On Signal Reconstruction from Its Spectrogram

Radu Balan

University of Maryland

Department of Mathematics, Center for Scientific Computation

And Mathematical Modeling, and Norbert Wiener Center

College Park, MD 20850

rvbalan@math.umd.edu

Abstract—This paper presents a framework for discrete-time signal reconstruction from absolute values of its short-time Fourier coefficients. Our approach has two steps. In step one we reconstruct a band-diagonal matrix associated to the rank-one operator $K_x = xx^*$. In step two we recover the signal x by solving an optimization problem. The two steps are somewhat independent, and one purpose of this paper is to present a framework that decouples the two problems. The solution to the first step is connected to the problem of constructing frames for spaces of Hilbert-Schmidt operators. The second step is somewhat more elusive. Due to inherent redundancy in recovering x from its associated rank-one operator K_x, the reconstruction problem allows for imposing supplemental conditions. In this paper we make one such choice that yields a fast and robust reconstruction. However this choice may not necessarily be optimal in other situations. It is worth mentioning that this second step is related to the problem of finding a rank-one approximation to a matrix with missing data.

I. INTRODUCTION

This paper presents a framework for discrete-time signal reconstruction from absolute values of its short-time Fourier coefficients. In this paper we consider 1-dimensional signals in the infinite dimensional Hilbert space $H = l^2(\mathbf{Z})$. The term "signal" means a vector in this Hilbert space.

The short-time Fourier transform is defined using a "window" $g \in H$ as follows. Fix two integer: *number of frequencies* F, and *time step* b. For every vector $x \in H$ its *short-time Fourier transform* is defined by

$$c_{k,f} = \langle x, g_{k,f} \rangle := \sum_{t \in \mathbf{Z}} e^{-2\pi i ft/F} x(t) \overline{g(t - kb)} \quad (1)$$

with $f \in \mathbf{Z}_F$, $k \in \mathbf{Z}$. Thus the short-time Fourier transform maps $H = l^2(\mathbf{Z})$ into $l^2(\mathbf{Z} \times \mathbf{Z}_F)$, as is well-known from square integrable representation theory of the discrete Weyl-Heisenberg group (see [1], [2]). Let L denote the length of the support of g, and assume the following support condition $supp(g) \subset \{0, 1, \ldots, L-1\}$. The problem we study in this paper is the following: given $\{|c_{k,f}|, (k,f) \in \mathbf{Z} \times \mathbf{Z}_F\}$ for $b = 1$ and $F \geq 2L$, we want to reconstruct x up to a global phase factor. We will analyze an algorithm for this task, whose origins are inspired by the analysis done in [3], [4]. Some parts of this algorithm are similar to the reconstruction algorithm proposed in [5]. Our approach has two steps (or phases). In step one we reconstruct a band-diagonal matrix associated to the rank-one operator $K_x = xx^*$. In step two we recover the signal x by solving an optimization problem. The two steps are somewhat independent, and one purpose of this paper is to present a framework that decouples the two problems. The solution to the first step is connected to the problem of constructing frames for spaces of Hilbert-Schmidt operators. This latter topic seems well studied in the context of Gabor multipliers (see [6]). The second step is somewhat more elusive. Due to inherent redundancy in recovering x from its associated rank-one operator K_x, the reconstruction problem allows for imposing supplemental conditions. In this paper we make one such choice that yields a fast and robust reconstruction. However this choice may not necessarily be optimal in other situations. It is worth mentioning that this second step is related to the problem of finding a rank-one approximation to a matrix with missing data. The organization of the paper is the following. Section II presents the two step approach for our problem, and contains fundamental properties for the objects of interest. Section III analyzes the specific case of STFT in $H = l^2(\mathbf{Z})$ presents reconstruction methods for the truncated K_x. Section IV analyzes the second step and presents a causal algorithm that solves this problem. Section V contains numerical examples and conclusions.

II. NONLINEAR EMBEDDING AND INVERSE MAPS

Consider a frame $\mathcal{F} = \{f_i; i \in I\}$ for the Hilbert space H, indexed by a countable index set I. As it is

well known in frame theory, the (linear) analysis map:

$$x \in H \mapsto T(x) = \{\langle x, f_i \rangle\}_{i \in I} \in l^2(I)$$

admits a left inverse (possibly non-unique) $R : l^2(I) \to H$ implementable via a linear synthesis formula:

$$R(c) = \sum_{i \in I} d_i f_i' \ , \ c \in l^2(I)$$

where $\{f_i'; i \in I\}$ is a dual frame. Thus $RT(x) = x$. Define $|.| : l^2(I) \to l^2(I)$ the nonlinear map that applies the absolute value on each entry, $|c| = \{|c_i|\}_{i \in I}$. The problem is now to construct a left inverse for the nonlinear analysis map $|T| : H \to l^2(I)$, where, explicitly, $|T|(x) = \{|\langle x, f_i \rangle|\}_{i \in I}$.

