The Self-Duality of Discrete Short-Time Fourier Transform and Its Applications

Tzu-Hsien Sang, *Member, IEEE*

***Abstract*—The self-duality of short-time Fourier transform (STFT) is an elegant property and is useful in shedding light on the construction of STFT and its resolution capability. In this paper, the discrete version of self-duality is studied, and the property is interpreted in the context of resolution capabilities of time frequency distributions. In addition, two applications are provided as showcases of these insights obtained from the interpretation. In the first application, the problem of STFT synthesis is considered, and self-duality serves as an important indication of whether the synthesis problem at hands is properly formulated. In the second application, a new kind of high-resolution time-frequency distribution is constructed based on the understandings obtained by contrasting two of the most popular time-frequency analysis tools, namely, the STFT and the Wigner distribution.**

***Index Terms*—Discrete short-time Fourier transform (STFT), self-duality, signal synthesis, time-frequency analysis.**

I. Introduction

THE self-duality of short-time Fourier transform (STFT) has been known for some time, for example, in the context of radar signal processing [1]. It is an elegant result, yet the usefulness of it has been far from fully demonstrated. To describe the self-duality of STFT of continuous signals, a modified definition of STFT was used:

$$G_{x,h}(t,\omega) = e^{j(1/2)t\omega}\frac{1}{\sqrt{2\pi}}\int x(\tau)h^*(t-\tau)e^{-j\omega\tau}\,d\tau \quad (1)$$

where $x(t)$ is the signal-to-be-analyzed, $h(t)$ is the window function, and $G_{x,h}(t,\omega)$ denote the STFT. Notice the extra phase term $e^{j(1/2)t\omega}$ which guarantees the symplectic covariance [2]–[4] and does not exist in the usual definition. With this little change, the following equation is easily shown to hold:

$$\frac{1}{2\pi}\int\int G_{x,h}(t,\omega)e^{j\theta t}e^{j\tau\omega}\,dt\,d\omega = 2G_{x,\tilde{h}}(2\tau,-2\theta) \quad (2)$$

where $\tilde{h}$ is the time-reversal of h. In particular, if $h(t)$ is time-symmetric, then the two-dimensional inverse Fourier transform of the STFT is proportional to the scaled version of the STFT itself; hence we call the relation self-duality. The short-hand notation $(t,\omega) \longleftrightarrow (2\tau,-2\theta)$ is used to denote such a relation.

Notice that the commonly known Fourier duality between any two orthogonal directions on the time-frequency plane refers to the conjugation of pairs of operations along each direction. A quick review can be found in [5]. Here the self-duality refers to, however, the functional resemblance between the Fourier pairs in the $t-\omega$ and $\theta-\tau$ domain. One typical Fourier dual pair is the scaling operation, namely, expanding a function in one domain leads to the contraction of its Fourier pair in the other domain and vice versa. Applying the scaling duality to the STFT and its Fourier pair, one would get a duality which is, with a figurative short-hand notation, $(t,\omega) \longleftrightarrow (1/\theta, 1/\tau)$. Apparently this relation is quite different from the self-duality which reads $(t,\omega) \longleftrightarrow (2\tau,-2\theta)$. Nevertheless, they are not conflicting each other, as matters will be explained with more details in Section II.

In this paper, the self-duality is established for STFT based on discrete-time Fourier transform (DTFT). In Section II sufficient conditions for this version self-duality to hold will be identified and the implications will be discussed. A pure discrete version of self-duality of STFT based on discrete Fourier transform (DFT) is also formulated for the implementation purpose. It can be viewed as an approximation to the exact self-duality, as long as the sampling and truncation of the signal and window function are done with precautions. Next, an interesting interpretation of self-duality, focusing on the supposed poor resolution capability of STFT/spectrogram via contrasting it with the Wigner distribution (WD), raises the possibility of linking self-duality and the resolution issue. Furthermore, in Sections III and IV two algorithms based on the aforementioned understanding of self-duality were developed for applications in STFT synthesis and the construction of high-resolution time-frequency distribution (TFD). The algorithms powerfully demonstrate the beauty and usefulness of this elegant property.

II. Self-Duality of Discrete STFT

For a discrete signal $x(n)$, the STFT based on DTFT is defined as

$$G(n,\omega) = e^{j(1/2)n\omega}\sum_{m=-\infty}^{\infty} x(m)h^*(n-m)e^{-j\omega m} \quad (3)$$

in which the signal $x(n)$ and the window function $h(n)$ are assumed to be band-limited within frequency bands $[-B,B]$ and $[-W,W]$ respectively, and both are properly sampled. The slice of $G(n,\omega)$ along the frequency axis is a continuous and periodic function, while along the time axis is a discrete sequence.

Manuscript received October 21, 2008; accepted August 20, 2009. First published September 09, 2009; current version published January 13, 2010. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Maria Hansson-Sandsten. This work was supported by NSC of Taiwan, R.O.C., under Grant NSC 98-2220-E-009-043.

The author is with the Department of Electronics Engineering, National Chiao-Tung University, Hsin-Chu 300, Taiwan, R.O.C. (e-mail: tzuhsien54120@faculty.nctu.edu.tw).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TSP.2009.2032038

To demonstrate the self duality, first take a mixed two-dimensional inverse Fourier transform of $G(n,k)$:

$$\hat{G}(\theta,q) \equiv \frac{1}{2\pi}\int_{\omega=-\pi}^{\pi}\sum_{n=-\infty}^{\infty} G(n,\omega)e^{j\theta n}e^{jq\omega}d\omega. \quad (4)$$

It is now shown that $\hat{G}(\theta,q) = 2G(2q,-2\theta)$ if certain conditions are met. Continue with (4) [see (5), shown at the bottom of the page]. The inner-most integration produces a sinc function; therefore,

$$\hat{G}(p,q) = \sum_{m=-\infty}^{\infty} x(m) \sum_{n=-\infty}^{\infty} h^*(n-m)e^{j\theta n}\mathrm{sinc}\left(m-q-\frac{n}{2}\right). \quad (6)$$

The inner summation can be viewed as the modulated sequence $h^*(n)e^{j\theta n}$ being convolved with a low-pass sinc filter. At the first glance, it seems necessary that $h^*(n)e^{j\theta n}$ shall be within the pass-band of the sinc filter, i.e., $W + |\theta| \leq \pi/2$, to further proceed the derivation of self-duality. With a closer scrutiny, however, it becomes clear that the condition is unnecessarily strict. The reason is that, as long as $B + W \leq \pi$, the convolution of $x(m)$ with the filtered version of $h^*(n)e^{j\theta n}$ will have the same result as in the case where $h^*(n)e^{j\theta n}$ is within the passing band of the sinc filter.

