

Sub-Nyquist Sampling of Short Pulses: Part I

Ewa Matusiak, Yonina C. Eldar, *Senior Member, IEEE*

Abstract

We develop sub-Nyquist sampling systems for analog signals comprised of several, possibly overlapping, finite duration pulses with unknown shapes and time positions. Efficient sampling schemes when either the pulse shape or the locations of the pulses are known have been previously developed. To the best of our knowledge, stable and low-rate sampling strategies for a superposition of unknown pulses without knowledge of the pulse locations have not been derived. The goal in this two-part paper is to fill this gap. We propose a multichannel scheme based on Gabor frames that exploits the sparsity of signals in time and enables sampling multipulse signals at sub-Nyquist rates. Our approach is based on modulating the input signal in each channel with a properly chosen waveform, followed by an integrator. We show that, with proper preprocessing, the Gabor coefficients necessary for almost perfect reconstruction of the input signal, can be recovered from the samples using standard methods of compressed sensing. In addition, we provide error estimates on the reconstruction and analyze the proposed architecture in the presence of noise. The resulting scheme is flexible and exhibits good noise robustness. The first part in this series is focused on the basic underlying principles. The second part generalizes the present sampling system in several directions. In particular, we consider practical implementations from a hardware perspective and extend the architecture to efficiently sample radar-like signals that are sparse in both time and frequency.

I. INTRODUCTION

One of the common assumptions in sampling theory suggests that in order to perfectly reconstruct a bandlimited analog signal from its samples, it must be sampled at the Nyquist rate, that is twice its highest frequency. In practice, however, all real life signals are necessarily of finite duration, and consequently cannot be perfectly bandlimited, due to the uncertainty principle [1]. The Nyquist rate is therefore dictated by the essential bandwidth of the signal, that is by the desired accuracy of the approximation: the higher the rate, meaning the more samples are taken, the better the reconstruction.

In this paper we are interested in sampling time limited signals. There are two standard approaches in the literature to sample such functions. One is to acquire pointwise samples and approximate the signal using Shannon's interpolation formula [2], [3]. The second, is to collect Fourier samples and approximate the signal using a truncated Fourier series. Both strategies require the Fourier transform of the signal to be integrable. Moreover, exact pointwise samples needed for Shannon's method requires implementing a very high bandwidth sampling filter. Here we show that these problems can be alleviated using Gabor frames [4].

The authors are with the department of electrical engineering, Technion–Israel Institute of Technology, Haifa, Israel (phone: +972-4-8294682, fax: +972-4-8295757, e-mail: ewa.matusiak@univie.ac.at, yonina@ee.technion.ac.il)

Gabor samples, which are inner products of a function with shifted and modulated versions of a chosen window, are a good compromise between exact pointwise samples and Fourier samples. In particular, we show that all square-integrable time limited signals, without additional conditions on their Fourier transforms, can be well approximated by truncated Gabor series. The price to pay is a slightly greater number of samples necessary for approximation, that comes with using frames, namely, overcomplete dictionaries. The use of frames is a result of the fact that Gabor bases are not well localized in both time and frequency [5]. In all three approaches (pointwise, Fourier, Gabor) the number of samples necessary to represent an arbitrary time limited signal is dictated by the essential bandwidth of the signal and the desired approximation accuracy.

Recently, there has been growing interest in efficient sampling of a special class of time limited functions: signals consisting of a stream of short pulses, referred to as multipulse signals [6], [7], [8]. This interest is motivated by a variety of different applications such as digital processing of certain radar signals, which are superpositions of shifted and modulated versions of a single pulse. Another example is ultrasound signals, that can be modeled by superpositions of shifted versions of a given pulse shape. Multipulse signals are also prevalent in communication channels, bio-imaging, and digital processing of neuronal signals. Since the pulses occupy only a small portion of the signal support, intuitively less samples should suffice to reconstruct the signal. Our main goal in this paper is to design a minimal rate sampling and reconstruction scheme for multipulse signals that exploits the inherent structure of these signals, without knowing the pulse shapes and their locations.

A special case of this model was considered in [7] and [6] in which the signal is composed of shifts of a single known pulse shape. Such signals are completely characterized by the time delays and amplitudes of the pulse, and are therefore defined by a finite number of parameters. This model falls under the class of finite rate of innovation (FRI) signals introduced in [8]. The sampling scheme proposed in [6] operates at the minimal sampling rate required for such signals, determined by the rate of innovation [8], which equals the number of unknown time delays and amplitudes. In this case without noise, perfect recovery is possible due to the finite dimensionality of the problem.

A more general class of multipulse signals results when the pulse shapes are not known. In this scenario perfect reconstruction from a finite number of samples is impossible, as there are generally infinitely many parameters defining the signal. Nonetheless, it is clear intuitively that the reconstruction error can be made sufficiently small with just a finite number of samples, when the signal is sampled dense enough. However, this strategy obviously results in many pointwise samples that are zero, leading to unnecessary high rates. Since the Fourier transform does not account for local properties of the signal, this method cannot be used to reduce the sampling rate and exploit the signal structure.

In this paper we consider sampling of multipulse signals when neither the pulses nor their locations are known. The pulses can have arbitrary shapes and positions, and may overlap. The only knowledge we assume is that our signal is comprised of N pulses, each of maximal width W . Despite the complete lack of knowledge on the signal shape, we show that using Gabor frames and appropriate processing, such signals can be sampled in an efficient and robust way, using far fewer samples than that dictated by the Nyquist rate. The number of samples is proportional to WN , that is, the actual time occupancy. More precisely, we need about $4\mu^{-1}\Omega'WN$ samples, where Ω' is related

to the essential bandwidth of the signal and $\mu \in (0, 1)$ is the redundancy of the Gabor frame used for processing. In contrast, Nyquist-rate sampling in this setting requires about $\Omega'\beta$ samples, where β is the duration of the signal. If the signal occupies only a small portion of its time duration, such that $4\mu^{-1}WN \ll \beta$, then our scheme can result in a substantial gain over Nyquist-rate sampling.