The approach we study here is the following. Denote by $HS(H)$ the Hilbert space of Hilbert-Schmidt operators over H endowed with the inner product

$$\langle A, B \rangle_{HS} = trace(AB^*) = \sum_{i,j} A_{i,j} \bar{B}_{j,i}$$

where $(A_{i,j}), (B_{i,j})$ are the matrix representations of the compact operators A, B with respect to a fixed orthonormal basis $\{e_i\}$ of H. Consider the nonlinear embedding K of H into $HS(H)$ given by

$$x \mapsto K_x \ , \ K_x(y) = \langle y, x \rangle x$$

that associates a rank-one operator to vector x.

Note the following key relation:

$$|\langle x, y \rangle|^2 = trace(K_x K_y) = \langle K_x, K_y \rangle_{HS}$$

Thus, the information contained in the magnitude of frame coefficients $|T(x)|$ is equivalent to the following scalar products:

$$|T(x)|_i^2 = \langle K_x, K_{f_i} \rangle_{HS}$$

This simple observation implies the following result:

Theorem 2.1 ([7]): Assume $\{K_{f_i}\}_{i \in I}$ is a frame for $HS(H)$. Denote by $\{\tilde{K}_i; i \in I\}$ one of its dual frames. Let $d_i = |\langle x, f_i \rangle|$. Then for every vector $y \in H$ so that $\langle y, x \rangle \neq 0$ there exists a unit complex number z (i.e. $|z| = 1$) so that:

$$R = \sum_{i \in I} d_i^2 \tilde{K}_i \tag{2}$$

$$zx = \frac{1}{\sqrt{\langle R(y), y \rangle}} R(y) \tag{3}$$

However in the case of STFT coefficients, the set K_{f_i} turns out not to be frame for the entire Hilbert-Schmidt space $HS(H)$. Instead, the linear combination in (2) yields PK_x, the orthogonal projection of K_x onto the span of K_i's.

III. GABOR-LIKE FRAME SEQUENCES FOR HILBERT-SCHMIDT OPERATORS

Consider the specific case $H = l^2(\mathbf{Z})$, $I = \mathbf{Z} \times \mathbf{Z}_F$, and the frame vectors $g_{k,f}(t) = e^{2\pi i ft/F} g(t - k)$. As mentioned in the introduction, we consider the case of 1-sample step size ($b = 1$) and frequency oversampling factor $F \geq 2L$ where $supp(g) \subset \{0, 1, \ldots, L - 1\}$.

The nonlinear embedding $K : l^2(\mathbf{Z}) \to HS(l^2(\mathbf{Z}))$ with respect to the canonical basis of $l^2(\mathbf{Z})$ has the following representation: $(K_x)_{t_1, t_2} = \overline{x_{t_1}} x_{t_2}$. The rank-one operator associated to each $g_{k,f}$ is represented by the following matrix:

$$(G_{k,f})_{t_1,t_2} = e^{2\pi i f(t_2 - t_1)/F} \overline{g(t_1 - k)} g(t_2 - k)$$

Consider the frame operator associated to $\{G_{k,f}\}_{(k,f) \in \mathbf{Z} \times \mathbf{Z}_F}$, $S : HS(l^2(\mathbf{Z})) \to HS(l^2(\mathbf{Z}))$,

$$S(X) = \sum_{(k,f) \in \mathbf{Z} \times \mathbf{Z}_F} \langle X, G_{k,f} \rangle_{HS} G_{k,f}$$

Denote by $\mathcal{T}$ and Σ the unitary operators acting on $HS(l^2(\mathbf{Z}))$ via conjugation with the shift operator T and modulation operator M respectively,

$$\mathcal{T}(X) = T^* X T \ , \ \Sigma(X) = M^* X M$$

where $T, M : l^2(\mathbf{Z}) \to l^2(\mathbf{Z})$ are $Tf(t) = f(t - 1)$, $Mf(t) = e^{2\pi i t/F} f(t)$. Note the frame operator S commutes with $\mathcal{T}$ and Σ, and $S = \sum_{(k,f) \in \mathbf{Z} \times \mathbf{Z}_F} \mathcal{T}^k \Sigma^f(G_{0,0})$. Via Borel functional calculus, for any Borel function ϕ on the spectrum of S, there is a Hilbert-Schmidt operator G^ϕ so that

