra-wideband systems. In such systems, the rate of innovation can be up to several orders of magnitude lower than the Nyquist rate. Therefore, sampling close to such a rate lead to very efficient processing schemes.

The outline of the paper is as follows. In Section 2, we introduce the stochastic approach based on a Bernoulli-Gaussian model. In Section 3, we formulate the optimization problem in the frequency domain. Section 4 compares the Bernoulli-Gaussian approach with the commonly used subspace methods. In Section 5 we explore some attractive applications, focusing in particular to the problem of ultra-wideband communication.

2. NOISY SIGNAL MODEL: A BERNOULLI-GAUSSIAN APPROACH

Consider a discrete-time stream of K Diracs affected by noise

$$x_l = \sum_{k=1}^{K} a_k \delta(l - l_k) + \epsilon_l, \quad l = 0, \dots, L-l \qquad (1)$$

where $l_k \in \{0, \dots, L-1\}$ and $a_k \in \mathbb{R}$, are respectively, the positions and the amplitudes of the K spikes, while ϵ_l, is the noise. The latter is modeled as a the realization of a Gaussian white noise with variance σ_ϵ^2, noted $\mathcal{E}_l$ (in the following we will note a random variable and its realization with a small and a capital letter respectively).

The stream of Diracs is modeled as one realization of a Bernoulli Gaussian process. Such a process can be thought of as a sequence of independent pairs of random variables (Q_l, A_l), $l = 0, \dots, L-1$, where Q_l is a binary random variable modeling the presence of a spike at position l, with $\mathrm{P}(Q_l = 1) = \lambda$, for all l, and A_l is a zero mean Gaussian random variable modeling the amplitude, with variance $q_l \sigma_a^2$, where q_l it the realization of the binary random variable Q_l. Hence, if for some l, we have $q_l = 1$, then there is a Dirac at position l and the corresponding amplitude is the realization of a centered Gaussian variable with variance σ_a^2. Otherwise ($q_l = 0$), there is no Dirac at position l, and indeed the signal amplitude is the realization of a zero-mean degenerated Gaussian variable with a null variance which is equal to zero with probability one.

Following the approach of [1], K positions and amplitudes of Diracs are estimated from samples of the low pass approximation of the noisy stream of Diracs, that is,

$$y_m = \sum_{j=0}^{L-1} h_{m-l} \left(x_l + \epsilon_l \right), \quad m = 0, \dots, M-1, \qquad (2)$$

where $2K+1 \leq M \leq L$. We recall that $2K+1$ is the minimum number of low-passed samples required to perfectly reconstruct a stream of K Dirac in the absence of noise and it corresponds to the degree of freedom of such a signal [1]. As we shall see in the following, M is an even number fixed by the lowpass filter bandwidth $[-(M-1)/2, (M-1)/2]$. It represents the number of non-zero Fourier coefficient of the lowpass approximation and, therefore, the number of low-pass samples used for the reconstruction. Since $M \leq L$, we shall refer to y_m, as sub-samples of the low-pass approximation.

Let $\boldsymbol{y} = \{y_0, \dots, y_{M-1}\}$, $\boldsymbol{x} = \{x_0, \dots, x_{L-1}\}$, $\boldsymbol{\epsilon} = \{\epsilon_0, \dots, \epsilon_{L-1}\}$ and define the matrix $H_{(l,m)} = h_{m-l}$, where $l = 0, \dots, L-1, \quad m = 0, \dots, M-1$. We can then compactly write (2) as,

$$\boldsymbol{y} = H \left(\boldsymbol{x} + \boldsymbol{\epsilon} \right). \qquad (3)$$

3. MAP ESTIMATION OF POSITIONS AND AMPLITUDES

Positions and amplitudes are sequentially estimated from $\boldsymbol{y}$ using a maximum *a posteriori* approach [6, 7]: first we estimate the position through maximization of the posterior marginal likelihood

$$\widehat{\boldsymbol{q}} = \operatorname{argmax} \mathrm{P}\left(\boldsymbol{q} \mid \boldsymbol{y} \right) \propto f\left(\boldsymbol{y} \mid \boldsymbol{q} \right) \qquad (4)$$

and then the amplitudes are determined through maximization the following *a posteriori* likelihood

$$\widehat{\boldsymbol{a}} = \operatorname{argmax} f\left(\boldsymbol{a} \mid \boldsymbol{y}, \widehat{\boldsymbol{q}} \right) \propto f\left(\boldsymbol{y} \mid \boldsymbol{a} \right) f\left(\boldsymbol{a} \mid \widehat{\boldsymbol{q}} \right).$$

Such a sequential approach is preferable since a joint MAP estimation of positions and amplitudes gives poor estimates [8]. We remark that since the number K of Diracs is known, the *a priori* Bernoulli distribution $\mathrm{P}(\boldsymbol{q})$ of the positions does not appear in expression (4).