The criteria we consider for sampling multipulse signals are: a) minimal sampling rate that allows almost perfect reconstruction, b) no prior knowledge on the locations of the pulses or their shapes, and c) numerical stability in the presence of mismodeling and noise. To achieve these goals we combine the well established theory of Gabor frames [4] with the recently proposed Xampling paradigm [9], [10] which is a framework for sub-Nyquist sampling of analog signals. Gabor samples, taken with respect to a window that is well localized in time and frequency, provide information about local behavior of any square integrable function and reflect the sparsity of a function either in time or frequency. The scheme we propose consists of a multichannel system that modulates the input signal in each channel with a parametric waveform, based on a chosen Gabor frame, and integrates the result over a finite time interval. We show that by a proper selection of the waveform parameters, the Gabor samples can be recovered, from which the signal is reconstructed. We further prove that the proposed system is robust to noise and model errors, in contrast with techniques based on exact pointwise samples.

Our development follows the philosophy in much of the recent work in analog compressed sensing, termed Xampling, which provides a framework for incorporating and exploiting structure in analog signals to reduce the sampling rates, without the need for discretization [9], [10]. A pioneer sub-Nyquist system of this type is the modulated wideband converter (MWC) introduced in [11] based on the earlier work of [12]. This scheme targets low rate sampling of multiband signals, namely bandlimited signals whose frequency content is concentrated on a few bands. The MWC enables perfect recovery of any multiband function from its samples at rates far below Nyquist, without knowledge of the band locations. Sub-Nyquist sampling is achieved by applying modulation waveforms to the analog input prior to uniformly sampling at the low rate. A hardware prototype of the MWC is reported in [10]; this is a first example of a wideband prototype that implements compressed sensing in the analog domain. Our sampling system is Fourier dual to the MWC in the sense that we treat signals concentrated on a few intervals in time.

Another system that falls into the Xampling paradigm is that of [6] which treats multipulse signals with a known pulse shape. The proposed sampling scheme is based on modulation waveforms as in the MWC, but the purpose of the waveforms is different. In the MWC the modulations are used to reduce the sampling rate relative to the Nyquist rate, while in the setting of [6] the modulations serve to simplify the hardware implementation and improve robustness. In the second paper of this series, we use similar modulation waveforms to reduce the sampling rate for multipulse signals that are also frequency sparse.

We note here, that Gabor frames were recently used to sample short discrete pulses in [13]. The authors analyzed standard compressed sensing techniques for redundant dictionaries, and applied their results to radar-like signals. The important difference between the scheme in [13] and our problem is that the former can handle only discrete time signals, which are already sampled. In contrast, our method directly reduces the sampling rate without the

need for discretization.

In the first of this two-part series we develop the basic sampling scheme, and investigate its performance in the presence of noise. Since time limited functions cannot be perfectly reconstructed from a finite number of samples, we provide error bounds on the approximation depending on the chosen Gabor frame. In the second part, we generalize our results to allow for more efficient hardware implementation. We then demonstrate how this extension can be used to efficiently sample radar signals. Such signals are an important subclass of the multipulse model in which all pulses share the same (unknown) structure. The resulting signal is a finite superposition of one shifted and modulated pulse. By designing proper modulating waveforms in our sampling system we exploit the resulting sparsity in time and frequency to further reduce the number of samples and reconstruct the signal almost perfectly.

The current paper is organized as follows. In Section II we introduce the notation and basic problem definition. Section III describes existing methods of sampling multipulse signals and their drawbacks. Since the main tool in our analysis is Gabor frames, in Section IV we recall basic facts and definitions from Gabor theory and show that truncated Gabor series provide a good approximation for time limited functions. Based on this observation, in Section V, we introduce a sub-Nyquist sampling scheme for multipulse signals. We analyze its performance in the presence of noise in Section VI. Section VII points out connections to recently developed sampling methods. The important part of our design are Gabor windows, which we review in Section VIII. In particular, we summarize several methods to generate compactly supported Gabor frames. We demonstrate our theory by several numerical examples in Section IX. Hardware considerations and further applications are discussed in Part II of this series.

II. PROBLEM FORMULATION AND MAIN RESULTS

A. Notation

We will be working throughout the paper with the Hilbert space of complex square integrable functions $L_2(\mathbb{R})$, with inner product

$$\langle f, g \rangle = \int_{-\infty}^{\infty} f(t)\overline{g(t)} dt \quad \text{for all } f, g \in L_2(\mathbb{R}) \quad (1)$$

where $\overline{g(t)}$ denotes the complex conjugate of $g(t)$. The norm induced by this inner product is given by

$$\|f\|_2^2 = \langle f, f \rangle. \quad (2)$$

The Fourier transform of a square integrable function $f(t)$ is defined as

$$\widehat{f}(\omega) = \int_{-\infty}^{\infty} f(t)e^{-2\pi i\omega t} dt \quad (3)$$

and is also square integrable with $\|\widehat{f}\|_2 = \|f\|_2$. In Section III we assume additionally that the Fourier transform of a signal $f \in L_2(\mathbb{R})$ is an integrable function, meaning $\widehat{f} \in L_1(\mathbb{R})$. The $L_1(\mathbb{R})$ norm is given by

$$\|f\|_1 = \int_{-\infty}^{\infty} |f(t)| dt. \quad (4)$$

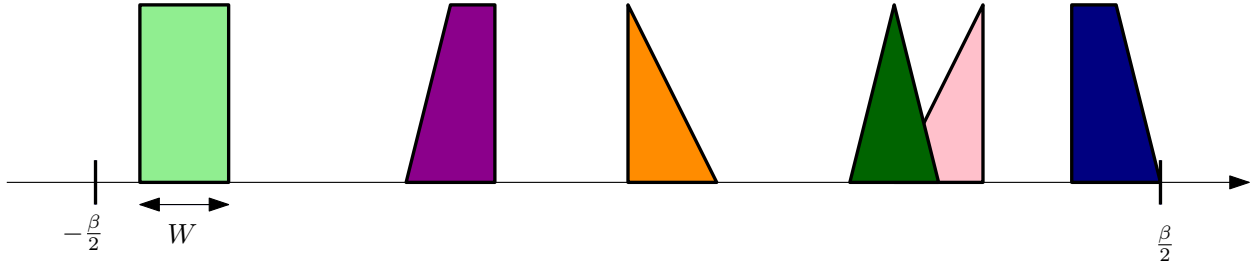


Fig. 1: Schematic example of a multipulse signal with $N = 6$ pulses each of width no more than W . In the example, two of the pulses are overlapping.

