

# Cross-Term-Free Time–Frequency Distribution Reconstruction via Lifted Projections

ZEYNEL DEPREM

A. ENİS ÇETİN, Fellow, IEEE  
Bilkent University  
Ankara, Turkey

**A crucial aspect of time–frequency (TF) analysis is the identification of separate components in a multicomponent signal. The Wigner–Ville distribution is the classical tool for representing such signals, but it suffers from cross-terms. Other methods, which are members of Cohen’s class of distributions, also aim to remove the cross-terms by masking the ambiguity function (AF), but they result in reduced resolution. Most practical time-varying signals are in the form of weighted trajectories on the TF plane, and many others are sparse in nature. Therefore, in recent studies the problem is cast as TF distribution reconstruction using a subset of AF domain coefficients and sparsity assumption. Sparsity can be achieved by constraining or minimizing the  $l_1$  norm. In this article, an  $l_1$  minimization approach based on projections onto convex sets is proposed to obtain a high-resolution, cross-term-free TF distribution for a given signal. The new method does not require any parameter adjustment to obtain a solution. Experimental results are presented.**

Manuscript received January 30, 2014; revised May 31, 2014; released for publication August 25, 2014.

DOI: No. 10.1109/TAES.2014.140080.

Refereeing of this contribution was handled by K. T. Wong.

The work of A. E. Çetin is funded by the Turkish Scientific and Technical Research Council (TUBITAK).

Authors’ address: Bilkent University, Electrical and Electronics Engineering, Bilkent, Ankara 06800, Turkey, E-mail: (zdeprem@ee.bilkent.edu.tr).

0018-9251/15/\$26.00 © 2015 IEEE

## I. INTRODUCTION

Signals with time-varying frequency content are encountered in many areas, such as AM/FM communication, radar, sonar applications, medicine (electroencephalograms), gravitational analysis, speech, and audio. An important aspect of time–frequency (TF) analysis is the identification of separate components in a multicomponent signal. High-resolution TF representations and methods based on instantaneous frequency (IF) are needed for analysis, detection, and classification of this type of signals. TF signal representations enable separation of time-varying components overlapping in both time and frequency domains. It may not be possible to isolate some signal components in one domain using ordinary frequency domain filtering.

An important application area of high-resolution TF analysis is radar imaging [1–3]. The image of the target is obtained from the spatial distribution of the reflected signal. But tracking a moving target is a challenging task due to Doppler spectra estimation.

In inverse synthetic-aperture radar [4], signals from many small apertures are coherently combined to obtain a large-aperture antenna, and the reflectivity distribution of the target is obtained by Doppler spectra. Doppler information is difficult to obtain using only the Fourier transform, because the scatters may not remain in their range cells during the imaging time. When the scatters drift out of their range cells, the Doppler frequency shift will be time varying and the radar image will be blurred. A high-resolution joint-TF analysis is necessary to obtain a focused image of the target [2, 3, 5, 6].

In addition to constant Doppler frequency shift caused by bulk motion of some target, the target itself or any structure on it may have some micromovements. Mechanical vibrations and rotations are examples of such micromovements. These micromovements cause a modulation in Doppler frequency, and the effect is called a micro-Doppler effect [7]. Micro-Doppler radar signature detection [8–10], micro-Doppler removal [11, 12], and FM signal classification [13, 14] all require high-resolution TF analysis of signals.

Another application area requiring high-resolution TF analysis is fault detection in electric motors operating under nonstationary conditions in aerospace and transportation. TF analysis is used to extract nonstationary fault signatures from brushless DC motor current. For this purpose, the Motor Current Signature Analysis technique [15, 16] is used. Accurate detection of signatures requires cross-term-free and high-resolution TF analysis [17, 18].

The classical tool for TF analysis is the Wigner–Ville (WV) distribution [19, 20]. Smoothed versions of the WV distribution are grouped under the name of Cohen’s class of distributions [21]. The WV distribution is a quadratic TF representation, which provides a good time–frequency resolution especially for chirp-type signals. Because of its quadratic definition, for multicomponent signals [22] the

WV representation shows cross-terms together with actual components or auto-terms. Since the cross-terms result from cross correlation of different components, they have an oscillatory shape on the TF plane. Therefore, in smoothed versions of the WV distribution they are attenuated or completely removed—but at the expense of highly reduced resolution. Because of this trade-off between resolution and cross-term reduction, there have been many smoothing efforts, which try to reduce cross-terms but obtain a good TF resolution [23].

Compressive sensing (CS) is a recently introduced concept which tries to recover a signal from a limited number of random measurements with the assumption that the signal under consideration is sparse in some transform domain [24–26, 41]. In CS problems, a sparsity assumption is imposed on the recovered signal by minimizing a cost function based on the  $l_0$  or  $l_1$  norm. Frequency-modulated (FM) signals used in radar signal processing can be considered to be sparse in the TF plane. The problem of obtaining a high-resolution and cross-term-free TF distribution is studied in [27] using the CS perspective [28]. In this approach, a sparse TF distribution is obtained using  $l_1$  minimization among all TF distributions whose Fourier transform coefficients are equal to a given subset in the ambiguity domain. The cost function used in [27] consists of two terms linearly combined or an upper bound variance of the error. A proper choice of the mixture parameter (regularization parameter) is required to obtain a sparse solution. Regularization-parameter selection is left as an open problem in [27].

In this article we use the POCS framework to solve the high-resolution and cross-term-free TF distribution estimation problem. A lifted-domain POCS method is developed. The new method does not require any parameter adjustment as in [27, 28]. It does not require any a priori bounds on the  $l_1$  norm of the signal [29], either. The new algorithm is designed based on making orthogonal projections onto the epigraph set of the  $l_1$  cost function. It successively imposes constraints on iterates in the TF and ambiguity function domains, as in the well-known Papoulis–Gerchberg algorithm [30, 31].

The paper is organized as follows. In section II, the TF distribution concept and WV distribution are reviewed. Section III defines the TF distribution reconstruction problem with a CS perspective. Section IV explains how the cross-term-free TF distribution is reconstructed using the lifted-projections method. In section V, TF estimation examples are presented and results are compared with other methods in terms of localization and similarity measures. In section VI, some conclusions are drawn.

## II. REVIEW OF TIME–FREQUENCY REPRESENTATIONS AND NOTATION

Linear and quadratic representations are the most widely used TF representations for signals with

time-varying frequency content [32]. The windowed short-time Fourier transform (STFT) of a signal  $x(t)$  is given by

$$STFT_x(t, f) = \int x(\tau) w^*(\tau - t) e^{-j2\pi f\tau} d\tau, \quad (1)$$

where  $w(t)$  is the window or kernel function of the transformation. Linearity is a favored property in analysis, but the selection of the window length is the main challenge for the STFT. While a long window provides good frequency resolution, it reduces the time resolution, and vice versa. There have been efforts [33, 34] to adapt the window length to the signal so that signal dependency and better TF resolution are obtained.

The representative of the quadratic group is the Wigner–Ville distribution (WV) [19, 20], which is defined as follows:

$$W_x(t, f) = \int x\left(t + \frac{\tau}{2}\right) x^*\left(t - \frac{\tau}{2}\right) e^{-j2\pi f\tau} d\tau. \quad (2)$$

Cohen’s class of TF distributions [21] are generalized versions of the VW distributions:

$$P_x(t, f) = \frac{1}{4\pi^2} \iint A_x(\tau, \theta) \Phi(\tau, \theta) e^{-j2\pi(\theta t + f\tau)} d\theta d\tau. \quad (3)$$

where  $A_x(\tau, \theta)$  is the ambiguity function (AF) of the signal  $x$  and is given by

$$A_x(\tau, \theta) = \int x\left(t + \frac{\tau}{2}\right) x^*\left(t - \frac{\tau}{2}\right) e^{j2\pi\theta t} dt. \quad (4)$$

The AF is a 2-D correlation function which correlates the signal  $x(t)$  by its time- and frequency-shifted versions. The parameter  $\tau$  (time lag) and the Doppler parameter  $\theta$  (frequency lag) represent the time and frequency shifts, respectively. In (3),  $\Phi(\tau, \theta)$  is the kernel of the Cohen’s class TF distribution. The WV distribution corresponds to  $\Phi(\tau, \theta) = 1$ . Therefore, the AF and WV are related to each other via the 2-D Fourier transformation (FT). Multiplication of  $A_x(\tau, \theta)$  by  $\Phi(\tau, \theta)$  corresponds to 2-D convolution of  $W_x(t, f)$  with the 2-D Fourier transformed kernel function in the TF plane. Therefore,  $P_x(t, f)$  is a smoothed version of  $W_x(t, f)$ . Among many other nice properties of the WV distribution, the main appreciated feature is its high resolution. But high resolution is achieved at the expense of the cross-terms. Because of its quadratic definition, for multicomponent signals the WV representation shows cross-terms together with actual components or auto-terms [22]. Even with a mono component signal having a nonlinear IF function, the WV distribution may have cross-terms. The kernel function has the role of shaping the ambiguity function  $A_x(\tau, \theta)$ . Cross-terms are located between auto-terms and are oscillatory in nature. Therefore the smoothed version  $P_x(t, f)$  of  $W_x(t, f)$  will have cross-terms attenuated or even removed, depending on the kernel function  $\Phi(\tau, \theta)$ . Spectrogram [32], which is the square magnitude of STFT, is an example of the smoothed WV distribution. In

spectrogram analysis, window length causes a trade-off between time and frequency resolution. The smoothed pseudo-Wigner distribution [32] is a solution to this trade-off via selection of independent smoothing functions for time and frequency parameters.