$$\phi(S) = \sum_{(k,f) \in \mathbf{Z} \times \mathbf{Z}_F} \mathcal{T}^k \Sigma^f(G^\phi)$$

where the series converges weakly. If ϕ is continuous then the series converges in operator norm. In particular, if $\{G_{k,f}\}_{k,f}$ is a frame for its span, then the spectrum of S is included in $\{0\} \cup [A, B]$ for some $0 < A \leq B < \infty$, and the orthogonal projection P onto the span of $G_{k,f}$'s is given by

$$P = \sum_{(k,f) \in \mathbf{Z} \times \mathbf{Z}_F} \mathcal{T}^k \Sigma^f(\widetilde{G_{0,0}}) S$$

where $\widetilde{G_{0,0}} = S^{-1} G_{0,0}$.

In the continuous time case (that is for $H = L^2(\mathbf{R})$) this result was first obtained by H.Feichtinger in a series of papers that started around 2000 (see e.g. [8], [9]). In that case, the authors proved that the set $\{G_{k,f}\}$ if frame for its span if and only if it is Riesz basis for its span, hence there is no redundancy. In the case studied in

this paper the situation is different. It turns out (we will explore this result in a separate paper) the coherent set $\{G_{k,f}\}_{k,f\in\mathbf{Z}\times\mathbf{Z}_F}$ may be frame for its span without being Riesz basic sequence. The reason for this phenomenon lays in the fact that the dual group of $\mathbf{Z}\times\mathbf{Z}_F$ has nontrivial connected components.

Let us analyze the span of $\{\mathcal{G}_{k,f}\}$. Note that for any Hilbert-Schmidt matrix $X \in HS(l^2(\mathbf{Z}))$ so that $X_{t_1,t_2} = 0$ for $|t_1 - t_2| \geq L$, $\langle X, G_{k,f}\rangle_{HS} = 0$. Hence only information about the L-band part of a matrix X is contained in coefficients $\langle X, G_{k,f}\rangle_{HS}$. Denote by E the closure of the span of $\{G_{k,f}\}_{(k,f)\in\mathbf{Z}\times\mathbf{Z}_F}$ in $HS(l^2(\mathbf{Z}))$, and by M_L the set of L-band matrices, $M_L = \{X \in HS(l^2(\mathbf{Z})) , X_{t_1,t_2} = 0 \ \forall|t_1 - t_2| \geq L\}$. Theorem 3.1 below gives necessary and sufficient conditions so for E to coincide with M_L.

Let $z_t^a = \overline{x_{t+a}}x_t$ be the ath-diagonal of the rank-one operator K_x. Set $h_t^a = g(-t)\overline{g(a-t)}$, $H^a(\omega) = \sum_{t\in\mathbf{Z}} e^{-2\pi i\omega t}h_t^a$ and let $c_{k,f} = \langle x, g_{k,f}\rangle$. Then a direct computation shows that:

$$\frac{1}{F}\sum_{f=0}^{F-1} e^{2\pi ifa/F}|c_{t,f}|^2 = \sum_{s\in\mathbf{Z}} h_{t-s}^a z_s^s \qquad (4)$$

In turn this yields the following:

Theorem 3.1: Assume $supp(g) \subset \{0, 1, \ldots, L-1\}$ and consider the functions $H^a : [0,1) \to \mathbf{C}$ defined above. Then:

1) $E = M_L$ if and only if for every $0 \leq a \leq L-1$ and $\omega \in [0,1)$, $H^a(\omega) \neq 0$;

2) The set $\{G_{k,f} ; (k,f) \in \mathbf{Z}\times\mathbf{Z}_F\}$ is frame for M_L with frame bounds A, B is and only if $A \leq |H^a(\omega)|^2 \leq B$ for all $\omega \in [0,1)$ and $0 \leq a \leq L-1$;

3) The set $\{G_{k,f} ; (k,f) \in \mathbf{Z}\times\mathbf{Z}_F\}$ is frame for its span with frame bounds A, B if and only if for every $\omega \in [0,1)$ and $0 \leq a \leq L-1$ either $H^a(\omega) = 0$ or $A \leq |H^a(\omega)|^2 \leq B$.