A main tool in our derivations are Gabor frames, which we review in Section IV-A. Two important operators that play a central role in Gabor theory, are the translation and modulation operators defined for $x, \omega \in \mathbb{R}$ as

$$T_x f(t) := f(t - x), \quad (5)$$

$$M_\omega f(t) := e^{2\pi i \omega t} f(t), \quad (6)$$

respectively. The composition $M_\omega T_x f(t) = e^{2\pi i \omega t} f(t - x)$ is called a time-frequency shift operator and gives rise to the short-time Fourier transform. For a fixed window $g \in L_2(\mathbb{R})$, the short time Fourier transform of $f \in L_2(\mathbb{R})$ with respect to $g(t)$ is defined as

$$V_g f(x, \omega) := \langle f, M_\omega T_x g \rangle. \quad (7)$$

Many derivations, and especially input-output relations for our sampling systems, will be presented in compact form of matrix multiplications. We denote matrices by boldface capital letters, for example \mathbf{C} , \mathbf{D} , and vectors by boldface low case letters, such as \mathbf{x} , \mathbf{z} .

B. Problem Formulation

We consider the problem of sampling and reconstructing signals comprised of a sum of short, finite duration pulses. A schematic representation of such a signal is depicted in Fig. 1. We do not assume any knowledge of the signal besides the maximum width (support) of the pulses. More formally, we consider real valued signals $f(t)$ of the form

$$f(t) = \sum_{n=1}^N h_n(t), \quad \text{where} \quad \max_n |\text{supp } h_n| \leq W. \quad (8)$$

The number of pulses N and their maximal width W are assumed known. Note that the pulses may overlap in time, as in Fig. 1. Clearly, $f(t)$ is of finite duration. We assume that it is supported on an interval $[-\beta/2, \beta/2]$ with $NW \ll \beta$. Our goal is to recover $f(t)$ from the minimal number of samples possible.

Due to the uncertainty principle, finite duration functions cannot be perfectly bandlimited. However, in practice the main frequency content is typically confined to a finite interval. We refer to such signals as essentially bandlimited.

More formally, we say that $f(t)$ is essentially bandlimited, or ϵ_Ω -bandlimited to $F = [-\Omega/2, \Omega/2]$, if for some $\epsilon_\Omega < 1$

$$\left(\int_{F^c} |\hat{f}(\omega)|^2 d\omega \right)^{1/2} \leq \epsilon_\Omega \|f\|_2. \quad (9)$$

The symbol F^c denotes the complement of the set F . The adjective ‘essential’ refers to the fact that the energy of $\hat{f}(\omega)$ outside $[-\Omega/2, \Omega/2]$ is very small. We denote the set of multipulse signals (8) timelimited to $[-\beta/2, \beta/2]$ and essentially bandlimited to $[-\Omega/2, \Omega/2]$ by $\mathcal{MP}(N, W, \beta, \Omega)$.

There are three interesting special cases that fall into the model (8). The first is when $h_n(t)$ are shifts of a known pulse $h(t)$, so that $h_n(t) = \sigma_n h(t - t_n)$ for some $t_n, \sigma_n \in \mathbb{R}$. In this case, the problem is to find $2N$ parameters, the amplitudes σ_n and shifts t_n . This setting can be treated within the class of finite rate of innovation problems [8]. Recently a method was proposed to sample such signals efficiently at the minimal rate [7], [6]. We will return to this scenario in Section VII and discuss the relation to our work in more detail. A second class, is when the location of the pulses $h_n(t)$ are known but the pulses themselves are not. The third, most difficult scenario, is when neither the locations nor the pulses are known. Our goal is to develop an efficient, robust, and low-rate sampling scheme for this most general scenario. We will see later that our system can be used to sample signals from the other two cases as well, at their respective minimal rates.

We aim at designing a sampling system for signals from the model $\mathcal{MP}(N, W, \beta, \Omega)$ that satisfies the following properties:

- (i) the system has no prior knowledge on the locations or shapes of the pulses;
- (ii) the number of samples should be as low as possible;
- (iii) the reconstruction from the samples should be simple;
- (iv) the reconstructed signal should be close to the original signal.

C. Main Results

The multi-channel sampling method we propose, depicted in Fig. 6, is a mixture of ideas from Gabor theory and Xampling [10], which lies at the heart of sub-Nyquist sampling of analog signals. Since we do not know the pulses making up the signal it is necessary to choose a frame/basis to represent the signal. For efficient sampling, the frame should be chosen such that the coefficients of the signal in this frame reflect its sparse structure in time. It is well known that Gabor coefficients mirror well the local behavior of any square integrable function. However, the Balian-Low theorem [5] precludes Gabor Riesz bases with good time-frequency localization. Therefore one has to settle with a certain degree of redundancy to be able to design well concentrated Gabor atoms.

Our sampling scheme consists of a set of modulators with functions $p_r(t)$, followed by integrators over the interval $[-\beta/2, \beta/2]$. The system depends on an appropriately chosen Gabor frame with redundancy degree $\mu \in (0, 1)$, generated by a compactly supported window that is well localized in the frequency domain. This frame provides a sparse representation for $\mathcal{MP}(N, W, \beta, \Omega)$. The modulating waveforms $p_r(t)$, formally defined in (33), are different, finite superpositions of shifted versions of the chosen Gabor window. The goal of the modulators is to mix together

all windowed pieces of the signal with different weights, so that, a sufficiently large number of mixtures will allow to almost perfectly recover relatively sparse multipulse signals. The analog input signal is first modulated with these waveforms before passing through an integrator. The resulting samples are weighted superpositions of Gabor coefficients of the signal with respect to the chosen frame. Compressed sensing methods [14], [15], [16], [17] are then used to recover the relevant nonzero signal coefficients from the given samples. The number of the rows in the resulting compressed sensing system is about $4N\mu^{-1}$; it is a function of the number of pulses present in the signal and the redundancy of the frame. The number of columns is a function of the desired accuracy of the approximation, and equals about ΩW .

