Robust Time-Frequency Analysis based on the L-estimation and Compressive Sensing

L. Stanković, *Fellow IEEE*, S. Stanković, *Senior Member IEEE*, I. Orović, *Member IEEE,* M. Amin, *Fellow IEEE*

Abstract- **The L-estimate transforms and time-frequency representations are presented within the framework of compressive sensing. The goal is to recover signal or local auto-correlation function samples corrupted by impulse noise. The signal is assumed to be sparse in a transform domain or in a joint-variable representation. Unlike the standard L-statistics approach, which suffers from degraded spectral characteristics due to the omission of samples, the compressive sensing in combination with the L-estimate permits signal reconstruction that closely approximates the noise free signal representation.**

Index Terms – robust time-frequency distributions, L-estimation signal reconstruction, compressive sensing

I. INTRODUCTION

Robust statistics have been shown to deal effectively with impulsive type of noise. Based on the maximum likelihood (ML) and Huber's estimation theory, Katkovnik has introduced the robust M-Fourier transform (FT) [1], which is calculated using absolute error loss function and iterative procedure. The robust approach to the definition of signal representations was extended to the time-frequency (TF) analysis in [2]. However, the M-estimates are quite sensitive to a priori assumption of disturbance's pdf, and they do not produce closed form solutions. Hence, other robust statistics approaches were introduced, such as the L- and the R-estimate to provide more robust representations without the need for iterative procedures [2]. In quadratic and higher order TF representations [3]-[6], which may be considered as the FTs of the local auto-correlation functions (LAFs), the impulse kind of noise may appear as a result of signal multiplications, even in the case of Gaussian input noise [2]. This makes robust TF analysis rather imperative.

The L-estimation form is of particular importance among all robust forms owing to its efficiency when noise pdf is unknown. It can also produce the median and the standard signal representation forms as the special cases. In the most commonly used L-estimation form of signal representations, the α-trimmed form, some values of product between the signal and basis functions are discarded. In quadratic and higher order analysis, the L-estimation forms are applied to the LAFs. However, an incomplete set of samples produces an effect similar to noise. It is shown analytically in this paper that such noise may have a significant variance which strongly depends on the number of omitted samples. In order to improve the quality of representation, the missing samples of signal or LAF should be properly reconstructed. This problem of recovering random samples is under-determined and thus amenable to CS application [7], [8]. It is noteworthy that this could be, in some way, related to the requirements present in image filtering [9], [10], where the noisy pixels are reconstructed by using other available pixels and specialized regularization methods.

Compressive sensing (CS) is a new approach for signal acquisition and compression which has found several applications using different sensing modalities [11]-[14]. In CS, the signal, which is sparse in certain transform domain, can be reconstructed from a small set of measurements by using convex optimization algorithms. This reduction can be due to difficulty of achieving high sampling rates or obtaining high quality of the samples. In the latter, the signal samples are all available, but some must be discarded due to high noise levels or strong interference contaminations. This is the case considered in this paper, where impulsive noise can occur either as additive component to the original signal or can arise later within the LAF in the case of quadratic or higher order TF distributions. Impulse noise in CS applications has been treated using myriad projections and the Lorentzian norm, to improve results over the standard Basis Pursuit (BP) method [15]. However, as stated therein, the Lorentzian norm optimization requires complex parameter selection and exhaustive search procedures, while myriad projections are computationally demanding compared to linear projections [15]. An approximate solution [16] is to combine l_0 least absolute deviation regression model with the weighted median regression. Again, it requires many computations, with parameter selections obtained by trial and error. The aim of this letter is to improve the L-estimation using the CS theory. The data samples affected by impulsive noise, and thus discarded by the L-estimation, can be recovered by employing l_1 sparse reconstruction techniques (e.g. based on standard BP). This method is adopted for the L-estimate forms of both the short-time FT and the Wigner distribution (WD), leading to significant enhancement of their performance compared to the standard L-estimation only approach.