This theorem suggests a reconstruction algorithm for (z_t^a) using linear filters. The block diagram of this scheme is included in Figure 1 where the linear filters W^a are inverse filters of H^a. Explicitly, if $W^a(\omega) = \sum_t e^{-2\pi i\omega t}w_t^a$ then the tap-t coefficient w_t^a of filter W^a is given by

$$w_t^a = \int_0^1 \frac{e^{2\pi iat\omega}}{H^a(\omega)}d\omega \qquad (5)$$

IV. RANK ONE APPROXIMATIONS

Previous section shows how to estimate $z_t^a = \overline{x(t+a)}x(t)$. In this section we present an on-line least-square estimate of the signal $x = (x_t)_t$. Specifically,
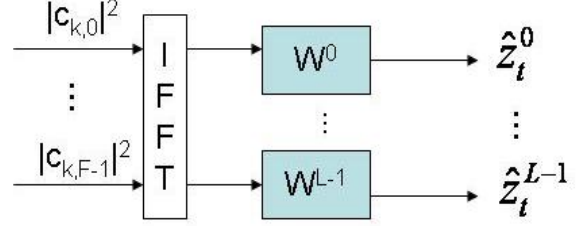

Fig. 1. Block diagram for linear reconstruction of sequences $(z_t^a)_t$.

assume we have estimated the previous J samples (for some $J < L$), $\hat{x}_{t-1}, \ldots, \hat{x}_{t-J}$, then the estimator for x_t is obtained by minimizing:

$$I(u) = p_0||u|^2 - |\widehat{z_t^0}|| + \sum_{k=1}^{J-1} p_k|\overline{(u)}\hat{x}_{t-k} - \widehat{z_{t-k}^k}|^2$$

over u. Here the weights $p_0, \ldots, p_{J-1}$ can be set as variances of the estimators of $z_t^0, \ldots, z_t^{J-1}$. Explicitly, the solution is given by

$$\hat{x}_t = \begin{cases} u_+ & if \ |u_+|^2 \geq |\widehat{z_t^0}| \ and \\ & \quad J(u_+) \leq min(J(u_0), J(u_-)) \\ u_- & if \ |u_-|^2 \leq |\widehat{z_t^0}| \ and \\ & \quad J(u_-) \leq min(J(u_0), J(u_+)) \\ u_0 & if \ otherwise \end{cases} \qquad (6)$$

where

$$u_\pm = \frac{\sum_{k=1}^{J-1} p_k\overline{\hat{x}_{t-k}}\widehat{z_{t-k}^k}}{\sum_{k=1}^{J-1} p_k|\hat{x}_{t-k}|^2 \pm p_0} \qquad (7)$$

$$u_0 = \sqrt{|\widehat{z_t^0}|} \frac{\sum_{k=1}^{J-1} p_k\overline{\hat{x}_{t-k}}\widehat{z_{t-k}^k}}{|\sum_{k=1}^{J-1} p_k\overline{\hat{x}_{t-k}}\widehat{z_{t-k}^k}|} \qquad (8)$$

V. EXAMPLES AND CONCLUSIONS

The block diagram of the solution is depicted in figure 2.

Unfortunately, as observed with the algorithm proposed in [5], this reconstruction algorithm is not very robust to coefficient estimation error. One contribution of this paper is to explain the cause of this instability. The perfect reconstruction result hinges on the assumptions that $\{G_{k,f}\}_{(k,f)\in\mathbf{Z}\times\mathbf{Z}_F}$ is frame for its span and its span coincides with M_L. Theorem 3.1 expresses necessary and sufficient conditions for this to happen. However in many cases of interest the filters $H^a(\omega)$ have at least one zero on $[0,1)$ which introduces an instability in the
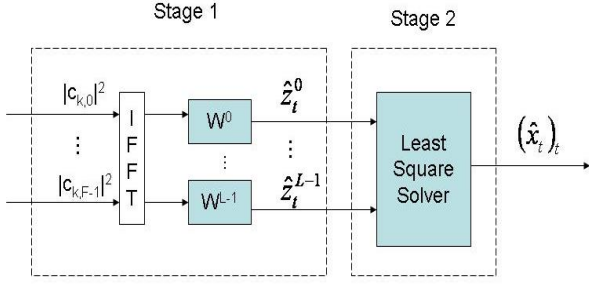
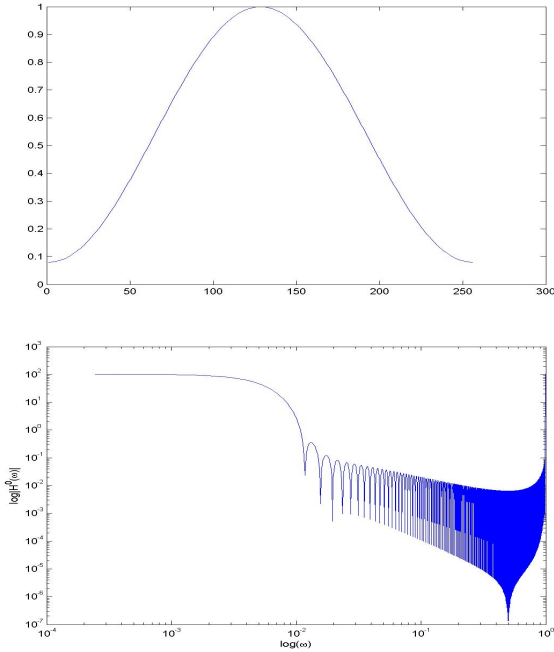Fig. 2.   Block diagram of the reconstruction algorithm.