Since lowpass filtering maintains Gaussian characteristics, determination of $\boldsymbol{a}$ is a classical MAP estimation problem under linear and Gaussian assumption and the solution can be expressed in closed form [6, 7]. Therefore, in the following we will focus on the estimation of $\boldsymbol{q}$.

In this work we formulate the optimization problem in the frequency domain. Using this approach we take into account only the non-zero Fourier components, thus avoiding numerical instability in the optimization process. We adopt the notation $\boldsymbol{\mathcal{X}}$ and $\boldsymbol{\mathcal{Y}}$ for the Fourier representation of the stream of Diracs and its low pass approximation, respectively. Note that the latter is characterized by M non-zero terms. The transformation from time to frequency domain is performed by mean of the Vandermonde matrix

$$V_{(l,m)} = e^{-i2\pi lm/L}, \quad i, m = 0, \dots, L-1.$$

Then, in the frequency domain, the noisy stream of Diracs reads $\boldsymbol{\mathcal{X}} = V\boldsymbol{x}$. Knowing $\boldsymbol{q}$, $\boldsymbol{\mathcal{X}}$ can be seen as the realization of a zero-mean Gaussian random vector with covariance matrix given by

$$R = V \Pi V' \sigma_a^2 + V V' \sigma_\epsilon^2$$

where Π is a diagonal matrix, with diagonal equal to $\boldsymbol{q}$. Therefore, the $[-(M-1)/2, (M-1)/2]$ low-pass approximation $\boldsymbol{\mathcal{Y}}$ of the noisy stream of Diracs can be modeled as the realization of a zero-mean Gaussian random vector with covariance matrix given by

$$R = \widetilde{V} \Pi \widetilde{V}' \sigma_a^2 + \widetilde{V} \widetilde{V}' \sigma_\epsilon^2$$

where $\widetilde{V} = \begin{bmatrix} V\left(1 : (M-1)/2, :\right) \\ V\left(L - (M-1)/2 : L, :\right) \end{bmatrix}$.

Then, giving M DFT values of the low pass approximation $\boldsymbol{\mathcal{Y}} = \left\{ \mathcal{Y}_1, \dots, \mathcal{Y}_{(M-1)/2}, \mathcal{Y}_{L-(M-1)/2}, \dots, \mathcal{Y}_L \right\}$, estimation of the positions $\boldsymbol{q}$ is achieved thorough the maximization of the log *a posteriori* likelihood function

$$LLF(\boldsymbol{q}) = -\boldsymbol{\mathcal{Y}} R^{-1} \boldsymbol{\mathcal{Y}} - \log(|R|).$$

This maximization can be globally performed by means of stochastic algorithms, such as simulated annealing, or locally performed by semi-stochastic or deterministic algorithm. Note that in the case of a small number of Diracs, simulated annealing with a finite annealing schedule [9] may be an attractive time-efficient approach. In view of the benchmark use of the numerical results, a direct exhaustive optimization search can provide an easy-to-implement optimization method.

4. NUMERICAL RESULTS AND COMPARISON

In order to illustrate the numerical performances of the method we have presented, we have considered discrete time streams of 2 weighted Diracs in a length 512 signal. The positions are randomly chosen according to a uniform distribution over $[1, 512]$, while the weights are sampled from i.i.d. zero mean Gaussian random variables with unit variance. Therefore, the average signal power is equal to $2/512$. The noise is supposed white and Gaussian and we shall consider three different variances, 2^8, 2^7 and 2^6, corresponding respectively to a SNR of 0db, -6.9db and -19.9db.

Estimation of the position has been achieved thought an exhaustive optimization search, while, as stated in the previous section, the amplitudes have been determined in a closed form.

The results obtained with the Bernoulli-Gaussian approach are compared to to those obtained by classic harmonic retrieval methods, such as ESPRIT, MUSIC and state-space method. As stated in [2], ESPRIT and MUSIC have similar statistical accuracy and in our numerical simulations they have provide very similar results. Therefore, only the ones obtained with ESPRIT are shown. We also consider the annihilating filter method, which is the approach adopted in [1] for estimating the Diracs in the absence of noise.

Each method is tested for the three different values of SNR and five different bandwidths of the low-pass filter, *i.e.* five different numbers of samples. More precisely, the lowest number of samples we consider is the one dictated by the degrees of freedom of the signal, *i.e.* the minimum number of samples required to perfectly reconstruct the signal in the absence of noise [1]. In particular, results are plotted versus the ratio between the number of samples and the number of degrees of freedom.