In principle, to reconstruct a time-limited function perfectly, one needs infinitely many samples. We show, that with only a finite number of Gabor samples, with about $4\Omega'WN\mu^{-1}$ samples, related to a somewhat larger interval $[-\Omega'/2, \Omega'/2] \supseteq [-\Omega/2, \Omega/2]$ in the frequency domain, we can reconstruct the signal almost perfectly. The size of the interval $[-\Omega'/2, \Omega'/2]$ is dictated by the chosen Gabor frame. The better frequency localization of the window and the closer the frame bounds to one, the smaller the value of Ω' . The oversampling degree μ dictates the number of samples in time. The bigger μ , the less time samples we need. For $\mu = 1$ it is not possible to construct a window that is well localized in frequency and forms a Gabor frame, therefore we focus on $\mu < 1$. The number $4\Omega'WN\mu^{-1}$ is worse by a factor of two from the minimal number of samples necessary to reconstruct a multipulse signal with the same Gabor frame, if we knew in advance the locations of the pulses. By the minimal number of samples we mean the lowest number of samples necessary to approximate the signal with the desired accuracy, in a frame with a fixed redundancy. Thus only a factor of two in the number of samples is needed in order to compensate for the unknown pulse locations.

After recovering the Gabor coefficients, we recover the signal using a dual Gabor frame. The function $\tilde{f}(t)$ reconstructed from the post-processed coefficients satisfies

$$\|f - \tilde{f}\|_2 \leq \tilde{C}_0(\epsilon_\Omega + \epsilon_B)\|f\|_2 + \tilde{C}_1\|\mathbf{n}_1\|_2 + \tilde{C}_2\|\mathbf{n}_2\|_2, \quad (10)$$

where $f(t)$ is the original signal, \tilde{C}_0 is a constant depending on the Gabor frame, and ϵ_B is related to the essential bandwidth of the chosen Gabor window. The first term in the error is due to the energy of the signal outside the essential bandwidth. The values of \mathbf{n}_1 and \mathbf{n}_2 reflect the noise level in the signal (mismodeling error) and the samples, respectively, while the constants \tilde{C}_1 and \tilde{C}_2 depend on the compressed sensing methods used for recovery of the Gabor coefficients. If the signal is perfectly multipulse and the sampling system is noise free, then $\mathbf{n}_1 = \mathbf{n}_2 = 0$.

We begin our development by first considering the existing sampling techniques and pointing out their drawbacks both in the case when the pulse locations are known and when they are not, in Section III. We show that the present approaches cannot be adapted to treat completely blind sampling, when both the pulse shapes and locations are unknown. In Section IV we briefly introduce the principles of Gabor theory needed in our development and show its usefulness in approximating finite duration signals through truncated Gabor series, with only a finite number of samples. Based on this observation we develop an efficient sampling and reconstruction scheme that treats all time

limited square-integrable signals and is robust to noise, in Section V.

III. PRIOR APPROACHES

The problem of sampling time limited signals is not new. It can be treated by existing methods, such as Fourier series or Shannon's interpolation formula, as we discuss below. However, when the signal exhibits further interior sparsity, as in Fig. 1, these methods cannot be improved to reduce the number of samples. In contrast, our approach allows to fully exploit this sparsity. In addition, even when no internal sparsity exists, our technique works for all square integrable signals and is robust to noise. We begin by reviewing existing approaches and point out their main shortcomings.

When the pulses are unknown, there are infinitely many parameters that allow to uniquely specify a time limited function $f(t)$. One standard method is to use the Fourier series. Every $f \in L_2(\mathbb{R})$ supported on the interval $[-\beta/2, \beta/2]$ can be written as

$$f(t) = \sum_{l \in \mathbb{Z}} \hat{f}\left(\frac{l}{\beta}\right) e^{2\pi i l t / \beta}, \quad \text{a.e. in } t \quad (11)$$

where $\hat{f}\left(\frac{l}{\beta}\right)$ are the Fourier coefficients given by

$$\hat{f}\left(\frac{l}{\beta}\right) = \frac{1}{\beta} \int_{-\beta/2}^{\beta/2} f(t) e^{-2\pi i l t / \beta} dt. \quad (12)$$

Therefore, infinitely many Fourier coefficients have to be used to represent $f(t)$ fully. If $f(t)$ is smooth, or at least $\hat{f} \in L_1(\mathbb{R})$, then the Fourier coefficients decay fast and $f(t)$ may be well approximated by a truncated Fourier series. The approximation error is given by

$$\left\| f(t) - \sum_{|l| \leq L_0} \hat{f}\left(\frac{l}{\beta}\right) e^{2\pi i l t / \beta} \right\|_2 \leq \sum_{|l| > L_0} |\hat{f}(l/\beta)| \quad (13)$$

where L_0 is an integer that specifies the desired approximation, and the number of samples. A sampling system that allows to obtain the desired Fourier coefficients is presented in Fig. 2(a) with parameters $\theta = 1/\beta$, $\tau = 0$ and $s(t) = \frac{1}{\beta} \chi_\beta(t)$, where χ_β is a rectangular function supported on $[-\beta/2, \beta/2]$.

The number $L = 2L_0 + 1$ of channels is related to the essential bandwidth of $f(t)$. If $f(t)$ is ϵ_Ω -essentially bandlimited to $[-\Omega/2, \Omega/2]$, then most of $\hat{f}(\omega)$, with respect to the $L_1(\mathbb{R})$ norm, is concentrated on the larger interval $[-\Omega''/2, \Omega''/2] \supseteq [-\Omega/2, \Omega/2]$. Therefore, L_0 in (13) has to be equal at least $\Omega''\beta/2$ to achieve a good approximation, meaning, at least $\Omega''\beta$ samples are needed to represent $f(t)$ sufficiently well. The smoother $f(t)$, the better the decay properties of $\hat{f}(\omega)$, and the smaller the gap between Ω'' and Ω . This approach however, does not explore additional information about $f(t)$, namely that it occupies only a small portion of the interval $[-\beta/2, \beta/2]$. Intuitively, since $f(t)$ is sparse in time, it should be possible to reduce the number of samples $\Omega''\beta$ without compromising the reconstruction quality.