With the assumption that the afore-mentioned condition holds, we can proceed with the case where $h^*(n)e^{j\theta n}$ is band-limited within the passing band of the sinc filter and the output is

$$2h^*(m-2q)e^{j\theta(m-2q)}. \quad (7)$$

If the window function is symmetric, then

$$\begin{aligned}\hat{G}(\theta,q) &= 2e^{-2j\theta q}\sum_{m=-\infty}^{\infty} x(m)h^*(m-2q)e^{j\theta m}\\ &= 2e^{-2j\theta q}\sum_{m=-\infty}^{\infty} x(m)h^*(2q-m)e^{j\theta m}\\ &= 2G(2q,-2\theta). \end{aligned} \quad (8)$$

That is, the self-duality based on discrete-time Fourier transform is established with the sufficient conditions that both $x(n)$ and $h(n)$ are band-limited and properly sampled, and their bandwidths satisfy $B + W \leq \pi$.

To facilitate algorithm implementation, a version of approximate self-duality based on discrete Fourier transform is needed. Next, such a version will be presented without elaborations on derivation. The details pretty much follow the derivation of (8). The result can be viewed as a sampled version of $\hat{G}(\theta,q) = 2G(2q,-2\theta)$, and the self-duality holds as long as the "edge effect" caused by substituting DTFT with DFT can be ignored.

Let $x(n)$ be a length-N discrete signal $x(n)$ and its discrete STFT is defined as

$$G(n,k) = e^{j\pi nk/N}\frac{1}{\sqrt{N}}\sum_{m=-N/2}^{N/2-1} x(m)h^*(n-m)e^{-j2\pi mk/N} \quad (9)$$

for $n = -N/2,\ldots,N/2-1$, and $k = -N/2,\ldots,N/2-1$. Notice that the range of indexes is chosen such that the indexing can be simplified. The discrete inverse Fourier transform of $G(n,k)$ is

$$\hat{G}(p,q) \equiv \frac{1}{N}\sum_{n=-N/2}^{N/2-1}\sum_{k=-N/2}^{N/2-1} G(n,k)e^{j2\pi np/N}e^{j2\pi kq/N}. \quad (10)$$

The self-duality states that

$$\hat{G}(p,q) \approx 2G(2q,-2p) \quad (11)$$

if the following conditions on $x(n)$ and $h(n)$ are met[1]:

1) the time supports should satisfy $a + b \leq N/2$;
2) the frequency supports should satisfy $B + W \leq N/2$.

These sufficient conditions agree with the intuition derived from the argument of avoiding "edge effects." The discrete version of self-duality is now established. In the following, an interpretation of self-duality, mainly focusing on the resolution issue on the time-frequency plane, is provided. Furthermore, the obtained insights will be utilized to develop two example applications that powerfully demonstrate the beauty and usefulness of self-duality.

A. Interpretation of Self-Duality and the Resolution Capability

An interesting interpretation of the self-duality is presented to provide useful insights on the resolution capability of STFT, and with analogy, of general time-frequency distributions. To facilitate discussions, a single Gabor logon centered at the origin of the time-frequency plane will be used as an example, and the continuous instead of discrete version will be discussed for the simpler notations. A single Gabor logon is

$$x(t) = e^{-t^2/2\sigma^2} \quad (12)$$

[1]The window function and the signal should be essentially time and band-limited, even though in strict sense this is not possible. Let their effective time supports be $[-a,a]$ and $[-b,b]$, and the effective frequency supports $[-B,B]$ and $[-W,W]$ respectively.

$$\begin{aligned}\hat{G}(\theta,q) &= \frac{1}{2\pi}\int_{\omega=-\pi}^{\pi}\sum_{n=-\infty}^{\infty}\sum_{m=-\infty}^{\infty} x(m)h^*(n-m)e^{j(1/2)n\omega}e^{-j\omega m}e^{j\theta n}e^{jq\omega}d\omega\\ &= \frac{1}{2\pi}\sum_{n=-\infty}^{\infty}\sum_{m=-\infty}^{\infty}\int_{\omega=-\pi}^{\pi} x(m)h^*(n-m)e^{j\theta n}e^{-j(m-q-n/2)\omega}d\omega. \end{aligned} \quad (5)$$

and its STFT is

$$G(t,\omega) = e^{-t^2/4\sigma^2} e^{-\sigma^2\omega^2/4}. \tag{13}$$

$G(t,\omega)$ satisfies the self-duality

$$e^{-t^2/4\sigma^2} e^{-\sigma^2\omega^2/4} \stackrel{\mathcal{F}}{\Longleftrightarrow} e^{-\tau^2/\sigma^2} e^{-\sigma^2\theta^2}. \tag{14}$$

The corresponding spectrogram, i.e., the squared magnitude of the STFT, is

$$e^{-t^2/2\sigma^2} e^{-\sigma^2\omega^2/2}. \tag{15}$$

In addition, the WD

$$W(t,\omega) \equiv \frac{1}{2\pi} \int x\left(t+\frac{\tau}{2}\right) x^*\left(t-\frac{\tau}{2}\right) e^{-j\tau\omega}\, d\tau \tag{16}$$

will be generated to demonstrate its superior resolution, which is regarded as its trademark [6]–[10], and its different yet somewhat similar self-duality relation. Through the exercise, it is hoped to demonstrate the correspondence between resolution capabilities and different self-duality relations.

The WD of the same Gabor logon is

$$W(t,\omega) = e^{-t^2/\sigma^2} e^{-\sigma^2\omega^2}. \tag{17}$$

Commonly, it is treated as an energy distribution function, and its corresponding "amplitude" distribution, which is to be compared with the STFT, is then

$$\sqrt{W}(t,\omega) = e^{-t^2/2\sigma^2} e^{-\sigma^2\omega^2/2}. \tag{18}$$

Notice that it satisfies a different self-dual relation:

$$e^{-t^2/2\sigma^2} e^{-\sigma^2\omega^2/2} \stackrel{\mathcal{F}}{\Longleftrightarrow} e^{-\tau^2/2\sigma^2} e^{-\sigma^2\theta^2/2} \tag{19}$$

or, in the short-hand notation, $(t,\omega) \longleftrightarrow (\tau,-\theta)$.

The contrast is clearly represented in Fig. 1. The different degrees of concentration of the "amplitude" distribution functions on the time-frequency plane can be easily observed. Different dual relations govern the two "amplitude" distribution functions: for STFT, it is $(t,\omega) \longleftrightarrow (-2\tau, 2\theta)$, i.e., the self-dual function is more concentrated on the $\theta-\tau$ plane; for $\sqrt{W}(t,\omega)$, it is $(t,\omega) \longleftrightarrow (-\tau,\theta)$, i.e., the self-dual function remains the same shape on the $\theta-\tau$ plane. Notice the observation that the STFT is more concentrated on the $\theta-\tau$ plane corresponds to its "excessive" spreads on the time-frequency plane, especially when compared to the "amplitude" function of the Wigner distribution.[2]