II. THEORETICAL BACKGROUND

A. Robust L-estimate transforms and TF representations

Consider a noisy signal $x(n)=s(n)+v(n)$, where $v(n)$ is a complex-valued noise. An efficient frequency analysis depends on noise pdf. A solution robust to pdf variations is based on L-statistics. The L-estimate FT can be written as:

$$X_L(k)=\sum_{i=0}^{M-1} a_i x_s(i), \quad \mathbf{x}_s = sort\{x(m)e^{-j2\pi mk/M},\ m=0,...,M-1\} \quad (1)$$

where $\sum_{i=0}^{M-1} a_i = 1$. The sorting operation (*sort*) could be performed with respect to either absolute values or real and imaginary parts independently (into non-decreasing order). In the case of impulse noise, the L-estimate can be used, such that $2\alpha(M-2)$ highest values are omitted, while the mean is calculated over the rest of the values, with the coefficients:

$$a_i = \begin{cases} \dfrac{1}{M(1-2\alpha)+4\alpha}, & \text{for } i\in[0, M(1-2\alpha)+4\alpha-1] \\ 0, & \text{elsewhere,} \end{cases} \quad (2)$$

where M is even, whereas α takes values within the range [0, 1/2]. If $M(1-2\alpha)+4\alpha$ is not an integer, then the nearest integer is used. Similar form is used when M is odd. Higher value of α provides better reduction of impulse noise, but reduces number of available samples for analysis. The proper value of α should be chosen according to the expected amount of samples with significant noise. Hence, if Q% of samples is assumed to be heavily corrupted by noise, then α should be chosen such that $(M(1-2\alpha)+4\alpha)/M \le 1-Q/100$. Details about α (or Q) selection and its adaptive form may be found in [2]. One approach starts with median form of signal transform, as a special case of the L-estimate, and iteratively increases the number of samples, thus improving representation. The influence of Q will be further discussed in Section IV.

Following the L-estimate form of the FT, the L-estimate STFT is defined as [2]:

$$STFT_L(n,k) = \sum_{i=-M/2}^{M/2-1} a_i s_i(n,k), \quad (3)$$
$$\mathbf{s}(n,k) = sort\left\{x(n+m)e^{-j2\pi mk/M}, m\in[-M/2, M/2-1]\right\}.$$

Similarly, the L-estimate robust WD can be written as:

$$WD_L(n,k) = \sum_{i=0}^{M-1} a_i w_i(n,k), \quad (4)$$
$$\mathbf{w}(n,k) = sort\left\{x(n+m)x^*(n-m)e^{-j4\pi mk/M}, m\in[-M/2, M/2-1]\right\}.$$

Here, the sorting and α trimming procedure is applied to the LAF. Quite similar procedure can be applied to other higher order TF representations. Boldface letters in the above equations denote corresponding vectors.

III. Compressive sensing based L-estimation

A. Statistical Analysis of Missing Samples Influence

In order to analyze the impact of samples removal using the L-estimation process, consider the standard FT of the non-noisy sinusoidal signal $\exp(j2\pi k_0 m/M)$:

$$X(k) = \sum_{m=0}^{M-1} x(m)e^{-j2\pi mk/M} = \sum_{m=0}^{M-1} e^{j2\pi m(k-k_0)/M}, \quad (5)$$

which reduces to delta function for integer k_0. If we remove arbitrary positioned terms, the resulting FT $X_L(k)$ would have random properties. Consider, for example, a certain percentage Q of the total number of terms in (5) has been removed (M_Q samples) so $N=M-M_Q=M(1-2\alpha)+4\alpha$ terms remain in (5). Then, depending on the value of k, two cases may arise.

Case 1: $k=k_0$ corresponds to the frequency of signal. At this frequency all terms within the sum are the same and equal to 1. Thus, the value of $X_L(k_0)$ does not depend on the positions of the removed samples and it is given by:

$$X_L(k_0) = (M-M_Q). \quad (6)$$

Case 2: $k=l+k_0$, $l\neq 0$. The removed samples in (5) assume values from set $\Theta = \{x_l(m) = \exp(j2\pi ml/M),\ m=0,1,\ldots,M-1\}$, with equal probability, for a given frequency $l=k-k_0$. The removal of samples in L-statistics can be modeled with a new additive noise terms $\varepsilon_L(m) = -x_l(m)$ in (5). Statistical mean of these values (with respect to m) is $E\{x_l(m)\}=0$ for $l\neq 0$, leading to $E\{X_L(l+k_0)\}=0$. The resulting statistical mean for any k is:

$$E\{X_L(k)\} = (M-M_Q)\delta(k-k_0). \quad (7)$$

The variance of the random variable taking values from the set Θ is $var\{x_l(m)\}=1$. The reconstructed values $X_L(k_0+l)$, $l\neq 0$, will be zero-mean random variables. Taking into account that variables in Θ are not independent (since their complete sum over all M samples is zero), the variance of $X_L(k_0+l)$ is:

$$\sigma_r^2 = \mathrm{var}\{\textstyle\sum_{m=1}^{M_Q} x_l(m)\} = M_Q\left(1-\frac{M_Q-1}{M-1}\right) \quad for \quad l\neq 0. \quad (8)$$

The DFT behaves as if the signal was noisy with resulting variance σ_r^2 in the DFT. For $M_Q \ll M$ we get $\sigma_r^2 \cong M_Q$. The ratio of $X_L(k)$ at $k\neq k_0$ and $X_L(k_0)$ satisfies:

$$\left|\frac{X_L(k)}{X_L(k_0)}\right| < \frac{\sqrt{6}}{(M-M_Q)}\frac{\sigma_r}{\sqrt{2}} = \sqrt{\frac{3M_Q}{(M-M_Q)(M-1)}}, \quad (9)$$

with probability of 0.95. Here, we assumed that large number of terms is removed. Thus, according to the central limit theorem the sum behaves as Gaussian random variable (its real and imaginary part, while the absolute value is Rayleigh distributed). As an example, consider M-M_Q=M/8 and M=128. Then $|X_L(k)/X_L(k_0)|<0.41$ with probability 0.95, meaning that 5% values of $X_L(k)$ are above $0.41X_L(k_0)$ due to this source of error. Having in mind the DFT linearity, this error analysis can be easily generalized to a sum of K sinusoidal signals.

B. CS based robust transforms and TF representations

In order to improve spectral characteristics in the case of robust L-estimate representations, we should recover the missing samples. For a small number of missing samples, we can perform a direct search within the range of the remaining values (e.g. from -1 to 1 with step 0.1 and over all phase values from 0 to 2π). The reconstructed values are estimated as the ones producing the best concentration in the transform domain, [4]. The direct search method, however, becomes computationally exhaustive and practically inapplicable when a high number of samples is missing. Alternatively, CS can be used to retrieve the missing samples and reconstruct the signal. Generally, the CS theory states that a signal with M samples can be reconstructed from its N randomly chosen samples ($N<M$), if the signal satisfies certain conditions, such as sparsity, when represented in a basis ψ [8],[17].

If we simply use a set of random observations from the noisy signal to perform CS reconstruction, this set will include, with a certain probability, at least some strong noisy peaks, which are sufficient to produce spreading in the FT domain. Hence, to provide an appropriate observation set, we use the L-estimate form given by (1) to obtain L-statistics based version of original signal x:

$$\mathbf{x}_{\alpha}=\{x(p(i)),\ for\ i\in[0,M(1-2\alpha)+4\alpha-1]\}, \qquad (10)$$

$$\mathbf{p}=\arg\left\{sort\left(x(m)e^{-j2\pi mk/M},m=0,...,M-1\right)\right\}.$$

Vector $\mathbf{p}$ contains the original positions of samples before the sorting operation. In the context of CS, the signal $\mathbf{x}_{\alpha}$, with $N=M-M_Q$ samples, represents a measurements vector. It can be written using the measurement matrix:

$$\mathbf{x}_{\alpha}=\mathbf{\Phi x}, \qquad (11)$$

where $\mathbf{\Phi}$ (NxM) selects samples from $\mathbf{x}$ according to the positions that remains after L-estimation. In other words, it contains only one value "1" per row on the positions defined by the L-estimation approach. Furthermore, the signal $\mathbf{x}$ can be represented as a linear combination of the orthonormal basis vectors as:

$$\mathbf{x}=\sum_{i=0}^{M-1}X(i)\boldsymbol{\psi}_i,\quad or\quad \mathbf{x}=\mathbf{\Psi X}. \qquad (12)$$