If additional information about the signal is available, namely time instances when the pulses appear, then the number of samples can be reduced using Shannon's interpolation formula. In [2], [3] the authors consider sampling time limited functions for which $\hat{f} \in L_1$, and provide error estimates on the reconstruction. If $f(t)$ is time-limited

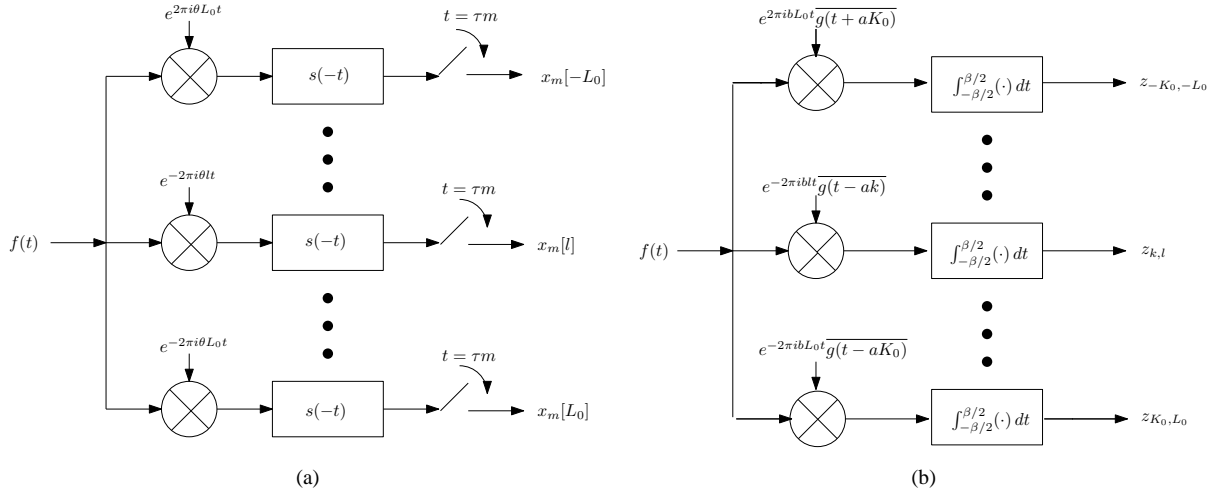


Fig. 2: A general multichannel sampling system (a) and a sampling systems associated with a Gabor frame $\mathcal{G}(g, a, b)$ (b). The systems are equivalent if in (a) the parameters are $\tau = a$, $m = -K_0, \dots, K_0$, $\theta = b$ and the filter $s(t) = \overline{g(t)}$.

to $[-\beta/2, \beta/2]$, then taking pointwise samples at the rate Ω'' , for some $\Omega'' > 0$, and interpolating between the points using Shannon's interpolation formula, gives a finite series

$$(S_{\Omega''} f)(t) = \sum_{|k| \leq K_0} f\left(\frac{k}{\Omega''}\right) \text{sinc}(\pi \Omega'' (t - k/\Omega'')), \quad (14)$$

where K_0 is the largest integer less than $\Omega''\beta/2$. For $k > K_0$, $f(k/\Omega'') = 0$ as $f(t)$ is of finite duration, so that about $\Omega''\beta$ pointwise values of $f(t)$ must be evaluated. It is shown in [2], [3] that the sum $(S_{\Omega''} f)(t)$ differs from $f(t)$ by at most

$$|f(t) - (S_{\Omega''} f)(t)| \leq \int_{|\omega| > \Omega''/2} |\widehat{f}(\omega)| d\omega. \quad (15)$$

The same sampling scheme of Fig. 2(a), but with different parameters, can be used to obtain the pointwise samples. The parameters in this case are $\theta = 0$, which means that we only need one branch, $\tau = 1/\Omega''$ and the filter $s(t) = \delta(t)$, where $\delta(t)$ is a delta function centered at zero.

Again, to obtain a good approximation, Ω'' is chosen with respect to the essential bandwidth of $f(t)$ as in the case of the truncated Fourier series. However, using this method, if the widths and locations of the pulses are known, then the number of samples can be reduced. Let $I_n = [a_n, a_n + W]$ be the active intervals corresponding to the support of each pulse $h_n(t)$. Then the pointwise samples of $f(t)$ can be evaluated only at points k/Ω'' where $k \in \bigcup_{n=1}^N [a_n \Omega'', a_n \Omega'' + W \Omega''] \cap \mathbb{Z}$. Meaning, only about $W \Omega''$ samples per pulse are required, adding up to $NW \Omega''$ samples for the whole signal.

The two sampling solutions presented above suffer from several drawbacks. First, the Fourier series approach does not take into account the sparsity of $f(t)$ in time. Second, if the sparsity is used to reduce the number of samples

using pointwise sampling, then the locations of the pulses have to be known in advance. Even then, the pointwise samples of $f(t)$ are in practice difficult to obtain, as it necessitates ideal sampling (or high analog bandwidth). Finally, both reconstructions require that the Fourier transform of $f(t)$ decays fast outside a certain interval, for the reconstruction to be accurate.

We conclude that a finite duration signal $f(t)$ can be well approximated either by a finite Fourier series or Shannon's interpolation formula as long as its Fourier transform $\hat{f}(\omega)$ decays at least like $|\omega|^{-1-\epsilon}$ for $\epsilon > 0$ and ω outside the essential bandwidth of $f(t)$. Moreover, if the signal is sparse in time, then to reduce the number of samples the locations of the pulses have to be known in advance. The reconstruction of $f(t)$ is either through a Fourier series or Shannon's interpolation formula. Since the exponentials $\{e^{2\pi i l t/\beta}\}_{l \in \mathbb{Z}}$ form an orthonormal basis for the space of continuous $L_2(\mathbb{R})$ functions supported on $[-\beta/2, \beta/2]$, the Fourier series expansion is an exact representation of $f(t)$ in that basis. However, the coefficients $\hat{f}\left(\frac{l}{\beta}\right)$ do not provide any information about local behavior of $f(t)$ in time. On the other hand, the coefficients $f\left(\frac{k}{\Omega''}\right)$ in Shannon's interpolation formula of (14) carry exact information about $f(t)$ in time, but the reconstruction through Shannon's series is exact only in the limit of $\Omega'' \rightarrow \infty$.