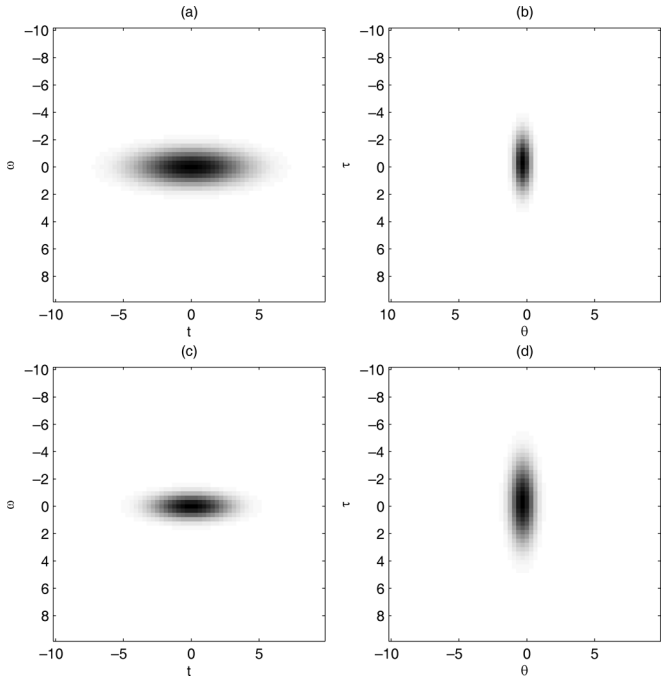

Fig. 1. Scaling relation between the time-frequency domain and the θ-τ domain. (a) shows an STFT and (b) is its two-dimensional inverse Fourier transform. (c) is the "amplitude" function of the Wigner distribution of the same signal, and (d) is its two-dimensional IFT.

On the alleged superior resolution of WD, there is an instructive equation that shows the construction of the Wigner distribution through auto-convolution of the STFT on the time-frequency plane: [see equation (21) at the bottom of the page], in which $\gamma = \int h^*(-t)h(t)\,dt$ and $G(t,\omega)$ is the STFT. This equation is a 2-dimensional extension of [13, eq. (3)] where the WD is related to the STFT via an auto-convolution in the frequency domain. There are two insights in (21) worth noting. The first is that, by the analogy to the claim that matching windows to localized signals results in high-resolution adaptive spectrograms [14], the auto-convolution in (21) also contributes to the superior resolution of WD. The second is that the oscillating term

[2]Please be noted that the self-duality relation presented here for the "amplitude" function of the WD is only for the single Gabor logon and cannot be extended to general signals. It should be viewed as merely an instructive example. For general signals, there is an interesting relation similar to what is found in [11], [12], regarding the "phased" WD, i.e., $e^{-j2t\omega}W(t,\omega)$. It goes like this:

$$\mathcal{IF}\{e^{-j2t\omega}W(t,\omega)\} = \frac{1}{2} e^{j\theta\tau/4} x\left(\frac{\tau}{2}\right) X^*\left(\frac{\theta}{2}\right) \tag{20}$$

in which $\mathcal{IF}$ denotes the two-dimensional inverse Fourier transform. Notice the "scaling" in the τ-domain signal and the θ-domain spectrum on the right-hand side. This equation conforms to the reasoning that expansion in the $\theta-\tau$ domain means tighter concentration of signal components and therefore a superior resolution in the time-frequency domain.

$$W(t,\omega) = \frac{1}{8\pi\gamma} \int\int G\left(t+\frac{\tau}{2},\omega+\frac{\theta}{2}\right) G^*\left(t-\frac{\tau}{2},\omega-\frac{\theta}{2}\right) e^{-j(\theta/2)t} e^{j(\tau/2)\omega}\, d\tau\, d\theta \tag{21}$$

$e^{-j(\theta/2)t}e^{j(\tau/2)\omega}$ gives indications on how the distinctive fast-oscillating "cross terms" arise in the midpoint of "auto terms."

The self-dual relation does not conflict with the well-known tradeoff of the resolution capability of Fourier analysis on the time-frequency plane. In the tradeoff, when the resolution capability along the frequency axis is enhanced through using a long window, the resolution along time axis is inevitably compromised in a similar way as shown by the Fourier transform pairs of Gabor logon $e^{-t^2/4\sigma^2} \stackrel{\mathcal{F}}{\Longleftrightarrow} e^{-\sigma^2\theta^2}$ and $e^{-\sigma^2\omega^2/4} \stackrel{\mathcal{F}}{\Longleftrightarrow} e^{-\tau^2/\sigma^2}$ which demonstrate the negotiation of spreads between two domains. This tradeoff can be clearly stated not only by the scaling duality but also the development of the uncertainty principle for time-frequency distributions [15]–[18] and may be figuratively understood through the short-hand notation $(t,\omega) \longleftrightarrow (1/\theta, 1/\tau)$. The scaling dual relation and the self-duality $(t,\omega) \longleftrightarrow (-2\tau, 2\theta)$ provide two different aspects to restrict the expression of STFT on the time-frequency plane.

The following sections present two applications of self-duality to general time-frequency analysis/synthesis. Notice that the main character of these applications is in recognizing the self-duality as an indication of the resolution limitation as well as a regulation on the expression of TFDs on the time-frequency plane.

III. Application I: Signal Synthesis From Given Time-Frequency Templates

The synthesis of STFT is to construct a valid STFT, that is, a 2-dimensional function which is mathematically related to a signal and a window function through (1), to match closely to a given time-frequency template. Some works have been done on related problems; for example, in [19], a finite-support discrete signal is reconstructed by using only the magnitude part of its STFT. TFDs other than STFT have also been studied in signal synthesis applications [20]–[22]. Here, the STFT synthesis problem is stated as follows.

1) The STFT Synthesis Problem: Given a power distribution template $K(n,k)$ on the time-frequency plane, find a legitimate STFT $G(n,k)$ such that the spectrogram $|G(n,k)|^2$ is closest to $K(n,k)$.

The self-dual (2) is a necessary condition for a function to be an STFT with symmetric window. Therefore, if there exists an STFT whose magnitude is close to $K(n,k)$. , there must exist self-dual functions whose magnitude is also close to the given template. The STFT synthesis problem may be solved by first find a self-dual function $C(n,k)$ such that its magnitude, i.e., $|C(n,k)|$, is close to $\sqrt{K(n,k)}$. The function $C(n,k)$ can be found with similar approaches adopted in the classical problem of two-dimensional signal retrieval using only the magnitude information of Fourier transform [23]–[29].

After the self-dual function is found, the next step is to construct a legitimate STFT $G(n,k)$ such that $\|G(n,k)-C(n,k)\|_2$ is minimized. Assume a given symmetric window function as a design parameter, a convex norm-minimization problem can be formulated and can be effectively solved through readily-available convex optimization software packages [30].