If the number of nonzero transform coefficients in $\mathbf{X}$ is $K<<M$, which is the case of sinusoids and Fourier basis then we can say that $\mathbf{x}$ is K sparse. Accordingly, we may write:

$$\mathbf{x}_{\alpha}=\mathbf{\Phi\Psi X}=\mathbf{AX}. \qquad (13)$$

The reconstructed signal $\mathbf{x}_r$ can be obtained as a solution of N linear equations with M unknowns. Having in mind that this system is undetermined and can have infinitely many solutions, optimization based mathematical algorithms should be employed to search for the sparsest solution, consistent with the linear measurements. A near optimal solution is achieved by using the l_1 norm based minimization as follows:

$$\min\|\mathbf{X}_r\|_{l_1}\quad subject\ to\quad \mathbf{x}_{\alpha}=\mathbf{AX}_r, \qquad (14)$$

where $\mathbf{X}_r$ is the DFT vector of reconstructed signal $\mathbf{x}_r$. The l_1 norm is convex, and linear programming can be used. In addition to the CS based L-estimate FT $\mathbf{X}_r$ we may define the CS based L-estimate TF representations. Hence, the STFT can be defined by applying the CS reconstruction to the signal:

$$STFT_r(n,k)=\sum_{m=-M/2}^{M/2-1}x_r(n+m)e^{-j2\pi mk/M}, \qquad (15)$$

where $x_r(n+m)$ is the recovered windowed signal. Similarly, the CS based WD can be defined as follows:

$$WD_r(n,k)=\sum_{m=-M/2}^{M/2-1}x_r(n+m)x_r^*(n-m)e^{-j4\pi mk/M}, \qquad (16)$$

where x_r* denotes the complex conjugate of x_r, while the CS reconstruction is applied to the samples of the LAF $R_r(n,m)=x_r(n+m)x_r^*(n-m)$. For the sake of simplicity, a rectangular window is assumed in previous relations. In the WD or higher order distributions, the resulting noise in LAF is impulsive in nature even for Gaussian input noise, [2]. Thus, the presented approach could be used to significantly improve these representations.

IV. EXAMPLES

Example 1: Consider a discrete noisy signal:

$$x(n)=e^{j2\pi k_{01}n/M}+e^{j2\pi k_{02}n/M}+\nu(n) \qquad (17)$$

where $M=64$, $k_{01}=8$, $k_{02}=52$, and $\nu(n)$ is impulse noise that significantly affects up to 35% of data samples (22 samples). According to the L-statistics 24 values are omitted out of 64. As will be shown, up to about 70% of omitted values will not significantly change the signal reconstruction performance. The STFT is illustrated for a single time instant (FT of one windowed signal part). The standard STFT of non-noisy and noisy signal are shown in Fig. 1.a and b, respectively. The L-estimate approach is given in Fig. 1.c. For comparison, the result of l_1 based robust processing for impulse noise reduction [1],[2],[14], is shown in Fig. 1.d. This result is close to the L-statistics based one, since it is a known efficient tool belonging to the M-estimate-based robust signal analysis. The proposed approach is given in Fig 1.e., which depicts almost the same performance as the original signal transform.

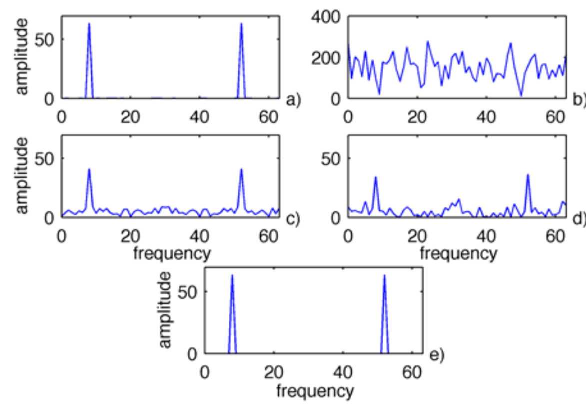

Fig 1. a) STFT of non-noisy signal, b) STFT of noisy signal, c) L-estimate STFT, d) robust l_1-norm based reconstruction, e) CS L-estimate STFT

The CS reconstruction in the proposed approach is performed using the BP primal-dual interior point method (Matlab toolbox l1-magic is adapted and used for this purpose).