To overcome these difficulties, we suggest a compromise between the two representations. We propose using frames for $L_2(\mathbb{R})$ in which multipulse signals are sparse and then use this sparsity, together with the Xampling methodology and compressed sensing techniques, to reduce the number of measurements necessary for a good reconstruction. Instead of taking pointwise samples in time or frequency, in other words, instead of exact localization, we consider averages of $f(t)$ on small intervals. More specifically, the signal $f(t)$ is first windowed with the shifts of some smooth compactly supported window $g(t)$ and then the Fourier transform of the windowed signal is pointwise sampled. This method is referred to as the short-time Fourier transform and is a main tool in Gabor frame theory. It is known that Gabor frames with well chosen windows reflect the local time-frequency concentration of every square integrable function. Therefore they are good candidates to explore time sparsity of multipulse signals. Moreover, the degree of approximation of $f(t)$ by a truncated Gabor series can be controlled both by the decay of $\hat{f}(\omega)$ and by the stability of the frame. In case the frame is tight with frame bounds equal to one, the degree of approximation is controlled by the decay of the Fourier transform of a chosen Gabor window. In Section IV we recall some basic facts and notions from Gabor theory that will be used throughout the paper, and then show how Gabor frames can be used to sample multipulse signal with known pulse locations. In Section V we expand the ideas to treat the unknown setting.

IV. SAMPLING USING GABOR FRAMES

A. Basic Gabor Theory

Regardless of our knowledge regarding the pulses making up the signal $f(t)$, every signal $f \in L_2(\mathbb{R})$ can be represented in some Gabor frame [4]. A collection $\mathcal{G}(g, a, b) = \{M_{bl} T_{ak} g(t) = e^{2\pi i b l t} g(t - ak); k, l \in \mathbb{Z}\}$ is a

Gabor frame for $L_2(\mathbb{R})$ if there exist constants $0 < A_1 \leq A_2 < \infty$ such that

$$A_1 \|f\|^2 \leq \sum_{k,l \in \mathbb{Z}} |\langle f, M_{bl} T_{ak} g \rangle|^2 \leq A_2 \|f\|^2 \quad (16)$$

for all $f \in L_2(\mathbb{R})$. The frame is called tight, if $A_1 = A_2$. By simple normalization every tight frame can be changed to a tight frame with frame bounds equal to one. Therefore, when we talk about tight frames we will mean frames with frame bounds $A_1 = A_2 = 1$.

A Gabor representation of a signal $f(t)$ comprises the set of coefficients $\{z_{k,l}\}_{k,l \in \mathbb{Z}}$ obtained by inner products with the elements of some Gabor system $\mathcal{G}(g, a, b)$ [4]:

$$z_{k,l} = \langle f, M_{bl} T_{ak} g \rangle = e^{2\pi i a k b l} \langle \widehat{f}, M_{-ak} T_{bl} \widehat{g} \rangle. \quad (17)$$

The coefficients $z_{k,l}$ are simply samples of a short-time Fourier transform of $f(t)$ with respect to $g(t)$ at points (ak, bl) . If $\mathcal{G}(g, a, b)$ constitutes a frame for $L_2(\mathbb{R})$, then there exists a function $\gamma \in L_2(\mathbb{R})$ such that any $f \in L_2(\mathbb{R})$ can be reconstructed from the coefficients $\{z_{k,l}\}_{k,l \in \mathbb{Z}}$ using the formula

$$f = \sum_{k,l \in \mathbb{Z}} z_{k,l} M_{bl} T_{ak} \gamma. \quad (18)$$

The Gabor system $\mathcal{G}(\gamma, a, b)$ is the dual frame to $\mathcal{G}(g, a, b)$. Consequently, the window $\gamma(t)$ is referred to as the dual of $g(t)$. Generally, there is more than one dual window $\gamma(t)$. The canonical dual is given by $\gamma = S^{-1}g$, where S is the frame operator associated with $g(t)$, and is defined by $Sf = \sum_{k,l \in \mathbb{Z}} \langle f, M_{bl} T_{ak} g \rangle M_{bl} T_{ak} g$. There are several ways of finding an inverse of S , including the Janssen representation of S , the Zak transform method or iteratively using one of several available efficient algorithms [4].

Throughout the article we will be working only with Gabor frames whose windows are compactly supported on some interval $[-\alpha/2, \alpha/2]$ and lattice parameters $a = \mu\alpha$, $b = 1/\alpha$ for some $\mu \in (0, 1)$. For such frames, the frame operator is a multiplication operator and takes on the particularly simple form

$$S(t) = \sum_{k \in \mathbb{Z}} |g(t - ak)|^2. \quad (19)$$

In this setting, the frame constants can be computed as $A_1 = \text{ess inf } S(t)$ and $A_2 = \text{ess sup } S(t)$. The canonical dual is then $\gamma(t) = bS^{-1}(t)g(t)$. For tight frames the dual atom is simply $\gamma(t) = A_1^{-1}bg(t)$. A necessary condition for $\mathcal{G}(g, a, b)$ to be a frame for $L_2(\mathbb{R})$ is that $ab \leq 1$, while Gabor Riesz bases can only exist if $ab = 1$ [4]. Thus the ratio $1/(ab)$ measures the redundancy of Gabor systems.

Since one key motivation for considering Gabor frames is to obtain a joint time-frequency representation of functions one usually attempts to choose the window $g(t)$ to be well localized in time and frequency. While the Balian-Low theorem [5] makes it impossible to design Gabor Riesz bases with good time-frequency localization, it is not difficult to design Gabor frames with excellent localization properties. For instance, if $g(t)$ is a Gaussian, then we obtain a Gabor frame whenever $ab < 1$. Therefore, to obtain a well localized window one needs to allow for certain redundancy of a Gabor system. Throughout the paper we will be working only with compactly supported windows. In Section VIII we discuss in detail how to construct such frames and their duals.