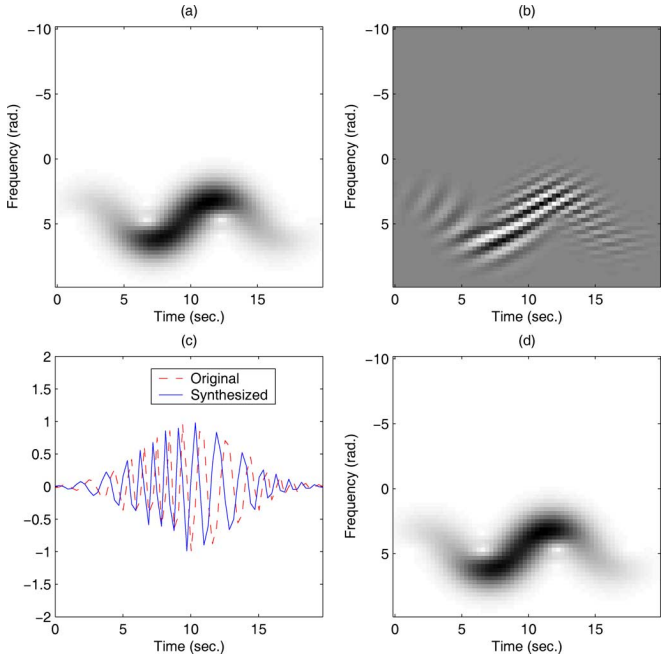

Fig. 2. STFT synthesizing algorithm is used to reconstruct a signal. (a) shows the time-frequency template function. (b) shows the real part of the self-dual function $C(n,k)$ generated by Step 1. The real part of the original signal (the dashed line) and the synthesized signal (the solid line) are shown in (c), and (d) shows the magnitude of the STFT generated by the reconstructed signal and the same window function.

The self-dual function $C(n,k)$ serves as an intermediate variable in the synthesis process. Notice that the phase-retrieval problem is nevertheless a non-trivial problem. In practice, most popular procedures are based on Iterative Fourier transform algorithm (ITF) [23], [26]. Here a variation of ITF is adopted as the core of the synthesis procedure. Start with an initial function $C(n,k) = \sqrt{K(n,k)}e^{\Phi(n,k)}$ where $\Phi(n,k)$ is a random phase function, we proceed to the following algorithm.

2) The STFT Synthesis Algorithm:

Step 1) Generate $\hat{C}(p,q) = \mathcal{IF}\{C(n,k)\}$, where $\mathcal{IF}$ denotes the 2-dimensional inverse Fourier transform. Set $|\hat{C}|(p,q) = 2\sqrt{K(-2q,2p)}$. Go to Step 2.

Step 2) Update $C(n,k)$ with $(1/2)C(n,k) + (1/2)\mathcal{F}\{\hat{C}(p,q)\}$ where $\mathcal{F}$ denotes the 2-dimensional Fourier transform. Set $C(n,k) = \sqrt{K(n,k)}$. Go back to Step 1 or go to Step 3 if the updating term is small.

Step 3) Given a window function $h(n)$ as a synthesis parameter, get $x(n)$ via solving a norm-minimization problem. The details are in Appendix I.

Two examples are provided to illustrate the STFT synthesis algorithm. The results are shown in Figs. 2 and 3. In the first example (shown in Fig. 2), the spectrogram of a frequency-modulated signal is used as the template. The STFT synthesis algorithm successfully reconstructs the signal up to a phase term, with only the magnitude of STFT is used. The next example demonstrates the role of the window function as a parameter in the synthesis process. Fig. 3(a) shows a template in the shape of a square with the length of 3.13 seconds and the width of 3.13 radians. Fig. 3(b) shows the real part of the self-dual function

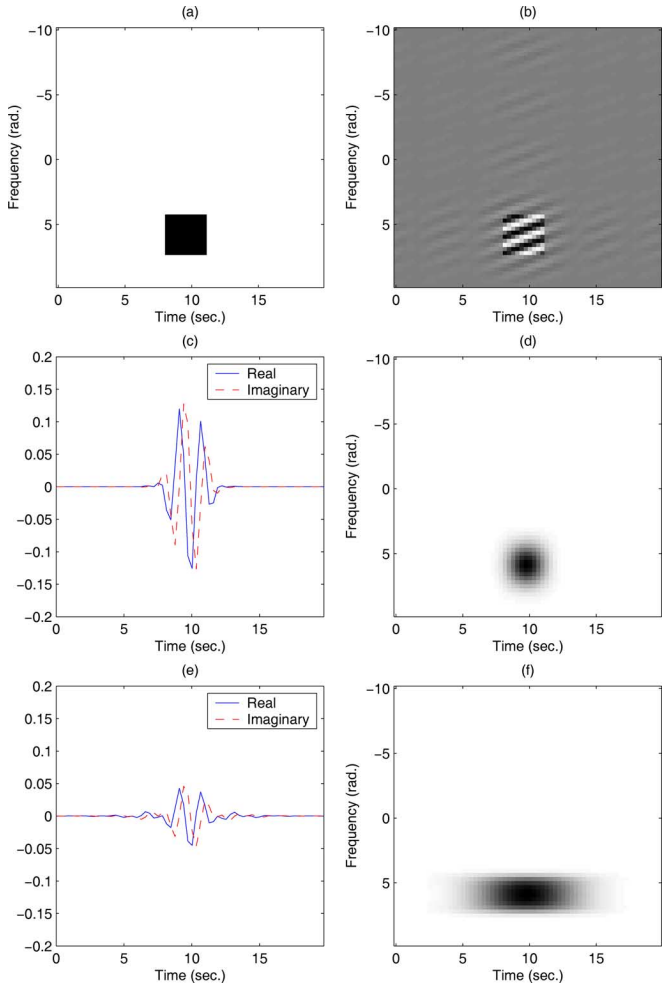

Fig. 3. (a) shows the time-frequency template function. (b) shows the real part of the self-dual function $C(n,k)$ generated by Step 1. (c) and (d) show the optimal signal and the spectrogram generated with a short window; (e) and (f) show the results with a longer window.

$C(n,k)$ found in Step 1. Fig. 3(c) and (e) shows, respectively, two norm-minimizing solutions obtained with a short and a long Gaussian window, and shown in Fig. 3(d) and (f) are the resulting spectrograms.

Since there are two duality relations (one is about self resemblance and the other scaling duality), i.e., $(t,\omega) \longleftrightarrow (1/\theta, 1/\tau)$ and $(t,\omega) \longleftrightarrow (-2\tau, 2\theta)$, that regulate the expression of STFT on the time-frequency plane, a given template can be judged whether is reasonable in the sense whether it is possible to exit a legitimate STFT having magnitude reasonably close to the given template. Indeed, since the self-duality is the necessary condition of an STFT with a symmetric window, Steps 1 and 2 can be used as a reality check for a given time-frequency template. If the magnitude of the generated self-dual function differs greatly from the given template, it means that the template function is ill chosen and there exists no signal that can possibly result in a spectrogram that is close to the given template.