Example 2: In order to test the proposed approach in the case of the Wigner distribution, let us observe a signal in the form $x(n)=\exp(j2\pi 48((n+75)/256)^2)+\nu(n)$. The 64-sample window size is used. The L-estimation is applied to the LAF. In order to illustrate the theoretical considerations given in Section II, different percentages Q of the omitted values, corresponding to different values of α, are examined. The MSE between the original and the estimated WD is calculated for a single time instant (Fig. 2.a). We observe that, according to the theory presented in Section II, the standard L-estimate approach degrades almost linearly as the number of omitted samples increases. The proposed approach significantly improves the results of the L-estimate WD, as long as the number of omitted samples is below 80% (Fig. 2.a). It means that at least 25-30% of samples is required for desirable results, which is achievable in most real cases. One may generally assume that at least 25-30% of samples will be slightly influenced by impulsive noise, and can thus be used for reconstruction. Now, we observe the MSE analysis in the presence of impulse noise (Fig. 2.b). For a small number of removed samples (small α), the impulse noise dominates. After removing the strong pulses, the MSE for the CS based reconstruction is almost negligible over a wide range of removed samples (from 30% to about 70%). Afterwards, the CS fails to provide any improvement. Thus, we may assert that when impulse noise is expected, about 70% of samples can be removed by using the L-statistics, and by applying the CS algorithm, the sparse signal reconstruction can be close to the original non-noisy signal.

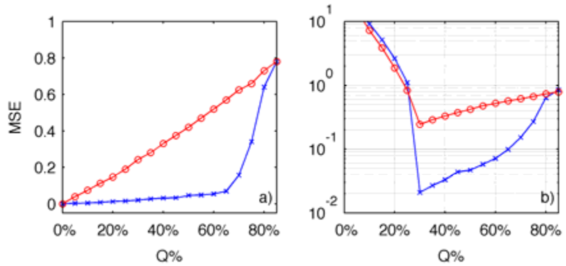

Fig 2. MSE between original and estimated WD (L-estimate WD - lines with "o", Proposed CS approach – lines marked with "+"): a) Non-noisy case, b) Noisy case (30% of samples are corrupted by noise) in log scale

The L-estimate WD and the CS based L-estimate WD, calculated for the entire signal in impulse noise are shown in Figs 3.a and b, respectively. We omit 50% of the data (15% more than the number of samples with significant noise). The case with Gaussian input noise is given in Fig. 3.c and d (70% samples are omitted).

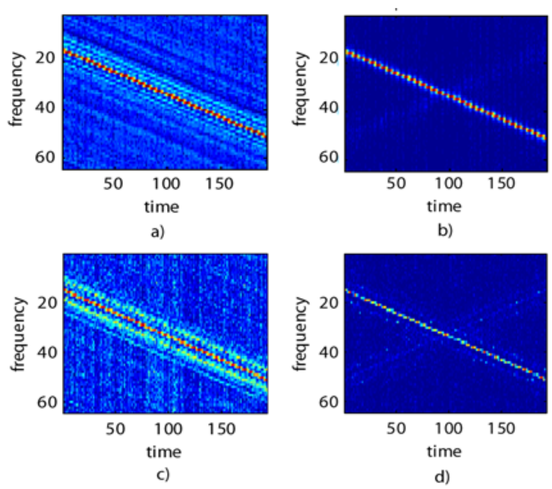

Fig 3. Impulse input noise: a) L-estimate WD , b) CS L-estimate WD, Gaussian input noise: c) L-estimate WD, d) CS L-estimate WD

An example with a real world frequency modulated signal is considered (Fig 4). The echolocation Daubenton's bat call is used (positive frequency range), with 60% of samples omitted.