With a Gabor system $\mathcal{G}(g, a, b)$ we associate a synthesis operator (or reconstruction operator) $A_g : \ell_2(\mathbb{Z}^2) \rightarrow L_2(\mathbb{R})$, defined as

$$A_g c = \sum_{k,l \in \mathbb{Z}} c_{k,l} M_{bl} T_{ak} g(t) \quad \text{for } c \in \ell_2(\mathbb{Z}^2). \quad (20)$$

The adjoint $A_g^* : L_2(\mathbb{R}) \rightarrow \ell_2(\mathbb{Z}^2)$ of A_g is called an analysis operator (or sampling operator), and is given by

$$A_g^* f = \{\langle f, M_{bl} T_{ak} g \rangle\} \quad \text{for } f \in L_2(\mathbb{R}). \quad (21)$$

We will be considering only windows $g(t)$ that are members of so-called Feichtinger algebra, denoted by S_0 [18]. Windows from this space guarantee that the synthesis and analysis mappings are bounded and consequently result in stable reconstructions. It also guarantees that the dual window is in S_0 . To define S_0 let $\varphi(t) = e^{-\pi t^2}$. Then

$$S_0 := \left\{ f \in L_2(\mathbb{R}) \mid \|V_\varphi f\|_1 = \iint |V_\varphi f(x, \omega)| dx d\omega < \infty \right\}, \quad (22)$$

with the norm given by $\|f\|_{S_0} := \|V_\varphi f\|_1$. The definition of S_0 is independent of the window φ , meaning we can take any other $g \in S_0$ instead of φ and we get the same space with equivalent norms. Examples of functions in S_0 are the Gaussian, B-splines of positive order, raised cosine, and any $L_1(\mathbb{R})$ function that is bandlimited or any $L_2(\mathbb{R})$ function that is compactly supported in time with Fourier transform in $L_1(\mathbb{R})$. Note that, the rectangular window is not a member of S_0 since its Fourier transform is not in $L_1(\mathbb{R})$.

For Gabor systems $\mathcal{G}(g, a, b)$ with $g \in S_0$ and $\gamma \in S_0$, the synthesis and analysis operators are bounded and satisfy

$$\|A_g^* f\|_{\ell_2} \leq C_{a,b} \|g\|_{S_0} \|f\|_2; \quad (23)$$

$$\|A_\gamma d\|_2 \leq C_{a,b} \|\gamma\|_{S_0} \|d\|_{\ell_2}, \quad (24)$$

where $C_{a,b} = (1 + 1/a)^{1/2} (1 + 1/b)^{1/2}$ is a constant associated with the time and frequency shifts a and b . These two relations will play a crucial role in estimating the accuracy of the reconstructions.

B. Truncated Gabor Series

We have seen in Section III that time limited $L_2(\mathbb{R})$ functions, whose Fourier transform is additionally in $L_1(\mathbb{R})$, can be well approximated with a finite number of samples using a Fourier series. We now show that the same is true for Gabor series, without assuming anything additional on the signal besides that it is square integrable.

Let $\mathcal{G}(g, a, b)$ be a Gabor frame with $g(t)$ compactly supported on an interval $[-\alpha/2, \alpha/2]$, $a = \mu\alpha$ and $b = 1/\alpha$ for some $\mu \in (0, 1)$. Procedures for constructing such frames with the desired smoothness of the window are presented in [19], [20]. We review these methods and provide some explicit examples in Section VIII. The reason for using compactly supported windows is that for every function $f(t)$ time limited to $[-\beta/2, \beta/2]$, the decomposition of (18) reduces to

$$f = \sum_{k=-K_0}^{K_0} \sum_{l \in \mathbb{Z}} z_{k,l} M_{bl} T_{ak} \gamma, \quad (25)$$

where $\gamma(t)$ is a dual window and K_0 denotes the smallest integer such that the sum in (25) contains all possible non-zero coefficients $z_{k,l}$. The exact value of K_0 is calculated by

$$-\frac{\alpha}{2} + (K_0 + 1)a \geq \frac{\beta}{2} \Rightarrow K_0 = \left\lceil \frac{\beta + \alpha}{2a} \right\rceil - 1. \quad (26)$$

If $a = \alpha$ and $b = 1/\alpha$, then $\mathcal{G}(g, \alpha, 1/\alpha)$ with $g(t)$ a rectangular window supported on $[-\alpha/2, \alpha/2]$ is an orthogonal basis for $L_2(\mathbb{R})$. Moreover, if $\alpha = \beta$, then the decomposition (18) reduces to the Fourier series of $f(t)$. We have already seen that the truncated Fourier series provides a good approximation of $f(t)$ as long as $\hat{f}(\omega)$ decays fast enough. A similar result holds for truncated Gabor expansions. Namely, for a chosen Gabor frame, if $f(t)$ is essentially time-frequency concentrated within $[-\beta'/2, \beta'/2] \times [-\Omega'/2, \Omega'/2]$ with respect to this frame, then $f(t)$ can be approximated with good accuracy by a finite Gabor expansion. The approximation does not demand extra conditions on $\hat{f}(\omega)$. The requirement on the decay of $\hat{f}(\omega)$ is transferred to a proper choice of the Gabor frame. The number of samples necessary for ϵ reconstruction are dictated by the pair of dual windows (g, γ) as incorporated in the following theorem.

Theorem IV.1. *Let $f(t)$ be a finite duration signal supported on the interval $[-\beta/2, \beta/2]$ and ϵ_Ω -bandlimited to $[-\Omega/2, \Omega/2]$. Let $\mathcal{G}(g, a, b)$ be a Gabor frame with $g \in S_0$ that is compactly supported on $[-\alpha/2, \alpha/2]$. Choose $a = \mu\alpha$ for some $\mu \in (0, 1)$, $b = 1/\alpha$ and let $\gamma \in S_0$ be the dual atom. Then for every $\epsilon_B > 0$ there exists an $L_0 < \infty$, depending on the dual window $\gamma(t)$ and the essential bandwidths of $g(t)$ and $f(t)$, such that*

$$\left\| f - \sum_{k=-K_0}^{K_0} \sum_{l=-L_0}^{L_0} z_{k,l} M_{bl} T_{ak} \gamma \right\|_2 \leq \tilde{C}_0(\epsilon_\Omega + \epsilon_B) \|f\|_2, \quad (27)$$

where $K_0 = \left\lceil \frac{\beta + \alpha}{2a} \right\rceil - 1$, $z_{k,l} = V_g f(ak, bl) = \langle f, M_{bl} T_{ak} g \rangle$ are the Gabor coefficients and $\tilde{C}_0 = C_{a,b}^2 \|\gamma\|_{S_0} \|g\|_{S_0}$ with $C_{a,b} = (1 + 1/a)^{1/2} (1 + 1/b)^{1/2}$ a constant depending on the chosen Gabor frame. Moreover, if $g_c(t)$ is a $[-B/2, B/2]$ -bandlimited approximation of $g(t)$ in S_0 , that is $\|g - g_c\|_{S_0} \leq \epsilon_B \|g\|_{S_0}$, then $L_0 = \left\lceil \frac{\Omega + B}{2b} \right\rceil - 1$.