Moreover, the scaling duality also regulates the concentration of signal energy, e.g., a signal cannot be concentrated along time axis without spreading on the frequency axis. With these two dual relations, sometimes it is straightforward to judge whether a given template can possibly be a suitable choice. For example,

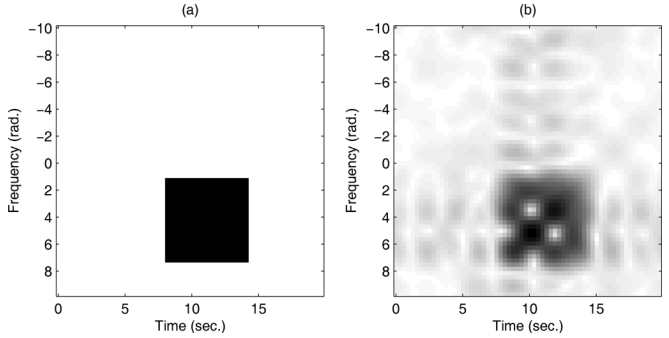

Fig. 4. Effect of ill-chosen templates. (a) shows a square template with the length of 6.26 s and the width of 6.26 rad, and the magnitude of the self-dual function generated in Step 1 is shown in (b).

consider a square plateau with the length and width of T. By self-duality, its two-dimensional IFT would have a magnitude function in the shape of a square plateau with the length and width of $T/2$. At the same time, by the scaling duality, the length and width of the significant part of the IFT should be around $2\pi/T$. Hence, for choosing a reasonable template, one should set $T/2 \approx 2\pi/T$, i.e., T should be around $2\sqrt{\pi}$; otherwise, the square plateau is not proper as a magnitude template for STFTs. Fig. 4 illustrates the effect of an ill-chosen template that has a plateau too large for a reasonable template. Fig. 4(a) shows such a template function, and the magnitude of the self-dual function generated in Step 1 is shown in Fig. 4(b). Since the template has a plateau much larger than what can be generated by a legitimate STFT, there is no self-dual function that can come close to the template and inevitably "hole" and "cracks" will develop in the magnitude of the self-dual function.

IV. Application II: High-Resolution Time-Frequency Distributions

The contrast between the self-dual relations regarding the spectrogram and the WD [(14) and (19)] and the fact that the WD holds a remarkable resolution capability despite the annoying cross terms [31] leads to an interesting question: Could the self-duality $(t,\omega) \longleftrightarrow (\tau,-\theta)$ be worked into the formulation of designing high-resolution TFDs with suppressed cross terms whose super-fine structure does not seem to conform to self-duality? In the following, such a plan is devised to find a 2-dimensional function, call it $H(t,\omega)$, for a given signal such that $H(t,\omega)$ satisfies the high-resolution version of self-duality $(t,\omega) \longleftrightarrow (\tau,-\theta)$ and also optimizes a certain criterion regarding the relevance between the signal and $H(t,\omega)$, and $H(t,\omega)$ will be shown to be worthwhile considering as a high-resolution TFD.

First the problem of formulating a proper criterion of relevance is addressed. The Wigner distribution, despite its problem of cross terms, has been widely regarded as a plausible time-frequency representation with high-resolution capability for time-varying signals. Therefore, a straightforward choice is to minimize the L_2 norm between $|H(t,\omega)|^2$ and $W(t,\omega)$, i.e., to minimize the term

$$\int_{t=-\infty}^{\infty}\int_{\omega=-\infty}^{\infty}[|H(t,\omega)|^2 - W(t,\omega)]^2\,dt\,d\omega. \qquad (22)$$

Intuitively, the approach can be understood as to create the function $H(t,\omega)$ by harnessing two forces which compensate each other. The minimization of the L_2 norm makes the resulting $H(t,\omega)$ assemble the Wigner distribution and hopefully keep the high resolution capability; while self duality eliminates spurious cross terms whose ultra-localized features on the time-frequency plane cannot meet the requirement. Later an example will be provided to illustrate this point. Take the formulation with continuous signals as a model, we now proceed to define the first version of the algorithm to construct high-resolution TFD for discrete signals.

1) The High-Resolution TFD Problem (I): Find a 2-D function $H(n,k)$ which satisfies the high-resolution version of self-duality, i.e., $(n,k) \longleftrightarrow (q,-p)$, and whose squared magnitude $|H(p,q)|^2$ should be as close to the Wigner distribution as possible.

A procedure based on the ITF algorithm can also be used to solve this problem. First, start with a discrete function $H(n,k) = \sqrt{\max(W(n,k),0)}e^{j\phi(n,k)}$ where $\phi(n,k)$ is an arbitrary random phase function. Then it goes to the following iterative procedure.

2) The High-Resolution TFD Algorithm (I):

Step 1) Generate $\hat{H}(p,q) = \mathcal{IF}\{H(n,k)\}$. Set $|\hat{H}(p,q)| = \sqrt{\max(W(q,-p),0)}$. Go to Step 2.

Step 2) Update $H(n,k)$ with $(1/2)H(n,k) + (1/2)\mathcal{F}\{\hat{H}(p,q)\}$. Set $|H(n,k)| = \sqrt{\max(W(n,k),0)}$. Go to Step 1 or exit if the updating term is small.

More often than not the TFD design problem involves time and frequency marginal functions, since they represent the information available on the time and frequency axes to which most attention is often given. These kinds of considerations usually are formulated as equality constraints in the TFD designing posed as a convex optimization problem [32].

Generalized marginal functions derived from the fractional Fourier transform [33]–[35] have also been proposed for the TFD design problem, for instance, in [36]. Here a novel approach towards considering multiple marginal functions is proposed, and it is called the collective marginal error. In short, the goal is to find a TFD that does not aim at satisfying a specific set of marginal functions, but to minimize the collective error of all generalized marginal functions. This criterion is particularly suitable for exploratory signal analysis where no preference for any marginal function is set *a priori*.

In [37], the authors proposed to use the Radon transform [38], to facilitate the kernel design [39] procedure for TFDs to achieve high resolution with suppressed cross terms. The Radon transform is performed on the modulus of the ambiguity function [6], [7] for observing of how the auto terms are distributed in the ambiguity domain (the $\theta - \tau$ domain) and use that information to design an adaptive kernel function. In this paper, the Radon transform will be used on the time-frequency domain to develop the idea of collective marginal error. In addition, the Radon transform only serves as a conceptual tool and is not calculated in the final algorithm.

Let $\mathcal{R}_\alpha$ denote the Radon transform on the axis obtained by counterclockwise rotating the time axis with the angle α and $M_\alpha(\xi)$ be the signal's generalized marginal on the rotated axis. The collective marginal error of a power distribution $|H(t,\omega)|^2$ is defined as

$$\mathcal{E} \equiv \int_{\alpha=0}^{\pi}\int_{\xi=-\infty}^{\infty}[(\mathcal{R}_\alpha|H|^2)(\xi) - M_\alpha(\xi)]^2\,d\xi\,d\alpha. \tag{23}$$

Let $D(t,\omega) = |H(t,\omega)|^2 - W(t,\omega)$, i.e., the difference between the target power distribution and the Wigner distribution, then the collective marginal error has an intuitive expression in the ambiguity domain as follows:

$$\mathcal{E} = \int_{\theta=-\infty}^{\infty}\int_{\tau=-\infty}^{\infty}\frac{|\hat{D}(\theta,\tau)|^2}{\sqrt{\theta^2+\tau^2}}\,d\theta\,d\tau \tag{24}$$

where $\hat{D}(\theta,\tau)$ is the 2-dimensional Fourier transform of $D(t,\omega)$. The mathematical derivation is presented in Appendix II.