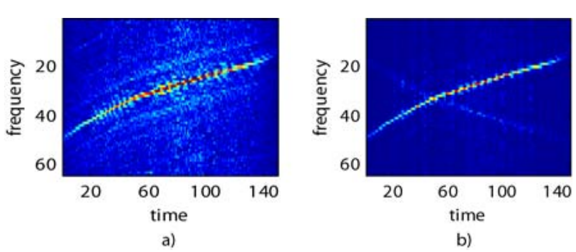

Fig 4. Echolocation bat sound: a) L-estimate WD, b) CS L-estimate WD

V. Conclusion

Compressive sensing was applied in combination with the L-statistics to achieve highly concentrated spectral representation of noisy signals. The L-estimation approach removes the noisy peaks, but produces perturbations which can be cast as noise with a significant variance. It has been shown that CS can effectively reconstruct missing samples, which improves representation. The CS based L-estimation forms of both the STFT and the WD outperform their counterparts based on the standard L-estimate approach. The application of this approach would be of crucial importance in higher order signal analysis, when noise is inherently of impulse nature.

References

[1] V. Katkovnik, "Robust M-periodogram," *IEEE Trans. SP*, vol.46, no.11, pp. 3104-3109, 1998

[2] I. Djurović, L. Stanković, J. F. Böhme, "Robust L-estimation based forms of signal transforms and time-frequency representations," *IEEE Trans. on SP*, vol. 51, no. 7, pp. 1753-1761, 2003.

[3] G. Viswanath, T. V. Sreenivas, "IF estimation using higher order TFRs," *Signal Processing*, vol. 82, no. 2, pp. 127-132, 2002.

[4] L. Stankovic, M. Dakovic, T. Thayaparan, *Time-frequency Signal Analysis with Applications*, Artech House, Boston, 2013.

[5] B. Boashash and P. J. O'Shea, "Polynomial Wigner-Ville distributions and their relationship to time-varying higher order spectra," *IEEE Trans. SP*, vol. 42, no.1, pp. 216-220, 1994.

[6] J. Lerga, V. Sucic, "Nonlinear IF Estimation Based on the Pseudo WVD Adapted Using the Improved Sliding Pairwise ICI Rule," *IEEE SP Letters*, vol.16, no.8, pp. 953-956, 2009.

[7] D. Donoho, "Compressed sensing," *IEEE Trans. IT*, vol. 52, no.4, pp. 1289 - 1306, 2006

[8] R. Baraniuk, "Compressive sensing," *IEEE SP Magazine*, vol. 24, no. 4, pp. 118-121, 2007

[9] X. Zeng, L. Yang, "Mixed impulse and Gaussian noise removal using detail-preserving regularization," *Optical Eng.*, vol. 49, no. 9, 097002, 2010.

[10] R.H. Chan, X.W. Ho, and M. Nikolova, "Salt-and-Paper Noise Removal by Median-type Noise Detectors and Detail-preserving Regularization," *IEEE Trans. IP*, vol.14, no. 10, pp. 1479-1485.

[11] F. Ahmad, M. G. Amin, "Through-the-wall human motion indication using sparsity-driven change detection," *IEEE Trans. GRS,* vol. 51, no.2, pp. 881-890, 2013

[12] P. Flandrin, P. Borgnat, "Time-Frequency Energy Distributions Meet Compressed Sensing," *IEEE Trans. SP*., vol. 58, No. 6, pp. 2974-2982.

[13] M. Lustig, D. Donoho, and J. M. Pauly, "Sparse MRI: the application of compressed sensing for rapid MR imaging," *Magnetic Resonance in Medicine*, vol. 58, no. 6, pp. 1182–1195, 2007.

[14] J. Meng, J. Ahmadi-Shokouh, H. Li, Z. Han, S. Noghanian, and E. Hossain, "Sampling rate reduction for 60 GHz UWB communication using compressive sensing," *Asilomar Conf. on Signals, Systems & Computers*, Monterey, CA, 2009.

[15] R. E. Carrillo , K. E. Barner , and T. C. Aysal, "Robust sampling and reconstruction methods for sparse signals in the presence of impulsive noise," *IEEE J. of Select. Topics in Sig. Proc.*, vol. 4, no.2, pp. 392–408, 2010

[16] J. L. Paredes, and G. R. Arce, "Compressive sensing signal reconstruction by weighted median regression estimates," *IEEE Trans. on SP*, vol. 59, no. 6, pp. 2585–260, 2011.

[17] S. Stankovic, I. Orovic, E. Sejdic, *Multimedia signals and systems*, Springer 2012.