Fig. 3.   The Hamming window (top) and the corresponding filter $H^0(\omega)$ (bottom) in log-log scale.

inverse filter W^a. For instance, for the Hamming window (figure 3, top plot):

$$g(t) = 0.54 - 0.46 cos(2\pi t/(L-1)) \, , \, 0 \le t \le L-1$$

with $L = 256$, the filter $H^0(\omega)$ is depicted in a log-log scale in the lower plot of Figure 3.

On the other hand, the exponential window

$$g(t) = e^{-3t/256} \, , \, 0 \le t \le 255$$

has the assocoated filter H^0 as depicted in Figure 4.

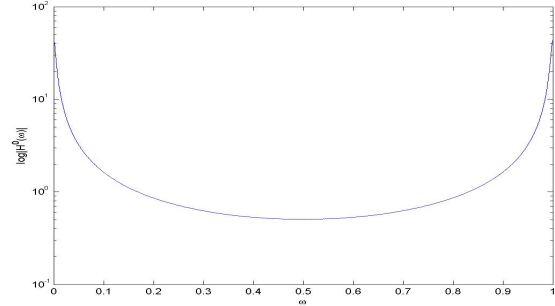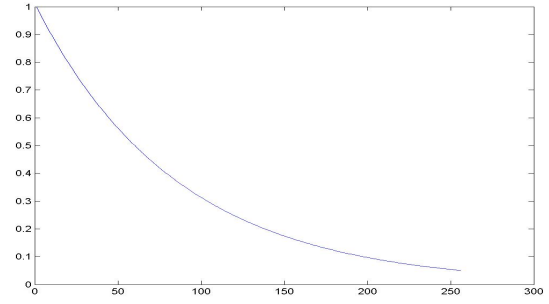For this window, the reconstruction algorithm works very well.

Fig. 4.   The exponential window (top) and the corresponding filter $H^0(\omega)$ (bottom).

REFERENCES

[1] H.G. Feichtinger and Thomas Strohmer, *Gabor Analysis and Algorithms. Theory and Applications.*, Birkhäuser, Boston, 1998.

[2] H.G. Feichtinger and Thomas Strohmer, *Advances in Gabor Analysis*, Birkhäuser, Basel, 2003.

[3] Radu Balan, Pete Casazza, and Dan Edidin, "Equivalence of reconstruction from the absolute value of the frame coefficients to a sparse representation problem," *IEEE Sig. Proc. Letters*, vol. 14, pp. 341–343, 2007.

[4] Radu Balan, Pete Casazza, and Dan Edidin, "On signal reconstruction from absolute value of frame coefficients," in *Frame isotropic multiresolution analysis for cardiac CT imaging*, et al. and Bernhard G. Bodmann, Eds., 2005, vol. 5914 of *Proceedings of the SPIE*, pp. 355–362.

[5] J.S. Lim S.H. Nawab, T.F. Quatieri, "Signal reconstruction from short-time fourier transform magnitude," *IEEE Trans. ASSP*, vol. 31, no. 4, pp. 986–998, August 1983.

[6] J.J. Benedetto and G.E. Pfander, "Frame expansions for Gabor multipliers," *Appl. Comput. Harm. Anal.*, vol. 20, no. 1, pp. 26–40, 2006.

[7] R. Balan, B. Bodman, P. Casazza, and D. Edidin, "Painless reconstruction from magnitudes of frame coefficients," *J.FourierAnal.Appl.*, 2009.

[8] H.G. Feichtinger, "Spline-type spaces in Gabor analysis," in *Wavelet Analysis: Twenty Years Developments. Proceedings of the international conference of computational harmonic analysis, Hong Kong, China, June 4–8, 2001*, D. X. Zhou, Ed., vol. 1 of *Ser. Anal.*, pp. 100–122. World Sci.Pub., River Edge, NJ, 2002.

[9] H.G. Feichtinger and K. Nowak, "A first survey of Gabor multipliers," in *Advances in Gabor Analysis*, H.G. Feichtinger and T. Strohmer, Eds., Appl. Numer. Harmon. Anal., pp. 99–128. Birkhäuser, 2003.