Notice that the collective marginal error emphasizes the low-pass part on the ambiguity, i.e., θ-τ, domain. Usually it is the case that the so-called cross terms occupy regions away from the origin of the ambiguity domain. Accordingly, the collective marginal error criterion is expected to be compatible with designing TFDs aimed at suppressing cross terms. Furthermore, notice that $\mathcal{E}$ is also a measure of how close is a TFD to the WD. Therefore, the resulting TFD is also expected to retain the resolution capability of the WD.

3) The High-Resolution TFD Problem (II): Find a 2-D function $H(n,k)$ which satisfies the self-duality $(n,k) \longleftrightarrow (q,-p)$ and whose squared magnitude $|H(p,q)|^2$ minimizes the discrete collective marginal error:

$$\mathcal{E} = \sum_{\substack{q=-N/2+1\\(p,q)\neq(0,0)}}^{N/2}\sum_{p=-N/2+1}^{N/2}\frac{|\hat{D}(p,q)|^2}{\sqrt{p^2+q^2}}. \tag{25}$$

Notice that $\hat{D}(p,q)|_{p=q=0}$ is made to be zero by forcing the total energy of $|H(n,k)|^2$ equal to that of $W(n,k)$.

The minimization of $\mathcal{E}$ only involves the magnitude of $H(n,k)$. Indeed $\mathcal{E}$ is a convex function of $|H(n,k)|^2$, and it is straightforward to find the $\mathcal{E}$-minimizing $|H(n,k)|^2$. Once given a magnitude template, via ITF, a self-dual function whose magnitude close to the template can be found. The overall procedure contains iterations between two operations. The operations result in the following heuristic algorithm.

4) The High-Resolution TFD Algorithm (II):

Step 1) Find the $\mathcal{E}$-minimizing mask function $|H_0(p,q)|^2$ via the steepest-gradient method.

Step 2) Generate $\hat{H}(p,q) = \mathcal{IF}\{H(n,k)\}$. Set $|\hat{H}(p,q)| = \min(|H_0(-q,p)|, |\hat{H}(p,q)|)$. Go to Step 3.

Step 3) Update $H(n,k)$ with $(1/2)H(n,k) + (1/2)\mathcal{F}\{\hat{H}(p,q)\}$. Set $|H(p,q)| = \min(|H_0(p,q)|, |H(p,q)|)$ and normalize the energy. Go back to Step 2 or exit if the updating term is small.

Simulation results show that this procedure performs better than the first algorithm and it does generate superb results

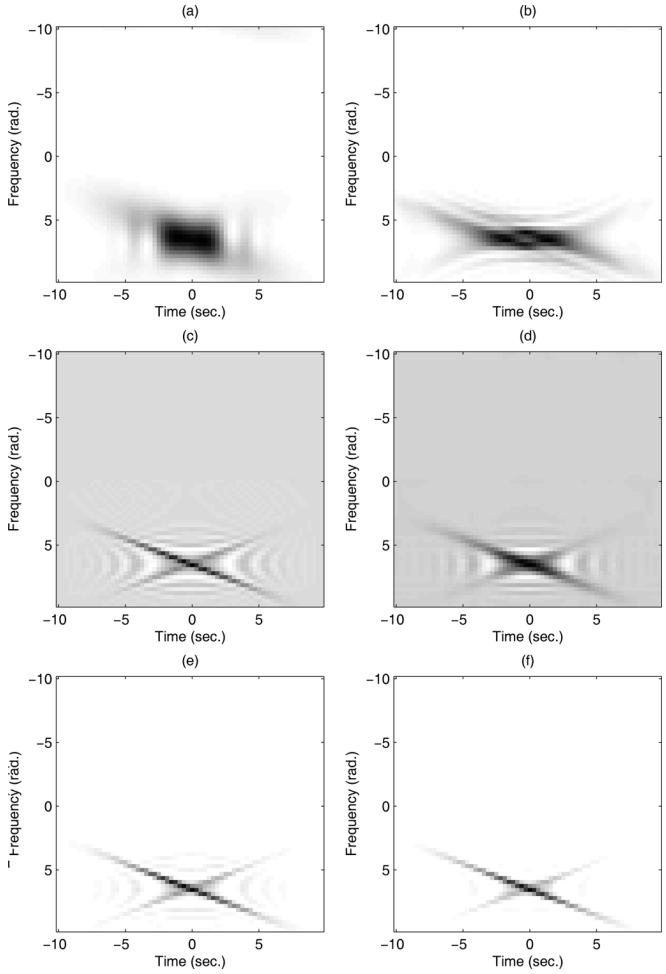

Fig. 5. Several time-frequency distributions of a pair of crossing chirped Gabor logons are shown for comparison. (a) and (b) are spectrograms generated with a short and a long window respectively. The Wigner distribution is shown in (c) and (d) shows the RID. Finally, (e) and (f) show the results generated with the high-resolution TFD algorithm I and II, respectively.

for signals that are typically difficult to be handled by other well-known time-frequency analysis tools. Fig. 5 illustrates an example of two chirped Gabor logons crossing each other. Fig. 5(a) and (b) shows the spectrograms generated with a relatively short and a long Gaussian window respectively. Either spectrogram fails to represent these chirped Gabor logons in a visually appealing way for an intuitive interpretation. The WD is shown in Fig. 5(c), and its signature features (the so-called cross terms) are so obtrusive that they obscure the superb resolution of the WD in representing the crossing chirp signals. The Choi–Williams distribution [39], a member of the reduced interference distribution (RID) class, is shown in Fig. 5(d) for comparison. While generating the RID, careful adjustment of the kernel function is made via minimizing the Rényi information measure [40]–[42] in an attempt to achieve a visual balance between showing details of the signal and reducing interference terms. However, for chirped signals, the end result is not very satisfying. Fig. 5(e) and (f) shows the results of high-resolution TFD algorithm I and II respectively. The TFDs approach the high resolution typically associated

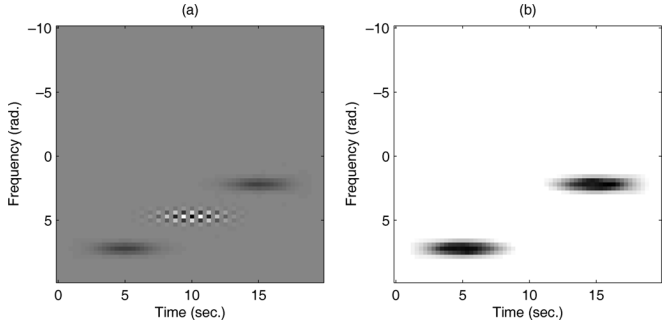

Fig. 6. WD is shown in (a) while the high-resolution TFD in (b). Notice that the spurious cross terms are substantially reduced in (b).

with the WD and substantially suppressing the annoying cross terms, especially in the case of Fig. 5(f).

Finally, an example of two well-separated Gabor logons is provided to show that fast-oscillating "cross terms" can be substantially reduced by the ITF which enforces self-duality. Notice that applying collective marginal error does reduce the cross terms; yet a heavy suppression is only achieved after the ITF.

V. Conclusion

The sufficient conditions for the approximate self-duality of discrete STFT to hold are presented. Self-duality is interpreted in the context of resolution capabilities of time-frequency distributions. Furthermore, two signal processing applications , the first be the STFT synthesis and the other the construction of high-resolution TFD, are provided to demonstrate the power and beauty of the concept.