Beside Cohen's class, there are other generalizations of the WV distribution which aim to achieve high resolution. One example is the L-class WV distribution proposed by Stankovic [35]. In this distribution, the FT of  $x^L(t + \frac{\tau}{2L})x^{*L}(t - \frac{\tau}{2L})$  is computed to get the distribution, where  $L$  is an integer. A value of  $L = 1$  corresponds to the WV distribution.

The WV distribution is ideally suitable to chirp-type signals which have linear frequency variation or second-order polynomial phase functions. The polynomial Wigner–Ville distribution (PWVD) [36, 37] is designed to localize higher order polynomial phase signals. But for multicomponent signals, it also suffers from cross-terms. In fact, in PWVD there exist nonoscillating cross-terms, which cannot be removed by smoothing. Therefore other approaches are needed to remove them. In [36], LPWVD—which is a combination of L-class WV and PWVD—is developed to solve this problem. In both PWVD and LPWVD, the order of transformation needs to be arranged according to the polynomial order of the polynomial phase signal. It is shown in [37] that the sixth-order PWVD achieves the delta function concentration for polynomial FM signals of up to the cubic order.

Among many other TF methods, the reassigned spectrum [38, 39] is the one which achieves the best localization. With this method, cross-term-free high-resolution TF distribution is obtained in two steps. In the first step, cross-terms are removed by a proper smoothing method, such as the spectrogram or smoothed pseudo-WV method. In the second step, each time and frequency point on the TF plane is moved to a new location according to the center of gravity of the neighboring region. In this way, the TF localization or resolution enhancement is obtained.

Since the AF is a correlation function, its values close to the origin correspond to auto-terms and those far from the origin are mainly due to the cross-terms. This is the main idea behind shaping the AF to obtain a cross-term-free distribution. Therefore efforts have been made to shape the ambiguity function by a kernel in a signal-dependent manner to both smooth the TF distribution and obtain a high resolution [40]. In the next section, the TF reconstruction problem using sparse signal processing is revisited.

### III. TF RECONSTRUCTION USING SPARSITY

Most practical time-varying signals are in the form of weighted trajectories on the TF plane. In this respect, although they are sparse neither in time nor in the frequency domain, they are sparse in the joint TF plane. A multicomponent [22] amplitude- and

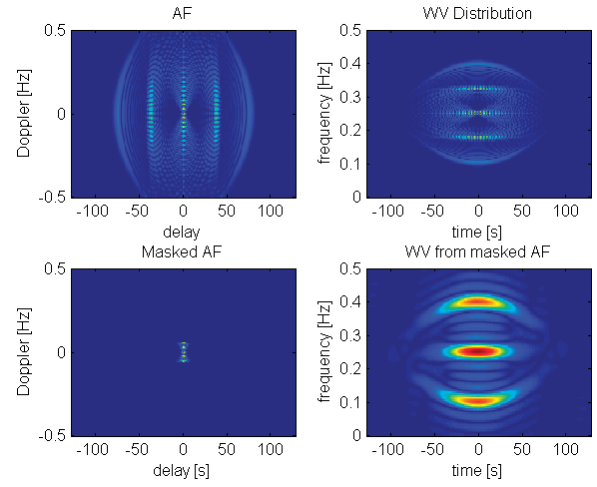


Fig. 1. Effect of shaping ambiguity function on WV distribution. Top left: Ambiguity function of time-varying signal. Top right: WV distribution. Bottom left: Masked ambiguity function. Bottom right: WV distribution corresponding to masked ambiguity function. Horizontal and the vertical axes show time and normalized frequency, respectively.

frequency-modulated signal expressed as

$$x(t) = \sum_{k=1}^L a_k(t) e^{j\vartheta_k(t)} \quad (5)$$

is an example of a signal that is sparse in the TF plane. In this expression,  $a_k(t)$  and  $\vartheta_k(t)$  are the amplitude and phase functions of the  $k$ th signal component. The TF distribution of the  $k$ th component can be expressed as

$$P_k(t, f) = a_k^2(t) \frac{1}{2\pi} \delta\left(f - \frac{d\vartheta_k(t)}{dt}\right). \quad (6)$$

This is a trajectory on the TF plane, with  $d\vartheta_k(t)/dt$  being the instantaneous frequency (IF) function and  $\delta(f)$  the Dirac delta function. Though not all the time-varying signals can be expressed in this form, most practical ones are sparse as in (5). In other words, they are localized in a small area of the TF plane.

The WV distribution is the 2-D Fourier transform (FT) of the AF, and the values of the AF around the origin are due to auto-terms of a multicomponent signal. Therefore, masking the AF with a filter around the origin and computing the 2-D FT may reduce the cross-terms in the WV distribution. But this approach also reduces the resolution, as shown in Fig. 1. The signal has three components or auto-terms in this example. However, the WV distribution has five components (top right). After masking the AF around the origin, the three main components are clearly visible (bottom right). Although the original WV distribution has high resolution, the three reconstructed components appear with a reduced resolution.

Due to the uncertainty principle, perfect localization cannot be obtained in both the TF and AF domains at the same time. Therefore, there is a trade-off between the TF domain resolution and cross-terms. In order to reduce the cross-terms of the TF distribution as much as possible, a

set of optimization problems is proposed by Flandrin and Borgnat [27] as follows:

$$\begin{aligned} \mathbf{P}^* &= \underset{\mathbf{P}}{\operatorname{argmin}} \|\mathbf{P}\|_1 \\ \text{s.t.} \quad \mathcal{F}^{-1}\{\mathbf{P}\} &= \mathbf{A}_x[k, l] \quad k, l \in \Omega, \end{aligned} \quad (7)$$

where  $\mathbf{P}$  and  $\mathbf{A}_x$  are matrices of size  $n \times n$  obtained by discretizing  $P(t, f)$  and  $A_x(\tau, \theta)$ , respectively, and  $n$  is the length of the discrete-time time-varying signal  $x$ . The

$l_1$ -norm is defined as  $\|\mathbf{P}\|_1 = \sum_{i=1}^n \sum_{j=1}^n |P_{i,j}|$ . The set  $\Omega$

defines the filter mask around the origin in the AF domain, and  $k$  and  $l$  are the discrete indices corresponding to the delay and Doppler parameters, respectively. It is established in CS theory that minimization of the  $l_1$ -norm of  $\mathbf{P}$  provides sparsity in the WV domain [27].

The second optimization problem is a relaxed version of (7):

$$\begin{aligned} \mathbf{P}^* &= \underset{\mathbf{P}}{\operatorname{argmin}} \|\mathbf{P}\|_1 \\ \text{s.t.} \quad \|\mathcal{F}^{-1}\{\mathbf{P}\} - \mathbf{A}_x[k, l]\|_2 &\leq \epsilon \quad k, l \in \Omega \end{aligned} \quad (8)$$

where the parameter  $\epsilon$  is a user-defined upper bound on the error variance between the inverse Fourier transform of the WV distribution  $\mathbf{P}$  and the ambiguity function  $\mathbf{A}_x$  over the filter mask  $\Omega$ . Obviously, (8) is equivalent to (7) when  $\epsilon = 0$ .

The third problem is a regularized optimization problem:

$$\mathbf{P}^* = \underset{\mathbf{P}}{\operatorname{argmin}} \lambda \|\mathbf{P}\|_1 + \frac{1}{2} \|\mathcal{F}^{-1}\{\mathbf{P}\} - \mathbf{A}_x[k, l]\|_2^2 \quad k, l \in \Omega, \quad (9)$$

where the regularization parameter  $\lambda$  is also a user-defined parameter adjusting the trade-off between the  $l_1$ -norm minimization and the error between the actual and estimated ambiguity functions. A large  $\lambda$  value corresponds to a sparse WV distribution in the TF plane, but it may correspond to a large deviation from the actual ambiguity function. Optimization problems (8) and (9) are actually equivalent to each other [27, 42]. It is always possible to find a  $\lambda$  value corresponding to each  $\epsilon$  value.

In Fig. 2, a reconstructed solution obtained by minimizing (7) is shown. The signal is the same as the signal in Fig. 1. A circular mask  $\Omega$  with radius  $r = n/16$  around the origin is applied to the AF as in Fig. 1.

The TF distribution in Fig. 2 (top) was obtained using the  $l_1$ -MAGIC TOOLBOX [28]. The 3-D plot of the WV distribution is shown in Fig. 3. The solution has high resolution and cross-terms are removed, but the reconstructed solution is too sparse to be called a TF distribution as stated in [27]. This is because the estimated distribution is not smooth at all; it is discontinuous and spiky.

Therefore, instead of solving the minimization problem (7), which has a strict constraint in the AF domain, the minimization problem (8) with relaxed constraints is solved in [27] to obtain an acceptable result.

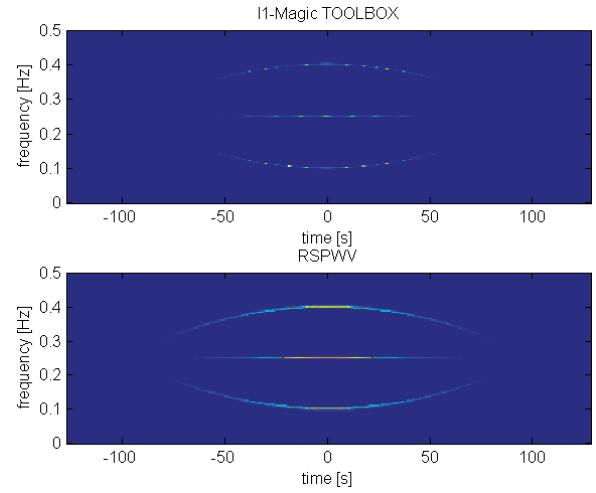


Fig. 2. TF distribution obtained by minimization of (7) using  $l_1$ -MAGIC TOOLBOX (top) and corresponding reassigned smoothed pseudo-Wigner-Ville (RSPWV) distribution [39] (bottom). Frequency is normalized according to sampling frequency in all TF figures.