We focus on the mean square error for the estimation of the positions only. Each method is tested on one thousand signals (ten stream of weighted Diracs, each one added to one hundred realizations of the noise). The corresponding one thousands estimation errors are then averaged and normalized by the length of the signal (relative error), *i.e.*

$$\text{error} = \frac{\sum_{j=1}^{1000} \|\widehat{\boldsymbol{q}}_j - \boldsymbol{q}_j\|_2}{1000} \times \frac{1}{512}.$$

Figure 1 shows the performances of the Bernoulli-Gaussian approach for the three values of SNR. We can see that reconstruction of the positions can still be achieved, with good precision, from samples taken at a rate closed to the one dictated by the number of degrees of freedom of the signal (abscissa 1). In our simulation example this is true in particular for SNRs higher than -13.9db.

Figures 2(a), (b) and (c) compares the performances of the Bernoulli-Gaussian approach with the classical subspace methods, respectively for the three values of SNR. Note that the overall performance of the stochastic method is clearly better than the ones of the classical subspace methods. Such a good performance is expected since with the stochastic approach the noise intrinsically taken into account in the model. Concerning the subspace methods, we can remark that performances do not improve consistently by increasing the number of samples. In particular, the annihilating filter performs quite badly being strongly noise dependent, while ESPRIT is penalized by the estimation of a covariance matrix which critically depends on the number of samples. The state-space method has a good overall performance, if compared to the annihilating filter and ESPRIT.

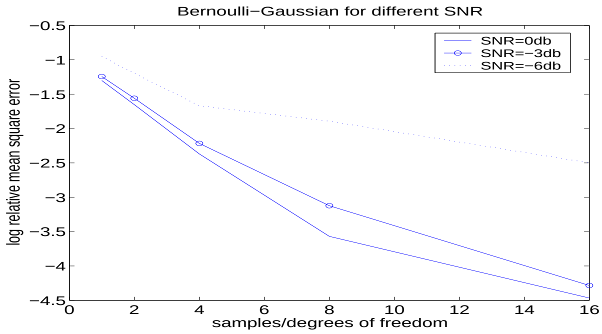

Fig. 1. Log relative mean square error versus the ratio between the number of samples and the number of degrees of freedom, obtained with the Bernoulli-Gaussian approach for three values of SNR: 0db, -6.9db and -19.9db

5. SAMPLING ULTRA WIDE BAND SIGNALS

We now turn our attention to the sampling of Ultra-Wideband (UWB) signals. UWB signals are designed to have very high bandwidth expansion, where the transmitted signal's bandwidth is many times the symbol rate. This in turn poses a challenge for the receiver, as the Nyquist rate required to sample the signal is exceedingly high. We note that the true innovation rate of the signal is dictated by the symbol rate. The framework developed in this work allows one to reliably sample close to the innovation rate or the symbol rate. This was first presented in [10, 11], where it was also demonstrated that noise robustness increases as a function of oversampling beyond the critical rate.

The main idea is as follows: we interpret UWB signals as a train of Diracs carrying the information at the symbol rate. This train of Diracs is then convolved with the