Appendix I
Norm Minimization in STFT Synthesis

Let $C(n,k)$ be the self-dual function found in Step 1 of the STFT synthesis algorithm. We would like to find the discrete signal $x(n)$ such that its STFT $G(n,k)$ with the given window function $h(n)$ minimized the norm

$$\sum_{n=-N/2}^{N/2-1}\sum_{k=-N/2}^{N/2-1}|G(n,k)-C(n,k)|^2. \tag{26}$$

The definition of STFT in (1) can be viewed as an inner product:

$$G(n,k)=\langle \underline{\mathbf{w}}_{n,k},\underline{\mathbf{x}}\rangle=\left[w_{n,k}\left(-\frac{N}{2}\right),\ldots,w_{n,k}\left(\frac{N}{2}-1\right)\right]\cdot\left[x\left(-\frac{N}{2}\right),\ldots,x\left(\frac{N}{2}-1\right)\right]^T \tag{27}$$

where

$$w_{n,k}(m)=\frac{1}{N}h^*(n-m)e^{j\pi k(n-2m)/N} \tag{28}$$

representing the combination of windowing operation and Fourier transform. The norm-minimizing $x(n)$ thus comes from the complex-valued least-squares solution of the equation

$$\mathbf{W}\underline{\mathbf{x}}=\underline{\mathbf{C}} \tag{29}$$

in which $\mathbf{W}$ is an $N^2\times N$ matrix stacked up with the row vectors $(\underline{\mathbf{w}}_{n,k})^T$, and $\underline{\mathbf{C}}$ is an $N^2\times 1$ vector stacked up with corresponding $C(n,k)$.

$$\mathcal{F}^{-1}_{\eta\to\xi}[\hat{D}(\eta\cos\alpha,\eta\sin\alpha)\otimes_\eta \hat{D}(\eta\cos\alpha,\eta\sin\alpha)] = \int_{\eta=-\infty}^{\infty}\int_{\mu=-\infty}^{\infty}\hat{D}(\mu\cos\alpha,\mu\sin\alpha)\hat{D}((\eta-\mu)\cos\alpha,(\eta-\mu)\sin\alpha)e^{j\xi\eta}\,d\mu\,d\eta, \tag{31}$$

$$\begin{aligned}
&\int_{\alpha=-\infty}^{\infty}\int_{\xi=0}^{\pi}\int_{\eta=-\infty}^{\infty}\int_{\mu=-\infty}^{\infty}\hat{D}(\mu\cos\alpha,\mu\sin\alpha)\hat{D}((\eta-\mu)\cos\alpha,(\eta-\mu)\sin\alpha)e^{j\xi\eta}\,d\mu\,d\eta\,d\xi\,d\alpha\\
&\quad=\int_{\alpha=0}^{\pi}\int_{\eta=-\infty}^{\infty}\int_{\mu=-\infty}^{\infty}\hat{D}(\mu\cos\alpha,\mu\sin\alpha)\hat{D}((\eta-\mu)\cos\alpha,(\eta-\mu)\sin\alpha)\delta(\eta)\,d\mu\,d\eta\,d\alpha\\
&\quad=\int_{\alpha=0}^{\pi}\int_{\mu=-\infty}^{\infty}\hat{D}(\mu\cos\alpha,\mu\sin\alpha)\hat{D}(-\mu\cos\alpha,-\mu\sin\alpha)\,d\mu\,d\alpha\\
&\quad=\int_{\alpha=0}^{\pi}\int_{\mu=-\infty}^{\infty}|\hat{D}(\mu\cos\alpha,\mu\sin\alpha)|^2\,d\mu\,d\alpha\\
&\quad=\int_{\theta=-\infty}^{\infty}\int_{\tau=-\infty}^{\infty}\frac{|\hat{D}(\theta,\tau)|^2}{\sqrt{\theta^2+\tau^2}}\,d\theta\,d\tau.
\end{aligned} \tag{32}$$

Notice that not all row vectors $(\underline{\mathbf{w}}_{n,k})^T$ of any n and k need to be selected for the equation system. A system with much reduced size can be constructed with only (n,k) corresponding to region of interests. According to the sampling theory of discrete STFT, as few as N equations may be sufficient to determine a legitimate STFT; as the number of selected equations goes small, however, solving the norm-minimization problem encounters the problem of numerical accuracy. The best selection of (n,k) and efficient numerical method to solve for optimal $x(n)$ is beyond the scope of this paper.

Appendix II
The Collective Marginal Error

First consider the case of continuous signals. The term $[(\mathcal{R}_\alpha D(\xi)]^2$ in the collective marginal error can be written as

$$\mathcal{F}^{-1}_{\eta\to\xi}\{[\mathcal{F}(\mathcal{R}_\alpha)](\eta)\otimes_\eta[\mathcal{F}(\mathcal{R}_\alpha)](\eta)\} \tag{30}$$

where $\mathcal{F}$ denotes the Fourier transform and $\otimes_\eta$ the convolution in the variable η. By invoking the projection theorem [38], the expression becomes [see (31), shown at the top of the page], and the collective marginal error $\mathcal{E}$ becomes [see (32), shown at the top of the page]. Notice that the final expression emphasizes the low-pass part of the error through the weighting factor $1/\sqrt{\theta^2+\tau^2}$; the suppression of high-pass components is a common feature of the so-called reduced-interference distribution [39].

As for discrete signals, the weighting term $1/\sqrt{p^2+q^2}$ with p and q being not simultaneously zero is used in evaluating the collective marginal error in double summation, and $\hat{D}(p,q)|_{p=q=0}$ is made to be zero by forcing the total energy of $|H(n,k)|^2$ equal to that of $W(n,k)$.

References

[1] C. A. Stutt, "Some results on real-part/imaginary part and ' magnitude-phase relations in ambiguity functions," *IEEE Trans. Inf. Theory*, vol. 10, no. 4, pp. 321–327, 1964.

[2] G. B. Folland, *Harmonic Analysis in Phase Space*. Princeton, NJ: Princeton Univ. Press, 1989.

[3] K. Gröchenig, *Foundations of Time-Frequency Analysis*. Boston, MA: Birkhäuser, 2000.

[4] E. Chassande-Mottin, I. Daubechies, F. Auger, and P. Flandrin, "Differential reassignment," *IEEE Signal Process. Lett.*, vol. 4, pp. 293–294, 1997.

[5] H. M. Ozaktas and U. Sümbül, "Interpolating between periodicity and discreteness through the fractional Fourier transform," *IEEE Trans. Signal Process.*, vol. 54, no. 11, pp. 4233–4243, Nov. 2006.

[6] J. Ville, "Theorie et applications de la notion de signal analytique," *Cables et Transmission*, vol. 2A, pp. 61–74, 1948.

[7] J. E. Moyal, "Quantum mechanics as a statistical theory," in *Proc. Camb. Phil. Soc.*, 1949, vol. 45, pp. 99–124.