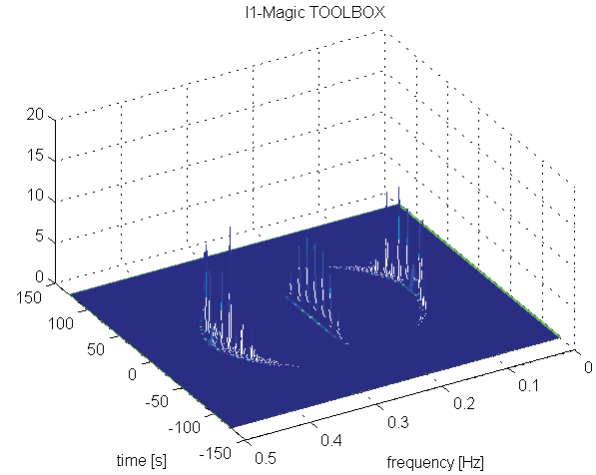


Fig. 3. 3-D plot of TF distribution obtained by minimization of (7) using  $l_1$ -MAGIC TOOLBOX. Frequency is normalized according to sampling frequency.

But in this modified problem, the parameter  $\epsilon > 0$  needs to be properly defined in advance. Therefore, the choice of the regularization parameter  $\lambda$  (or equivalently the upper bound  $\epsilon > 0$ ) is left as an open problem in [27].

Among many other TF representation, the reassigned spectrum [38, 39] results in good TF localization. In Fig. 2 (bottom), the TF was obtained with reassigned smoothed pseudo-WV (RSPWV) using the Time-Frequency Toolbox [43]. In Fig. 4, the 3-D plot of the result is shown.

The reassignment spectrum produces a good localization around the IF law, as shown in Figs. 2 and 4. This is similar to the result obtained by the  $l_1$ -MAGIC TOOLBOX, but it has a spiky nature. In this respect it deviates from the physical meaning of the signal being analyzed. However, it is still the best method in terms of TF localization.

In this paper a lifted POCS method [44, 45] which does not require any regularization parameter or upper bound on the  $l_1$ -norm of the signal is used to estimate the

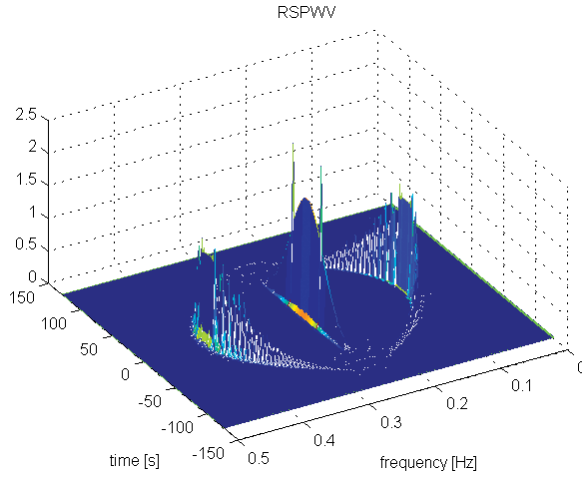


Fig. 4. 3-D plot of RSPWV, which is obtained by using Time-Frequency-Toolbox [43].

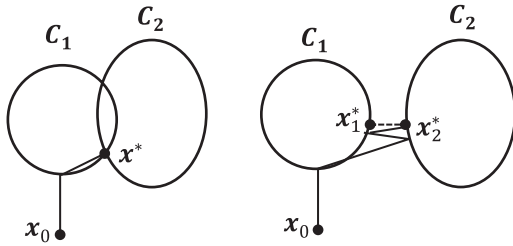


Fig. 5. Left: POCS iterates converge to a vector in intersection of convex sets  $C_1$  and  $C_2$ . Vector  $x_0$  is initial vector and  $x^*$  is in intersection of sets  $C_1$  and  $C_2$ . Right: Iterates oscillate between two vectors when intersection is empty. Vectors  $x_1^*$  and  $x_2^*$  minimize distance between sets  $C_1$  and  $C_2$ .

TF distribution. The algorithm is iterative. Iterates go back and forth between the Fourier and AF domains. In the AF domain, the masking filter is applied on the current iterate. In the TF domain, an orthogonal projection onto the epigraph set of the  $l_1$ -norm is performed. In the next section this method is explained.

#### IV. TF RECONSTRUCTION WITH LIFTED POCS

Bregman's (POCS) framework [30, 31] has been successfully applied to many inverse and design problems in signal and image processing [30, 31, 42, 44, 46–53]. POCS is an iterative signal reconstruction method in which the goal is to find a solution satisfying all the constraints of a given problem in a Hilbert space framework. The solution vector should be in the intersection of all the constraint sets corresponding to the constraints. If the constraint sets happen to be closed and convex, the algorithm globally converges regardless of the initial vector. In each step of the algorithm, an orthogonal projection onto one of the convex sets is performed. Bregman showed that iterates converge to a vector in the intersection of all the convex sets provided that the intersection of the constraint sets is nonempty. If the sets do not intersect, iterates oscillate between members of the sets [46–48]. This process is graphically illustrated in

Fig. 5 for both intersecting and nonintersecting cases.

Both  $x^*$  as well as  $x_1^*$  and  $x_2^*$  are accepted as solutions in inverse problems [46–48].

A POCS-based solution is proposed here based on the cost function ( $l_1$ -norm) of the TF reconstruction problem defined in (8):

$$f(\mathbf{P}) = \|\mathbf{P}\|_1 = \sum_{i=1}^n \sum_{j=1}^n |P_{i,j}|. \quad (10)$$

Since the TF distribution  $\mathbf{P}$  has  $n^2$  entries, it can be converted to a vector in  $\mathbf{R}^{n^2}$ .

In the lifted POCS approach, we increase the dimension of the vectors by one. In  $\mathbf{R}^{n^2+1}$ , any vector on the graph of the  $l_1$ -norm can be represented as follows:

$$\mathbf{w} = [\text{vec}(\mathbf{P})^T f(\mathbf{P})]^T \quad (11)$$

where  $\text{vec}(\mathbf{P}) \in \mathbf{R}^{n^2}$  is the vector form of the TF distribution matrix  $\mathbf{P}$  and the last entry represents the  $l_1$ -norm of the TF distribution  $\mathbf{P}$ . For the TF reconstruction problem, the epigraph set  $C_f$  of the  $l_1$ -norm is defined as

$$C_f = \left\{ \mathbf{w} = [\text{vec}(\mathbf{P})^T v]^T \in \mathbf{R}^{n^2+1} \mid f(\mathbf{P}) = \|\mathbf{P}\|_1 \leq v \right\}, \quad (12)$$

where  $\mathbf{w}$  is an arbitrary vector in the lifted domain  $\mathbf{R}^{n^2+1}$  and  $v$  is the last element of the vector  $\mathbf{w}$ . The epigraph set  $C_f$  contains all the vectors above the graph of the  $l_1$ -norm [54]. The epigraph set of a function is graphically illustrated in the appendix. Since the  $l_1$ -norm is a convex function, the epigraph set is a convex set in the vector space  $\mathbf{R}^{n^2+1}$ . The set  $C_f$  represents our TF domain constraint on the solution of the TF distribution estimation problem.

The second convex set is simply based on the AF domain information. It is the set of TF distributions whose 2-D inverse FT is equal to  $A_x[k, l]$  on the filter mask  $\Omega$ . It is defined as follows:

$$C_{AF} = \left\{ \mathbf{w} = [\text{vec}(\mathbf{P})^T v]^T \in \mathbf{R}^{n^2+1} \mid \mathcal{F}^{-1}\{\mathbf{P}\} = A_x[k, l], \right. \\ \left. k, l \in \Omega. \right\} \quad (13)$$

It can be shown that  $C_{AF}$  is also a closed and convex set.

It may not be possible to know if the sets  $C_f$  and  $C_{AF}$  intersect or not a priori. This depends on the values of the ambiguity function. But we can easily understand if they intersect or not during the implementation of the POCS algorithm. If the iterates converge to a single solution, they intersect. If they oscillate between the two solutions, then it means that the sets  $C_f$  and  $C_{AF}$  do not intersect.

Next we describe the orthogonal projection operations onto the sets  $C_f$  and  $C_{AF}$ .

Given an initial TF distribution  $\mathbf{P}_0$ , we construct a corresponding vector in  $\mathbf{R}^{n^2+1}$  by padding a 0 at the very end as follows:  $\mathbf{w}_0 = [\text{vec}(\mathbf{P}_0)^T 0]^T \in \mathbf{R}^{n^2+1}$ , whose orthogonal projection  $\mathbf{w}_1$  onto  $C_f$  is defined as

$$\mathbf{w}_1 = \min_{\mathbf{w} \in C_f} \|\mathbf{w} - \mathbf{w}_0\|_2^2. \quad (14)$$

The vector  $\mathbf{w}_1$  is the closest vector in  $C_f$  to  $\mathbf{w}_0$ . The solution TF distribution matrix  $\mathbf{P}_1$  is obtained from the first  $n \times n$  entries of  $\mathbf{w}_1 = [\text{vec}(\mathbf{P}_1)^T f(\mathbf{P}_1)]^T$ . The last entry of  $\mathbf{w}_1$  is  $f(\mathbf{P}_1)$  because the projection should be on the boundary of the convex set, which is the graph  $f(\mathbf{P}_1)$ . If  $\mathbf{w}_0$  is inside  $C_f$ , its projection is itself by definition. The projection of  $\mathbf{w}_0$  onto the epigraph set  $C_f$  can be also defined as follows:

$$\begin{aligned} \mathbf{w}_1 &= [\text{vec}(\mathbf{P}_1)^T \quad f(\mathbf{P}_1)]^T \\ &= \min_{\mathbf{P}} \|\text{vec}(\mathbf{P}) - \text{vec}(\mathbf{P}_0)\|_2^2 + f^2(\mathbf{P}), \end{aligned} \quad (15)$$

where the first term is obtained from the first  $n^2$  entries and the second term is obtained from the last entries of  $\mathbf{w}$  and  $\mathbf{w}_0$ , respectively. Notice that square of the  $l_1$ -norm  $f^2(\mathbf{P})$  is different from the  $l_2$ -norm. The solution of the minimization problem (15) is discussed in the appendix.