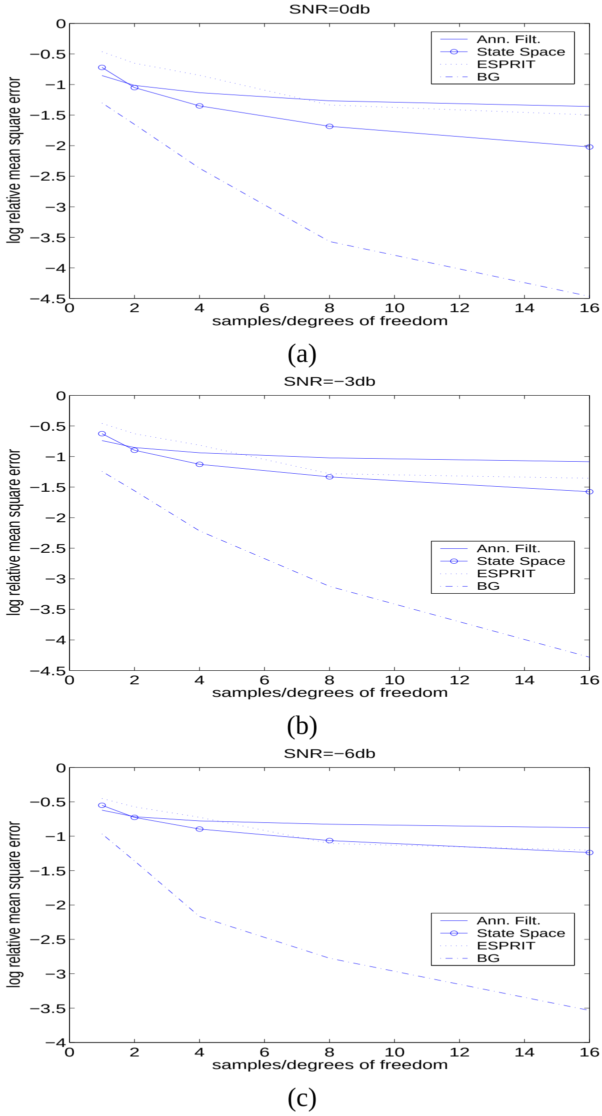

Fig. 2. Log relative mean square error versus the ratio between the number of samples and the number of degrees of freedom, for all the explored methods. From top to bottom, figures depict the results for a SNR respectively equal to 0db, -6.9db and -19.9db

pulse shape, and in turn subjected to the noisy effects of a linear time-invariant channel, and receive filter. We can jointly consider these effects and refer to them as the *compound channel*, as shown in Figure 3. Therefore, through

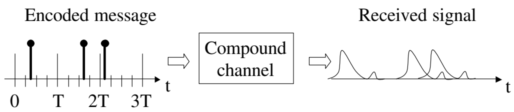

Fig. 3. Illustration of the compound channel.

equalization, the framework for this problem derives from the framework for sampling of signals with a finite rate of innovation with a linear time-invariant kernel [1].

Within the framework of UWB systems, results of Section 4 show that it is possible detect UWB signals by sampling well below the Nyquist rate, but above the rate of innovation. In particular, the Bernoulli Gaussian approach provides an error lower bound for the analysis of real time UWB detection methods.

Related work for UWB system with computational efficient methods is explored in [12].

6. CONCLUDING REMARKS

We have explored the retrieval of noisy weighted Diracs in the light of the recent work on sampling noiseless signals with finite rate of innovation [1]. The stochastic approach we have adopted, yields very good performances and allows for high-precision reconstruction from samples taken above the critical innovation rate, but still way below the standard Nyquist rate. The superior performance of the stochastic approach compared to the classical subspace methods makes it a natural candidate for benchmark real time oriented algorithms, where speed is achieved at the price of poorer performances. Moreover, this method is numerically stable. With increasing interest in bandwidth-expanding communication systems, this work gives a promising framework with which communication engineers can trade-off system performance with complexity, using sampling rates well below the Nyquist rate.

7. REFERENCES

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

[2] P. Stoica and R. Moses, *Introduction to Spectral Analysis*, Prentice Hall, 1997.

[3] S. Y. Kung, K. S. Arun, and D. V. Bhaskar Rao, "State-space and singular-value decomposition-based approximation methods for the harmonic retrieval problem," *J. Opt. Soc. Amer.*, vol. 73, no. 12, pp. 1799–1811, December 1983.

[4] C. E. Shannon, "A mathematical theory of communication," *Bell System Tech. J.*, vol. 27, pp. 379–423,623–656, 1948.

[5] P. Marziliano, *Sampling Innovations*, Ph.D. thesis, EPFL, 2001.

[6] J. J. Kormylo and J. M. Mendel, "Maximum-likelihood detection and estimation of Bernoulli-Gaussian processes," *IEEE Trans. Inform. Theory*, vol. 28, pp. 482–488, 1982.

[7] F. Champagnat, Y. Goussard, and J. Idier, "Unsupervised deconvolution of sparse spike trains using stochastic approximation," *IEEE Trans. Signal Process.*, vol. 44, no. 12, pp. 2988–2998, December 1996.

[8] J. M. Mendel, *Optimal Seismic Deconvolution*, Academic Press, New York, NY, U.S.A., 1983.

[9] Olivier Catoni, "Solving scheduling problems by sim-

ulated annealing," *SIAM J. Control Optim.*, vol. 36, no. 5, pp. 1539–1575, 1998.

[10] J. Kusuma, A. Ridolfi, and M. Vetterli, "Sampling of communications systems with bandwidth expansion," in *Proc. IEEE ICC*, 2002, vol. 3, pp. 1601–1605.

[11] J. Kusuma, I. Maravic, and M. Vetterli, "Sampling with finite rate of innovation: Channel and timing estimation for UWB and GPS," in *Proc. IEEE ICC*, 2003, pp. 3540–3544.

[12] I. Maravic, M. Vetterli, and K. Ramchandran, "High-resolution synchronization and channel estimation with sub-nyquist sampling and application to UWB systems," *Submitted to IEEE Trans. Signal Process.*, 2003.