[8] T. A. C. M. Claasen and W. F. G. Mecklenbräuker, "The Wigner distribution—A tool for time-frequency signal analysis, Part I: Continuous-time signals," *Philips J. Res.*, vol. 35, no. 3, pp. 217–250, 1980.

[9] T. A. C. M. Claasen and W. F. G. Mecklenbräuker, "The Wigner distribution—A tool for time-frequency signal analysis, Part II: Discrete-time signals," *Philips J. Res.*, vol. 35, no. 4/5, pp. 276–300, 1980.

[10] T. A. C. M. Claasen and W. F. G. Mecklenbräuker, "The Wigner distribution—A tool for time-frequency signal analysis, Part III: Relations with other time-frequency signal transformations," *Philips J. Res.*, vol. 35, no. 6, pp. 372–389, 1980.

[11] J. G. Kirkwood, "Quantum statistics of almost classical ensembles," *Phys. Rev.*, vol. 44, pp. 31–37, 1933.

[12] P. J. Loughlin, J. W. Pitton, and B. Hannaford, "Fast approximations to positive time-frequency distributions, with applications," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, 1995, vol. II, pp. 1009–1012.

[13] L. Stankovic, "A method for time-frequency analysis," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 42, no. 1, pp. 225–229, 1994.

[14] G. Jones and B. Boashash, "Generalized instantaneous parameters and window matching in the time-frequency plane," *IEEE Trans. Signal Process.*, vol. 45, no. 5, pp. 1264–1275, May 1997.

[15] W. Heisenberg, *Physics and Philosophy; The Revolution in Modern Science*. New York: Harper & Row, 1958.

[16] P. Flandrin, *Time-Frequency/Time-Scale Analysis*. San Diego, CA: Academic, 1999.

[17] P. Korn, "Some uncertainty principles of time-frequency transforms of the Cohen class," *IEEE Trans. Signal Process.*, vol. 53, no. 2, pp. 523–527, Feb. 2005.

[18] P. J. Loughlin and L. Cohen, "The uncertainty principle: Global, local, or both?," *IEEE Trans. Signal Process.*, vol. 52, no. 5, pp. 1218–1227, May 2004.

[19] S. H. Nawab, T. F. Quatieri, and J. S. Lim, "Signal reconstruction from short-time Fourier transform magnitude," *IEEE Trans. Signal Process.*, vol. 31, no. 4, pp. 986–998, Aug. 1983.

[20] A. Francos and M. Porat, "Analysis and synthesis of multicomponent signals using positive time-frequency distributions," *IEEE Trans. Signal Process.*, vol. 47, no. 2, pp. 493–504, Feb. 1999.

[21] S. Peleg and B. Friedlander, "Multicomponent signal analysis using the polynomial-phase transform," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 32, no. 1, pp. 378–387, 1999.
[22] L. Le and S. Krishnan, "Time-frequency signal synthesis and its application in multimedia watermark detection," *EURASIP J. Appl. Signal Process.*, vol. 2006, pp. 1–14, 2006.
[23] R. Gerchberg and W. O. Saxton, "A practical algorithm for the determination of phase from image and diffraction plane pictures," *Optik*, vol. 35, pp. 237–246, 1972.
[24] P. J. Napier and R. H. T. Bates, "Inferring phase information from modulus information in two-dimensional aperture synthesis," *Astron. Astrophys. Suppl.*, vol. 15, pp. 427–430, 1974.
[25] M. H. Hayes, J. S. Lim, and A. V. Oppenheim, "Signal reconstruction from phase or magnitude," *IEEE Trans. Acoust., Speech, and Signal Process.*, vol. ASSP-28, pp. 672–680, Dec. 1980.
[26] J. R. Fienup, "Phase retrieval algorithms: A comparison," *Appl. Opt.*, vol. 21, pp. 2758–2769, 1982.
[27] M. H. Hayes, "The reconstruction of a multidimensional sequence from the phase of magnitude of its Fourier transform," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. ASSP-30, pp. 140–154, Apr. 1982.
[28] Z. Mou-yan and R. Unbehauen, "Methods of reconstruction of 2-d sequences from Fourier transform magnitude," *IEEE Trans. Image Process.*, vol. 6, no. 2, pp. 222–233, Feb. 1997.
[29] A. E. Yagle and A. E. Bell, "One- and two-dimensional minimum and nonminimum phase retrieval by solving linear systems of equations," *IEEE Trans. Signal Process.*, vol. 47, no. 11, pp. 2978–1989, Nov. 1999.
[30] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.
[31] L. Cohen, *Time-Frequency Analysis*. Englewood Cliffs, NJ: Prentice-Hall, 1995.
[32] R. M. Nickel, T. H. Sang, and W. J. Williams, "New signal adaptive approach to positive time-frequency distributions with suppressed interference terms," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, 1998, vol. 3, pp. 1777–1780.
[33] V. Namias, "The fractional order Fourier transform and its application to quantum mechanism," *J. Inst. Math. Appl.*, vol. 25, pp. 241–265, 1980.
[34] L. B. Almeida, "The fractional Fourier transform and time-frequency representations," *IEEE Trans. Signal Process.*, vol. 42, no. 11, pp. 3084–3091, Nov. 1994.
[35] A. I. Zayed, "On the relationship between the Fourier and fractal Fourier transforms," *IEEE Signal Process. Lett.*, vol. 3, pp. 310–311, Dec. 1996.
[36] J. R. Fonollosa, "Positive time-frequency distributions based on joint marginal constraints," *IEEE Trans. Signal Process.*, vol. 44, no. 8, pp. 2086–2091, Aug. 1996.
[37] B. Ristic and B. Boashash, "Kernel design for time-frequency signal analysis using the radon transform," *IEEE Trans. Signal Process.*, vol. 41, no. 5, pp. 1996–2008, May 1993.
[38] J. S. Lim, *Two-Dimensional Signal and Image Processing*. Englewood Cliffs, NJ: Prentice-Hall, 1990.
[39] J. Jeong and W. J. Williams, "Kernel design for reduced interference distributions," *IEEE Trans. Signal Process.*, vol. 40, no. 2, pp. 402–412, Feb. 1992.
[40] A. Rényi, "On measures of entropy and information," in *Proc. 4th Berkley Symp. Math., Stat., Probab.*, 1961, vol. 1, pp. 547–561.
[41] T. Sang and W. J. Williams, "Rényi information and signal-dependent kernel design," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, 1995, vol. 2, pp. 509–512.
[42] S. Aviyente and W. J. Williams, "Minimum entropy time-frequency distributions," *IEEE Signal Process. Lett.*, vol. 12, no. 1, pp. 37–40, Jan. 2005.

Tzu-Hsien Sang (S'95–M'99) received the B.S.E.E. degree from the National Taiwan University in 1990 and the Ph.D. degree from the University of Michigan at Ann Arbor in 1999.

He previously worked on physical layer design for broadband technologies at Excess Bandwidth, Sunnyvale, CA, a start-up company. He is currently with the Department of Electronics Engineering, National Chiao-Tung University (NCTU), Taiwan. His research interests include signal processing for communications, time-frequency analysis for biomedicl signals, and RF circuit noise modeling.