Next, the vector  $\mathbf{w}_1$  is projected onto  $C_{AF}$ , producing the next iterate  $\mathbf{w}_2$ . The corresponding TF matrix  $\mathbf{P}_2$  satisfies  $\mathcal{F}^{-1}\{\mathbf{P}_2\} = \mathbf{A}_x[k, l], k, l \in \Omega$ . This projection corresponds to the AF domain constraint. It is implemented very easily using the 2-D inverse Fourier transform. The ambiguity function corresponding to  $\mathbf{P}_1$  is computed as follows:

$$\mathbf{A}_1 = \mathcal{F}^{-1}\{\mathbf{P}_1\}. \quad (16)$$

The ambiguity function  $\mathbf{A}_2$  is defined using the actual  $\mathbf{A}_x$  values in the mask  $\Omega$ ,

$$\mathbf{A}_2[k, l] = \mathbf{A}_x[k, l], \quad k, l \in \Omega, \quad (17)$$

and the remaining entries from  $\mathbf{A}_1$ :

$$\mathbf{A}_2[k, l] = \mathbf{A}_1[k, l], \quad k, l \notin \Omega. \quad (18)$$

Next,  $\mathbf{P}_2$  is obtained by computing the 2D FT of  $\mathbf{A}_2$ .

In the second round of POCS iterations,  $\mathbf{P}_2$  or equivalently  $\mathbf{w}_2 = [\text{vec}(\mathbf{P}_2)^T \quad f(\mathbf{P}_1)]^T$  is constructed where first  $n$  entries are taken from  $\mathbf{P}_2$  and  $(n+1)^{\text{th}}$  entry is taken from previous projection onto  $C_f$  because we have not changed the last entry during projection onto  $C_{AF}$ . This vector is projected back onto  $C_f$  to obtain  $\mathbf{P}_3$ . After this projection operation, the constraint (17) is probably no longer valid. Therefore,  $\mathbf{P}_3$  is projected back onto  $C_{AF}$  to obtain  $\mathbf{P}_4$ , and so on. The lifted POCS iterations continue in this manner. Assuming that the intersection of  $C_f$  and  $C_{AF}$  is nonempty, the iterations will converge to a point in the intersection set—or they will oscillate between  $C_f$  and  $C_{AF}$ , as shown in Fig. 5. Both cases are fine with us because we look for a compromise solution for the TF distribution.

The method proposed here also provides globally convergent solutions for other convex cost functions, such as total variation [55], filtered variation [49],  $l_1$ , and the entropic function, that are widely used in signal and image processing problems because all convex cost functions can be represented as closed and convex sets in a lifted vector space.

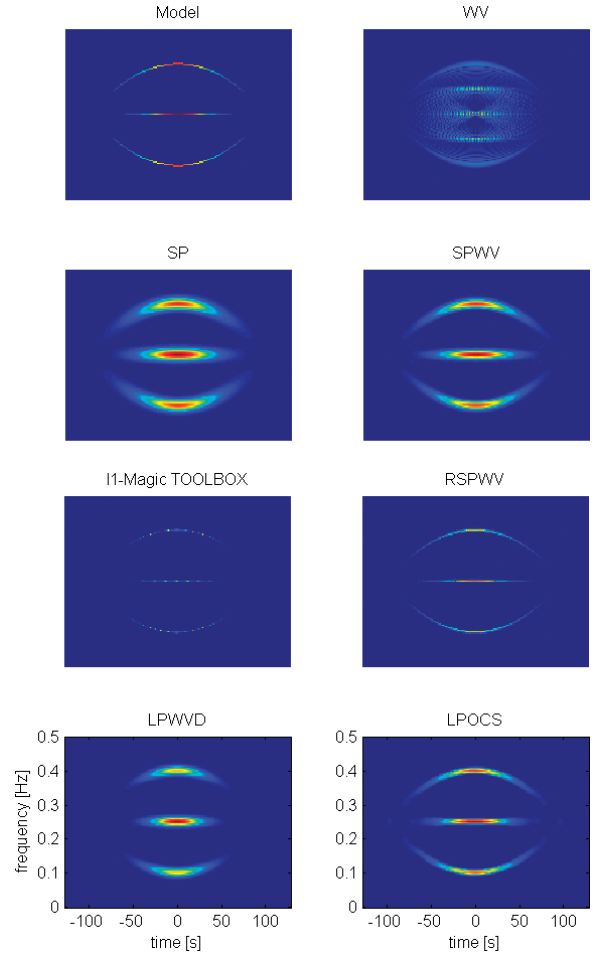


Fig. 6. Example 1: TF reconstruction using various methods. Left column: Ideal model, spectrogram,  $l_1$ -MAGIC TOOLBOX, L-class polynomial WV distribution (LPWVD); right column: WV distribution, smoothed pseudo-WV (SPWV) distribution, reassigned SPWV (RSPWV), lifted POCS. Frequency is normalized frequency.

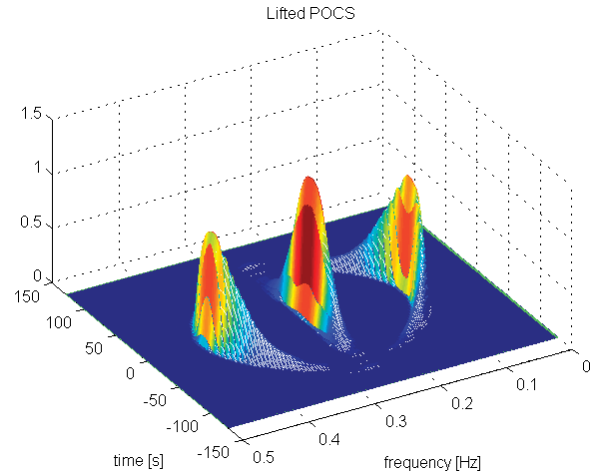


Fig. 7. 3-D plot of estimated TF distribution of Example 1 with lifted POCS method. Frequency is normalized frequency.

## V. EXPERIMENTAL RESULTS

In order to test the effectiveness of the lifted POCS method introduced in section IV, TF distributions for several example signals are estimated. The examples used

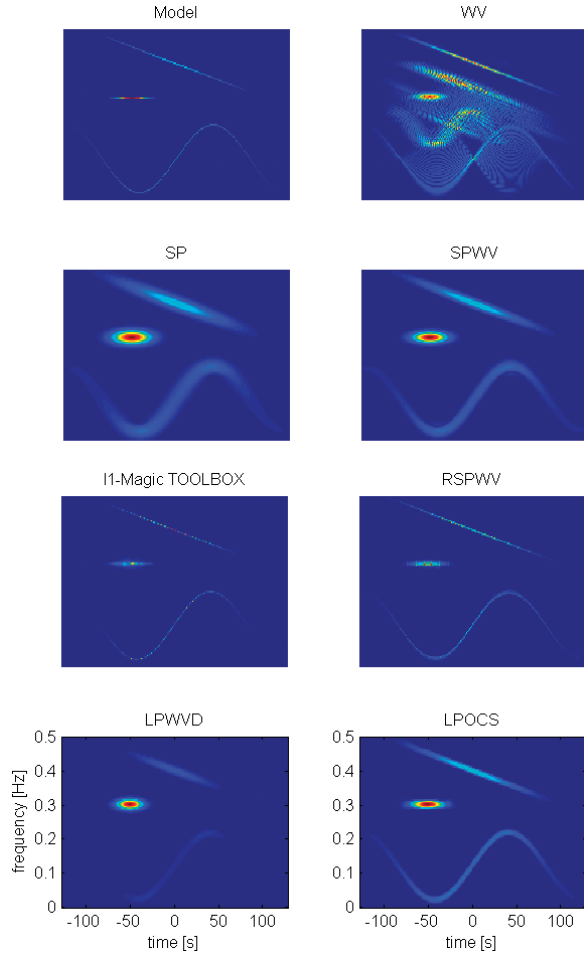


Fig. 8. Example 2: TF reconstruction using various methods. Left column: Ideal model, spectrogram,  $l_1$ -MAGIC TOOLBOX, LPWVD; right column: WV distribution, SPWV, RSPWV, lifted POCS. Frequency is normalized frequency.

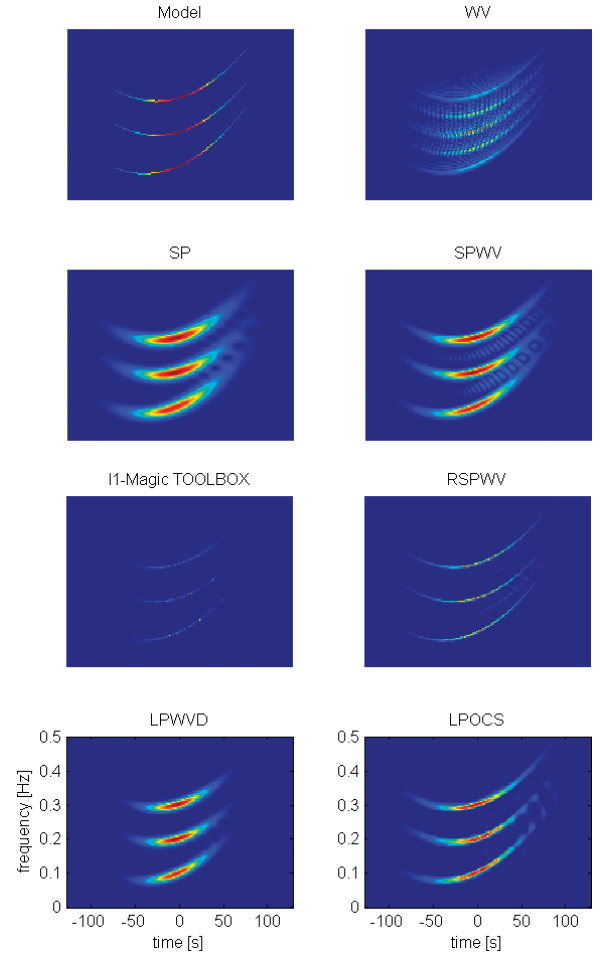


Fig. 10. Example 3: TF reconstruction using various methods. Left column: Ideal model, spectrogram,  $l_1$ -MAGIC TOOLBOX, LPWVD; right column: WV distribution, SPWV, RSPWV, lifted POCS. Frequency is normalized frequency.

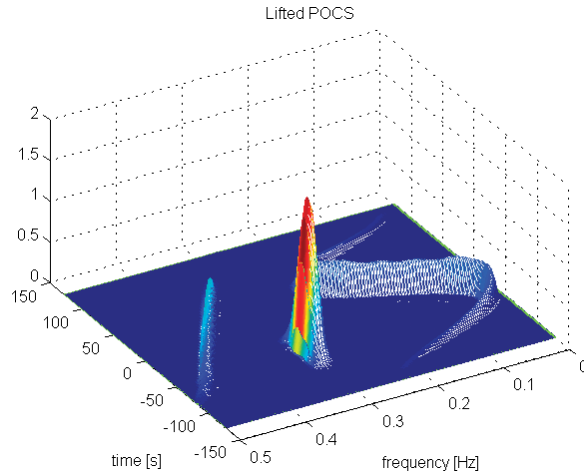


Fig. 9. 3-D plot of estimated TF distribution of Example 2 with lifted POCS method. Frequency is normalized frequency.

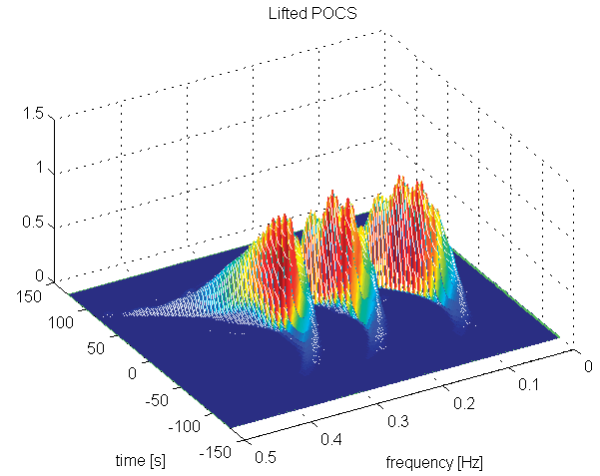


Fig. 11. 3-D plot of estimated TF distribution of Example 3 with lifted POCS method. Frequency is normalized frequency.

in [27] are also used here. Reconstruction results are shown in Figs. 6–13. In all the examples, the set  $\Omega$  is chosen as a circular mask around the origin in the ambiguity domain. The radius of the mask is selected as  $r = n/16$ , where  $n$  is the length of the discrete-time signal

as in [27]. Together with model TF and Lifted POCS, the results obtained using spectrogram (SP), SPWV, RSPWV[39], LPWVD [36], and TF reconstruction using  $l_1$ -MAGIC TOOLBOX (interior point methods) as in [27] are shown in Figs. 6, 8, 10, 12. For the purpose of

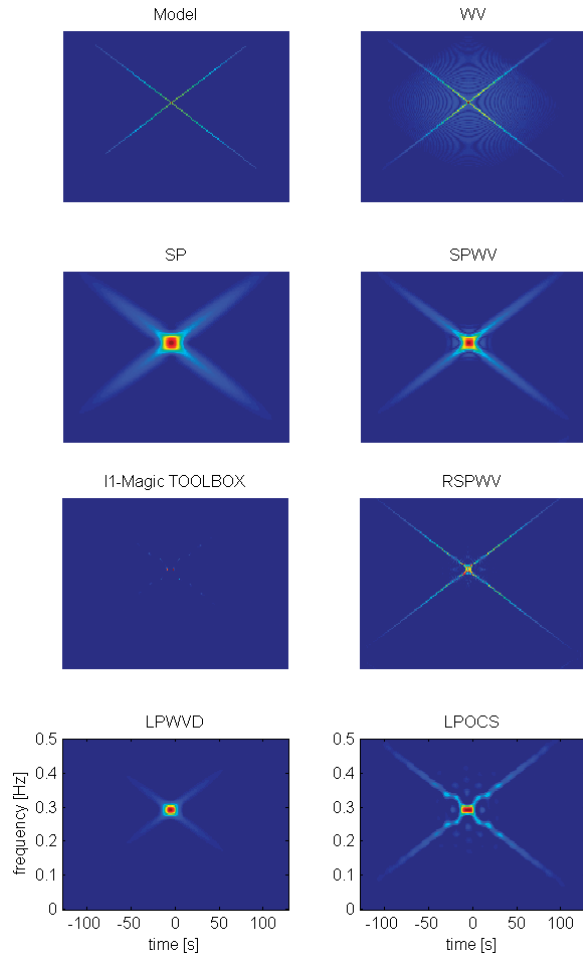


Fig. 12. Example 4: TF reconstruction using various methods. Left column: Ideal model, spectrogram,  $l_1$ -MAGIC TOOLBOX, LPWVD; right column: WV distribution, SPWV, RSPWV, lifted POCS. Frequency is normalized frequency.

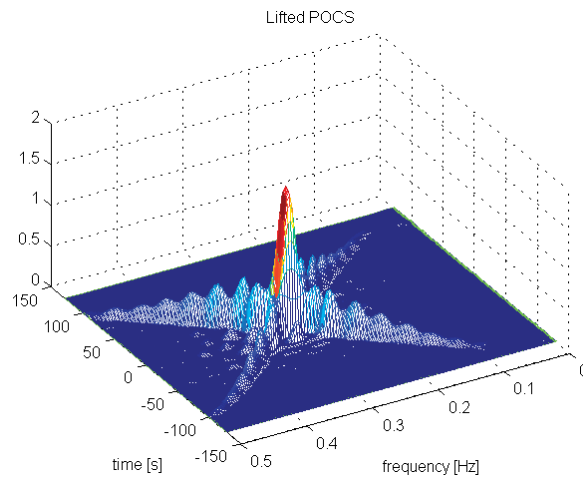


Fig. 13. 3-D plot of estimated TF distribution of Example 4 with lifted POCS method. Frequency is normalized frequency.

comparison, the desired ideal TF model of the signals is also included in the figures. The TF model is nothing but the TF constructed from the IF law of the signal components scaled by their power, as in (6). Though not all examples are polynomial phase signals and have

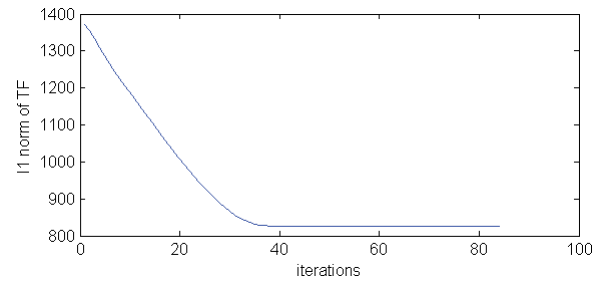


Fig. 14. Convergence plot of lifted POCS iterations for Example 1. Figure shows  $l_1$ -norm of TF distribution versus number of iterations.

time-varying amplitude, LPWVD [36] with order 6 was also used to obtain the related TF distribution. The related MATLAB code was obtained from [36]. In order to get, on average, good results, the SPWV time-smoothing window length for all the example signals at hand was set to the odd integer closest to  $n/10$ , and the length of the frequency smoothing filter at the time domain was set to the odd integer closest to  $n/4$ , with  $n$  being the signal length. In this way, any parameter adaptation to the signal was avoided.

The convergence of the lifted POCS method is monitored with the help of normalized error, defined by

$$err = \frac{\|vect(\mathbf{P}_i) - vect(\mathbf{P}_{i-1})\|_2}{\|vect(\mathbf{P}_{i-1})\|_2}. \quad (19)$$

The  $l_1$ -norm of the TF distribution versus the number of iterations is shown in Fig. 14 for Example 1.

Reconstruction results in Figs. 6–13 show that the solutions obtained with  $l_1$ -MAGIC TOOLBOX are too sparse. As pointed out by Borgnat and Flandrin [27] they cannot be accepted as a TF representation of the signal. RSPWV has good localization and better smoothness than the  $l_1$ -MAGIC TOOLBOX result, but it also has a spiky nature, as shown in Fig. 4. On the other hand, the lifted POCS (LPOCS) method generates better and acceptable results without adjusting any parameters during the optimization process. In this respect, the LPOCS method provides a good compromise between localization and smoothness, which is a physical property of the original signal. Both SPWV and LPOCS have good resolution and smoothness based on visual comparison. But the resolution of LPOCS is better than that of SPWV. SPWV additionally requires the time and frequency window lengths to be adapted to the signal for good resolution.

In Fig. 15, a reconstructed TF example obtained from a noisy signal is shown. The time-varying signal in Fig. 8 was corrupted by additive zero-mean white Gaussian noise. The signal-to-noise ratio is 10 dB. Signal auto-terms are clearly reconstructed and cross-terms suppressed by the lifted POCS method in Fig. 15 (bottom right). The result is comparable to that of the reassigned spectrum.

In Fig. 16, a signal example from frequency-hopping/ $M$ -ary frequency-shift-keyed (FH/MFSK) communication is shown. It has been shown [56] that using a cross-term-free TF representation, the parameters of

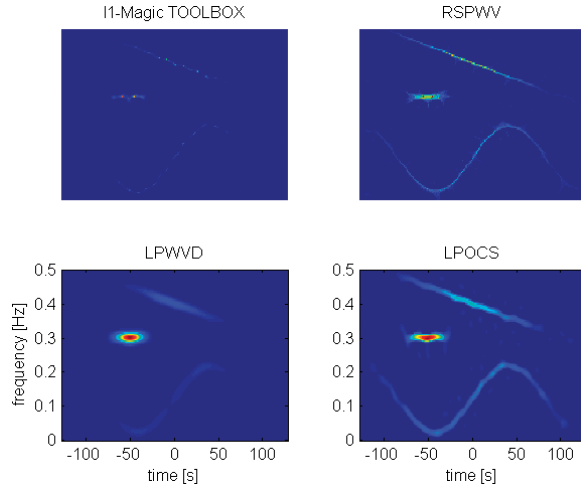


Fig. 15. Signal in Fig. 8 corrupted by additive zero-mean white Gaussian noise. Signal-to-noise ratio is 10 dB. TF reconstruction result obtained by lifted POCS method (bottom right) is comparable to reassigned smoothed pseudo-WV (top right). Frequency is normalized frequency.

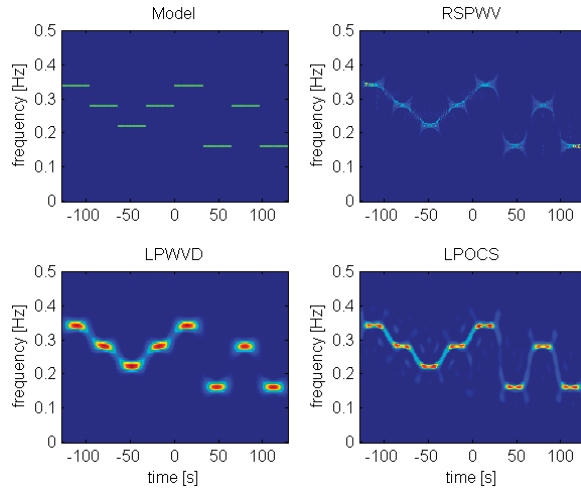


Fig. 16. Example 5: TF distribution of a frequency-hopping/ $M$ -ary frequency-shift-keyed signal. Top left: Ideal model. Top right: RSPWV. Bottom left: LPWVD. Bottom right: Lifted POCS solution. Frequency is normalized frequency.

FH/MFSK signal, which include hopping frequencies, hopping rate, hopping sequence, and modulation type can be estimated without making any assumption about the alphabet of hopping frequencies or the synchronization. It is observed that the LPOCS method clearly reveals the hopping frequencies and the hopping rate without adjusting any parameters. It provides better localization than LPWVD and is not spiky.

In Fig. 17, the TF of a short segment from a dolphin-click signal is shown. It is known that this acoustic signal should have three FM components starting at 0.1 Hz, 0.18 Hz, and 0.29 Hz (normalized frequency)—corresponding to actual 1100 Hz, 1980 Hz, and 3190 Hz, respectively—in the first half of the observation duration. Only the spectrogram and the

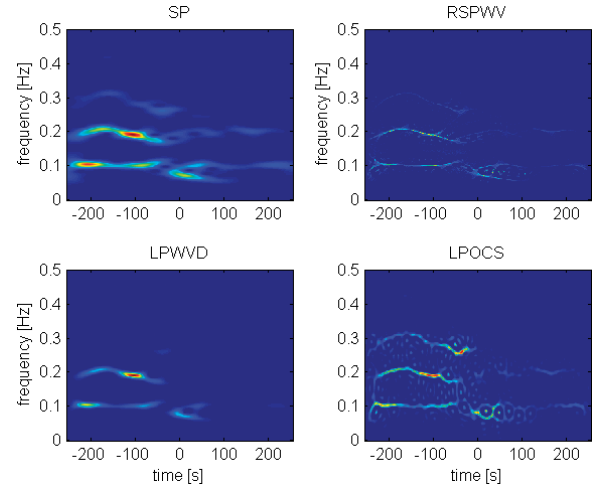


Fig. 17. Example 6: TF distribution of dolphin-click signal segment. Top left: Spectrogram. Top right: RSPWV. Bottom left: LPWVD. Bottom right: Lifted POCS solution. Frequency is normalized frequency.

LPOCS solutions reveal these three components clearly, and the LPOCS has a better resolution.

In order to measure the localization of each TF distribution in a quantitative way, we use the  $l_1$ -norm as a measure. Rényi entropy [57] is also a preferred method for measuring the localization. Rényi entropy is given by

$$R_P^\alpha = \frac{1}{1-\alpha} \log_2 \left\{ \sum_{n=0}^{N-1} \sum_{m=1}^{M-1} P^\alpha[n, m] \right\}, \quad (20)$$

where  $P[n, m]$  is the TF distribution and  $\alpha$  is the order of measure. Rényi entropy allows TF distribution to take negative values. The value of Rényi entropy is expressed in terms of bits. The lower the Rényi measure, the better the localization. A Rényi entropy of order 3 has been shown to be a good measure for localization [57].

Localization alone is not sufficient for a good comparison. We also need to know how similar the TF result is to the model TF we desire. Therefore, together with Rényi entropy, the Pearson correlation of the solution TF and the model TF is also used as a measure of similarity. The Pearson correlation between solution TF  $\mathbf{P}$  and model TF  $\mathbf{P}_{\text{model}}$  is given by

$$pcor = \frac{\vec{v}ec(\mathbf{P})^T \vec{v}ec(\mathbf{P}_{\text{model}})}{\|\vec{v}ec(\mathbf{P})\|_2 \|\vec{v}ec(\mathbf{P}_{\text{model}})\|_2}, \quad (21)$$

where  $\vec{v}ec(\mathbf{P})$  represents the vector form of  $\mathbf{P}$  with the mean value subtracted. Pearson correlation measures the shape similarity rather than any exact norm difference. A value of 1 indicates an exact shape match.

In Table I, the Pearson correlations between the model and the solution TF are computed, and in Table II, Rényi entropy is provided for all the examples studied in this paper.

For a meaningful comparison, we first should look at the final solution in terms of acceptability as a TF distribution related to the signal. Therefore, we should first check how similar the result is to the desired model. Then

TABLE I  
Pearson Correlation Between TF Distributions and the Model TF for  
Tested Examples

pcor	Model	WV	SP	SPWV	l1-Magic	RSPWV	LPWVD	LPOCS
Ex1	1,00	0,35	0,41	0,49	0,49	<b>0,73</b>	0,47	0,59
Ex2	1,00	0,24	0,38	0,43	0,29	<b>0,52</b>	0,40	0,50
Ex3	1,00	0,18	0,39	0,45	0,31	0,47	0,44	<b>0,55</b>
Ex4	1,00	<b>0,45</b>	0,34	0,40	0,09	0,42	0,28	0,40
Ex5	1,00	0,11	0,34	0,35	0,20	0,34	0,38	<b>0,44</b>
Ex6	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A

Note: A higher value shows better similarity to the model.

TABLE II  
Rényi Entropy of All the TF Distributions for Tested Examples

Rényi(3)	Model	WV	SP	SPWV	l1-Magic	RSPWV	LPWVD	LPOCS
Ex1	8,13	9,01	11,51	10,97	<b>7,18</b>	8,47	10,44	10,14
Ex2	<b>7,46</b>	9,56	11,18	10,66	8,06	8,74	8,84	9,99
Ex3	8,24	9,20	11,77	11,18	<b>7,59</b>	8,82	10,69	10,35
Ex4	8,04	8,51	11,01	10,41	<b>5,45</b>	8,57	8,46	9,65
Ex5	7,99	9,77	11,88	11,23	<b>6,53</b>	9,45	10,86	10,23
Ex6	N/A	10,99	13,11	12,39	9,53	10,73	11,28	11,70

Note: A lower value indicates better localization.

we consider the localization. From Table I we observe that the LPOCS and RSPWV methods are better than all the other methods in terms of the similarity measure.

When we compare the localization property of the methods, the TF obtained by  $l_1$ -MAGIC TOOLBOX has the highest localization. The second best is RSPWV, and third is LPOCS. But as we emphasized in Figs. 3 and 4, TF solutions obtained by RSPWV and  $l_1$ -MAGIC TOOLBOX methods are spiky and do not correspond to the physical reality of the actual signals [27]. In fact, from Table II we observe that  $l_1$ -Magic TOOLBOX provides overlocalized results which have lower Rényi entropy than the actual model TF in some cases. In Example 6, shown in Fig. 17, we observe that RSPWV is too weak to produce the spectral lines clearly. The LPOCS method has good results in terms of localization, similarity, and physical interpretation. Furthermore, the LPOCS method does not require any parameter adjustment nor parameter selection. Our overall assessment is that LPOCS is superior to RSPWV. However, the computational cost of RSPWV is lower than LPOCS.

## VI. CONCLUSIONS AND FUTURE WORK

In both the proposed lifted POCS method and in [27], the cardinality of the set  $\Omega$  is very low compared to the actual TF signal ( $n \times n$ ). In all the examples tried in this paper, the size of the  $\Omega$  set is selected as a circle with radius  $r = n/16$ . This is necessary to remove the

cross-terms [27]. When the sets  $C_f$  and  $C_{AF}$  intersect, there may be many solutions satisfying the constraints specified by those sets. In this case, the solution depends on the initial vector. In all the examples, iterations start with a 2-D distribution obtained from the actual AF by a masking window with radius  $r = n/16$ . In Example 1, the starting distribution for the iterative lifted POCS approach is shown in the bottom-left plot of Fig. 1. In other example, the related masked AF is used as the initial estimate. They are all relatively smooth VW distributions. When the sets  $C_f$  and  $C_{AF}$  do not intersect, the iterations converge to either one of the two unique distributions, as illustrated in Fig. 5.

We also tried different mask sizes in  $C_{AF}$ , ranging from  $r = n/12$  to  $n/24$ . This range of masks successfully removed cross-terms in all cases. Therefore, the choice of  $r$  is not very critical in estimating a cross-term-free VW distribution. However, the question of the optimal  $r$  value or the shape of the mask  $\Omega$  for a given time-varying signal remains an open problem.

Both for the lifted POCS method and for the method in [27], the computational cost is high compared to the classical VW and other AF shaping- or smoothing-based methods. This is because the optimization problems posed by both methods are solved in an iterative manner. Obviously, estimated VW distributions are better in terms of cross-terms compared to classical methods. Considering the good localization achieved in the TF plane without any regularization parameter, the lifted POCS method is a promising approach for TF distribution estimation.

The lifted POCS method has a flexible structure. It also allows other constraints to be incorporated into the reconstruction problem. One example is using the frequency marginal requirement for each time slice of the TF distribution [50] as a constraint. In each iteration cycle it is possible to satisfy this requirement.

## APPENDIX THE PROJECTION ONTO EPIGRAPH SET OF A CONVEX FUNCTION

In the lifted POCS method, orthogonal projection onto sets  $C_f$  and  $C_{AF}$  has to be performed. In this appendix, the projection operation onto the epigraph set  $C_f$  of a convex function  $f$  is described. While the projection onto the measurement set  $C_{AF}$  is obtained using Fourier transform relations in (16)–(18), the projection onto  $C_f$  given in (12) cannot be obtained in a closed form. The projection onto  $C_{AF}$  is implemented using successive projections onto supporting hyperplanes.

Given  $\mathbf{x} \in \mathbf{R}^n$ , with  $f : \mathbf{R}^n \rightarrow \mathbf{R}$  a convex function, the epigraph set of the function  $f$  is defined as

$$C_f = \left\{ \mathbf{w} = [\mathbf{x}^T v]^T \in \mathbf{R}^{n+1} \mid f(\mathbf{x}) \leq v \right\}, \quad (22)$$

where  $\mathbf{w} \in \mathbf{R}^{n+1}$  is a vector defined in the lifted domain and  $v \in \mathbf{R}$  is the last element of  $\mathbf{w}$ .

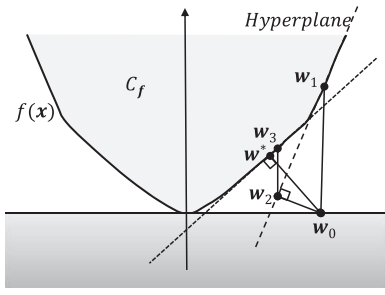


Fig. 18. Projection onto epigraph set  $C_f$  by successive projections onto supporting hyperplanes.

Given the initial point  $w_0 = [x_0^T \ 0]^T$ , a supporting hyperplane for  $C_f$  is defined at  $x_0$ . The supporting hyperplane is the set of points in  $\mathbf{R}^{n+1}$  satisfying  $a^T w = b$ , where  $a$  and  $b$  are given by

$$a = \begin{bmatrix} \nabla f(x_0) \\ -1 \end{bmatrix} \in \mathbf{R}^{n+1} \quad (23)$$

and

$$b = a^T \begin{bmatrix} x_0 \\ f(x_0) \end{bmatrix} \in \mathbf{R} \quad (24)$$

and  $\nabla f(x_0)$  is the gradient of the cost function  $f$  at  $x_0$ . The supporting hyperplane, as shown in Fig. 18, is tangent to  $C_f$  at  $w_1 = [x_0^T \ f(x_0)]^T$ . The vector  $w_0$  is projected onto this hyperplane and  $w_2 = [x_2^T \ v_2]^T$  is obtained.

Then a second supporting hyperplane is defined at  $x_2$ . This second hyperplane is again tangent to  $C_f$  at  $w_3 = [x_2^T \ f(x_2)]^T$ . The vector  $w_0$  is reprojected onto the second hyperplane and  $w_4 = [x_4^T \ v_4]^T$  is obtained. This iteration continues until the projected point  $w_k$  satisfies  $w^* = [x^{*T} \ f(x^*)]^T \in C_f$ . Since this is an iterative process, the iterations are stopped after a fixed number of steps or once there is no improvement between consecutive steps. Once an increase in distance is detected, a refinement should be done for the point at which the hyperplane is defined.

The distance  $\|w_i - w_0\|_2$  between the point to be projected and the current projection will not always decrease for high values of iteration  $i$ ; therefore, the distance need to be monitored. In this case, the hyperplane for  $f$  should be defined at the point  $(x_i + x_{i-1})/2$ .

If the gradient  $\nabla f(x_0)$  is not computable, then the concept of the subgradient can be used to determine a supporting hyperplane at  $x_0$ .

## REFERENCES

- [1] Gaunard, G. C., and Strifors, H. C. Signal analysis by means of time–frequency (Wigner-type) distributions—Applications to sonar and radar echoes. *Proceedings of the IEEE*, **84**, 9 (Sept. 1996), 1231–1248.
- [2] Chen, V. C., and Qian S. Joint time–frequency transform for radar range–Doppler imaging. *IEEE Transactions on Aerospace and Electronic Systems*, **34**, 2 (Apr. 1998), 486–499.
- [3] Xia, X.-G., Wang, G., and Chen, V. C. Quantitative SNR analysis for ISAR imaging using joint time–frequency analysis—Short time Fourier transform. *IEEE Transactions on Aerospace and Electronic Systems*, **38**, 2 (Apr. 2002), 649–659.
- [4] Wehner, D. R. *High-Resolution Radar* (2nd ed.). Boston: Artech House, 1994, chs. 6 and 7.
- [5] Li, Y., Xing, M., Su, J., Quan, Y., and Bao, Z. A new algorithm of ISAR imaging for maneuvering targets with low SNR. *IEEE Transactions on Aerospace and Electronic Systems*, **49**, 1 (Jan. 2013), 543–557.
- [6] Wang, Y., Ling, H., and Chen, V. C. ISAR motion compensation via adaptive joint time–frequency technique. *IEEE Transactions on Aerospace and Electronic Systems*, **34**, 2 (Apr. 1998), 670–677.
- [7] Chen, V. C., Li, F., Ho, S.-S., and Wechsler, H. Micro-Doppler effect in radar: Phenomenon, model, and simulation study. *IEEE Transactions on Aerospace and Electronic Systems*, **42**, 1 (Jan. 2006), 2–21.
- [8] Whiteloni, N., and Ling, H. Radar signature analysis using a joint time–frequency distribution based on compressed sensing. *IEEE Transactions on Antennas and Propagation*, **62**, 2 (Feb. 2014), 755–763.
- [9] Lim, H., Park, J. H., Yoo, J. H., Kim, C. H., Kwon, K. I., and Myung, N. H. Joint time–frequency analysis of radar micro-Doppler signatures from aircraft engine models. *Journal of Electromagnetic Waves and Applications*, **25**, 8–9 (2011), 1069–1080.
- [10] Gao, H., Xie, L., Wen, S., and Kuang, Y. Micro-Doppler signature extraction from ballistic target with micro-motions. *IEEE Transactions on Aerospace and Electronic Systems*, **46**, 4 (Oct. 2010), 1969–1982.
- [11] Stankovic, L., Thayakaran, T., Dakovic, M., and Popovic-Bugarin, V. Micro-Doppler removal in the radar imaging analysis. *IEEE Transactions on Aerospace and Electronic Systems*, **49**, 2 (Apr. 2013), 1234–1250.
- [12] Stankovic, L., Djurovic, I., and Thayakaran, T. Separation of target rigid body and micro-Doppler effects in ISAR imaging. *IEEE Transactions on Aerospace and Electronic Systems*, **42**, 4 (Oct. 2006), 1496–1506.
- [13] de Luigi C., and Jauffret, C. Estimation and classification of FM signals using time frequency transforms. *IEEE Transactions on Aerospace and Electronic Systems*, **41**, 2 (Apr. 2005), 421–437.
- [14] Lopez-Risueno, G., Grajal, J., and Sanz-Orsorio, A. Digital channelized receiver based on time–frequency analysis for signal interception. *IEEE Transactions on Aerospace and Electronic Systems*, **41**, 3 (July 2005), 879–898.
- [15] Blodt, M., Chabert, M., Regnier, J., and Faucher, J. Mechanical load fault detection in induction motors by stator current time–frequency analysis. *IEEE Transactions on Industry Applications*, **42**, 6 (Nov.–Dec. 2006), 1454–1463.

- [16] Kliman, G. B., and Stein, J.  
Methods of motor current signature analysis.  
*Electric Machines and Power Systems*, **20**, 5 (1992), 463–474.
- [17] Yazici, B., and Kliman, G. B.  
An adaptive statistical time–frequency method for detection of broken bars and bearing faults in motors using stator current.  
*IEEE Transactions on Industry Applications*, **35**, 2 (Mar.–Apr. 1999), 442–452.
- [18] Rajagopalan, S., Restrepo, J. A., Aller, J. M., Habetler, T. G., and Harley, R. G.  
Nonstationary motor fault detection using recent quadratic time–frequency representations.  
*IEEE Transactions on Industry Applications*, **44**, 3 (May–June 2008), 735–744.
- [19] Claasen, T. A. C. M., and Mecklenbrauker, W. F. G.  
The Wigner distribution—A tool for time–frequency signal analysis. Part III: Relations with other time–frequency signal transformations.  
*Philips Journal of Research*, **35**, 6 (1980), 372–389.
- [20] Boudreaux-Bartels, G., and Parks, T. W.  
Time-varying filtering and signal estimation using Wigner distribution synthesis techniques.  
*IEEE Transactions on Acoustics, Speech and Signal Processing*, **34**, 3 (June 1986), 442–451.
- [21] Cohen, L.  
Time–frequency distributions—A review.  
*Proceedings of the IEEE*, **77**, 7 (July 1989), 941–981.
- [22] Cohen, L.  
What is a multicomponent signal?  
In *IEEE International Conference on Acoustics, Speech, and Signal Processing*, San Francisco, CA, Mar. 1992, **5**, 113–116.
- [23] Krattenthaler, W., and Hlawatsch, F.  
Time–frequency design and processing of signals via smoothed Wigner distributions.  
*IEEE Transactions on Signal Processing*, **41**, 1 (Jan. 1993), 278–287.
- [24] Candès, E. J., Romberg, J., and Tao, T.  
Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information.  
*IEEE Transactions on Information Theory*, **52**, 2 (Feb. 2006), 489–509.
- [25] Donoho, D. L.  
Compressed sensing.  
*IEEE Transactions on Information Theory*, **52**, 4 (Apr. 2006), 1289–1306.
- [26] Candès, E. J., and Tao, T.  
Near-optimal signal recovery from random projections: Universal encoding strategies?  
*IEEE Transactions on Information Theory*, **52**, 12 (Dec. 2006), 5406–5425.
- [27] Flandrin, P., and Borgnat, P.  
Time–frequency energy distributions meet compressed sensing.  
*IEEE Transactions on Signal Processing*, **58**, 6 (June 2010), 2974–2982.
- [28] Candès, E. J., and Romberg, J.  
l1-MAGIC: Recovery of sparse signals via convex programming.  
[Online]. Available: <http://www.acm.caltech.edu/l1magic/2005>.
- [29] Candès, E. J., and Romberg, J.  
Practical signal recovery from random projections.  
*Proceedings of the SPIE*, **5674** (2005), 76–86.
- [30] Bregman, L.  
Finding the common point of convex sets by the method of successive projection.  
*Doklady Akademii nauk SSSR*, **7**, 3 (1965), 200–217.
- [31] Youla, D. C., and Webb, H.  
Image restoration by the method of convex projections: Part 1—Theory.  
*IEEE Transactions on Medical Imaging*, **1**, 2 (Oct. 1982), 81–94.
- [32] Hlawatsch, F., and Boudreaux-Bartels, G. F.  
Linear and quadratic time–frequency signal representations.  
*IEEE Signal Processing Magazine*, **9**, 2 (Apr. 1992), 21–67.
- [33] Cohen, L.  
*Time–Frequency Analysis*. Englewood Cliffs, NJ: Prentice Hall, 1995.
- [34] Durak, L., and Arkan, O.  
Short-time Fourier transform: Two fundamental properties and an optimal implementation.  
*IEEE Transactions on Signal Processing*, **51**, 5 (May 2003), 1231–1242.
- [35] Stankovic, L.  
L-class of time–frequency distributions.  
*IEEE Signal Processing Letters*, **3**, 1 (1996), 22–25.
- [36] Wang Y., and Jiang, Y. C.  
New time–frequency distribution based on the polynomial Wigner–Ville distribution and L class of Wigner–Ville distribution.  
*IET Signal Processing*, **4**, 2 (Apr. 2010), 130–136.
- [37] Boashash, B., and O’Shea, P.  
Polynomial Wigner–Ville distributions and their relationship to time-varying higher order spectra.  
*IEEE Transactions on Signal Processing*, **42**, 1 (Jan. 1994), 216–220.
- [38] Ram, S. S., and Ling, H.  
Application of the reassigned joint time–frequency transform to wideband scattering from waveguide cavities.  
*IEEE Antennas and Wireless Propagation Letters*, **6** (2007), 580–583.
- [39] Flandrin, P., Auger, F., and Chassande-Mottin, E.  
Time–frequency reassignment: From principles to algorithms.  
In *Applications in Time–Frequency Signal Processing*. A. Papandreou-Suppappola, Ed. Boca Raton, FL: CRC Press, 2003, ch. 5, 179–203.
- [40] Jones, D. L., and Baraniuk, R. G.  
An adaptive optimal-kernel time–frequency representation.  
*IEEE Transactions on Signal Processing*, **43**, 10 (Oct. 1995), 2361–2371.
- [41] Chen, S. S., Donoho, D. L., and Saunders, M. A.  
Atomic decomposition by basis pursuit.  
*SIAM Journal on Scientific Computing*, **20**, 1 (1998), 33–61.
- [42] Çetin, A. E.  
Reconstruction of signals from Fourier transform samples.  
*Signal Processing*, **16**, 2 (Feb. 1989), 129–148.
- [43] Auger, F., Flandrin, P., Gonçalves, P., and Lemoine, O.  
Time-Frequency Toolbox. Available: <http://tftb.nongnu.org>.
- [44] Çetin, A. E., Bozkurt, A., Gunay, O., Habiboğlu, Y. H., Köse, K., Onaran, I., Tofighi, M., and Sevimli, R. A.  
Projections onto convex sets (POCS) based optimization by lifting.  
In *2013 IEEE Global Conference on Signal and Information Processing*, Austin, TX, Dec. 3–5, 2013.
- [45] Chierchia, G., Pustelnik, N., Pesquet, J.-C., and Pesquet-Popescu, B.  
An epigraphical convex optimization approach for multicomponent image restoration using non-local structure tensor.  
In *2013 IEEE International Conference on Acoustics, Speech and Signal Processing*, Vancouver, Canada, May 2013, 1359–1363.

- [46] Gubin, L. G., Polyak, B. T., and Raik, E. V.  
The method of projections for finding the common point of convex sets.  
*USSR Computational Mathematics and Mathematical Physics*, **7**, 6 (1967), 1–24.
- [47] Combettes, P. L. and Pesquet, J.  
Image restoration subject to a total variation constraint.  
*IEEE Transactions on Image Processing*, **13**, 9 (2004), 1213–1222.
- [48] Çetin, A. E., Gerek, O. N., and Yardimci, Y.  
Equiripple FIR filter design by the FFT algorithm.  
*IEEE Signal Processing Magazine*, **14**, 2 (Mar. 1997), 60–64.
- [49] Kose, K., Cevher, V., and Çetin, A. E.  
Filtered variation method for denoising and sparse signal processing.  
In *2012 IEEE International Conference on Acoustics, Speech and Signal Processing*, Kyoto, Japan, Mar. 2012, 3329–3332.
- [50] Loughlin, P. J., Pitton, J. W., and Atlas, L. E.  
Construction of positive time–frequency distributions.  
*IEEE Transactions on Signal Processing*, **42**, 10 (Oct. 1994), 2697–2705.
- [51] Combettes, P. L.  
The foundations of set theoretic estimation.  
*Proceedings of the IEEE*, **81**, 2 (Feb. 1993), 182–208.
- [52] Sezan, M. I., and Stark, H.  
Image restoration by the method of convex projections: Part 2—Applications and numerical results.  
*IEEE Transactions on Medical Imaging*, **1**, 2 (1982), 95–101.
- [53] Stark, H. (Ed.) *Image Recovery: Theory and Application*. Orlando, FL: Academic Press, 1987.
- [54] Bauschke, H. H., and Combettes, P. L.  
*Convex Analysis and Monotone Operator Theory in Hilbert Spaces*. New York: Springer, 2011.
- [55] Chambolle, A.  
An algorithm for total variation minimization and applications.  
*Journal of Mathematical Imaging and Vision*, **20**, 1–2 (Jan. 2004), 89–97.
- [56] Chen, T.-C.  
Joint signal parameter estimation of frequency-hopping communications.  
*IET Communications*, **6**, 4 (Mar. 2012), 381–389.
- [57] Baraniuk, R. G., Flandrin, P., Janssen, A. J. E. M., and Michel, O. J. J.  
Measuring time–frequency information content using the Rényi entropies.  
*IEEE Transactions on Information Theory*, **47**, 4 (May 2001), 1391–1409.



**Zeynel Deprem** got his M.Sc. degree from Middle East Technical University, Ankara, in 1993. He worked in Turk Telekom’s R&D center between 1990 and 1998. He is currently pursuing a Ph.D. degree from Bilkent University and also working at Turk Telekom, mainly in access network planning projects.



**A. Enis Çetin** got his Ph.D. degree from the University of Pennsylvania in 1987. Between 1987 and 1989, he was an assistant professor of electrical engineering at the University of Toronto. He has been with Bilkent University, Ankara, Turkey, since 1989. Çetin was an associate editor of *IEEE Transactions on Image Processing* between 1999 and 2003. Currently, he is on the editorial boards of *IEEE Signal Processing Magazine*, *IEEE Transactions on Circuits and Systems for Video Technology*, and *Machine Vision and Applications* (IAPR), Springer. He is the editor-in-chief of *Signal, Image and Video Processing*, Springer. He is a fellow of IEEE. His research interests include signal and image processing, human–computer interaction using vision and speech, and audiovisual multimedia databases.