Proof: See Appendix A. ■

There are two important aspects of the theorem to note. First, the constant \tilde{C}_0 depends on the dual Gabor window, so that the approximation error is a function of the dual pair (g, γ) . For a fixed ϵ_B and fixed $\tilde{C}_0(\epsilon_\Omega + \epsilon_B)$ the number L_0 , respectively B , depends on the chosen dual window. If the S_0 norm of γ is big, then we need to increase L_0 , respectively B , to achieve the same degree of approximation even if $g(t)$ is well localized in frequency. When the frame is tight, the number of frequency coefficients depends only on the Gabor window and resembles the situation of Fourier series. If the dual window is the canonical dual, then the smallest L_0 is achieved for tight, or almost tight, Gabor frames with a lower frame constant close to one. This follows from the fact, that for dual Gabor windows $\|\gamma\|_{S_0} \leq A_1^{-1} b \|g\|_{S_0}$, and the shifts of the window forming the frame cover the signal well. Therefore, stable frames, with lower frame bound A_1 away from zero, are the optimal choice.

Second, note that the theorem excludes the case of $\mu = 1$, so that in particular a rectangular window is not allowed. The reason is that for $\mu = 1$, that is when there is no oversampling, it is impossible to construct a frame with a compactly supported window that is well localized in frequency [4]. Meaning, the Fourier transform of

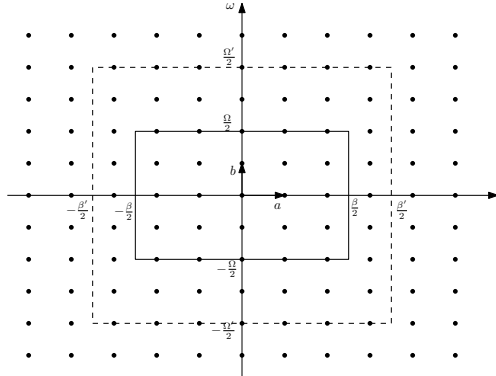


Fig. 3: The rectangular $[-\beta'/2, \beta'/2] \times [-\Omega'/2, \Omega'/2]$ represents the essential time-frequency support of a short-time Fourier transform of a signal $f(t)$ that is time limited to $[-\beta/2, \beta/2]$ and essentially bandlimited to $[-\Omega/2, \Omega/2]$. The short-time Fourier transform is taken with respect to a compactly supported window $g(t)$ that is essentially bandlimited. The size of Ω' depends on the essential bandwidths of $f(t)$ and $g(t)$ and the frame bounds $g(t)$ generates. The black dots denote the time-frequency lattice points (ak, bl) , $k, l \in \mathbb{Z}$.

such a window will not even be L_1 -integrable, therefore not in S_0 . When a window is not well localized in the frequency domain, we need to increase L_0 in order to obtain a small reconstruction error. On the other hand, for $\mu < 1$, it is easy to construct Gabor windows with the desired degree of smoothness or frequency localization [19], [20]. Note that decreasing μ , meaning increasing the redundancy of the frame, gives more freedom in designing well localized Gabor atoms. Therefore, there is a tradeoff between the number of samples in the frequency domain and the number of samples with respect to time. An optimal value of μ lies between $1/2$ and 1 . We will elaborate more on the proper choice of $\mu \geq 1/2$ and present a few constructions of Gabor windows with the desired degree of smoothness that result in stable Gabor frames in Section VIII.

Theorem IV.1 states that finite duration, essentially bandlimited signals, can be well approximated using just the dominant coefficients in the Gabor representation. The sampling system to obtain those coefficients is depicted in Fig. 2(b). An equivalent system is the one in Fig. 2(a) with parameters $\tau = a$, $\theta = b$, and $s(t) = \overline{g(t)}$. If $m = -K_0, \dots, K_0$, then the samples $x_m[l]$ of Fig. 2(a) equal the samples $z_{m,l}$ of Fig. 2(b).

For a fixed class of time limited, essentially bandlimited functions, the number of coefficients, equivalently the number of channels in the sampling scheme, depends on the chosen frame and desired accuracy of the approximation. To minimize that number, we need to choose $\mu \geq 1/2$ (that reduces the number of samples in time) and construct a window that is well localized in frequency and generates a frame with lower frame bound close to one (that reduces the number of samples in frequency). The total number of Gabor coefficients, meaning samples of the short-time Fourier transform, is related to a somewhat larger interval $[-\beta'/2, \beta'/2] \subseteq [-\beta/2, \beta/2]$, with $K \approx \frac{\beta'}{a}$, in the time domain and a larger interval $[-\Omega'/2, \Omega'/2] \subseteq [-\Omega/2, \Omega/2]$, with $L \approx \frac{\Omega'}{b}$, in the frequency domain. Overall, the

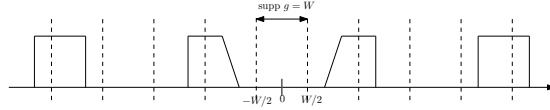


Fig. 4: The relation between f and the shifts of the support of g in the case when $\mu = 1$. When $\text{supp } T_{Wk}g$, for some k , does not overlap any of the pulses of f , then $z_{k,l} = 0$ for all l .

required number of samples is

$$KL \approx \frac{\beta' \Omega'}{a b} = \beta' \Omega' \mu^{-1}. \quad (28)$$

The relation between the respective intervals in the time-frequency plane is schematically depicted in Fig. 3. The lattice points that fall into the rectangular $[-\beta'/2, \beta'/2] \times [-\Omega'/2, \Omega'/2]$ represent the samples of the short-time Fourier transform, and there are exactly KL of them.

When μ is close to one, and $g(t)$ is well localized in frequency forming a tight frame, the number of required samples is close to $\Omega\beta$, by the relation (28). For a fixed μ and a chosen accuracy of approximation, the number of frequency samples in a tight frame depends on the decay properties of $\widehat{g}(\omega)$. Therefore, to minimize the number of channels, we need to choose a window $g(t)$ that exhibits good frequency localization. On the other hand, having already chosen a frame $\mathcal{G}(g, a, b)$, if we desire to improve the accuracy of approximation, then the number L_0 of ‘frequency’ coefficients has to increase.

C. Multipulse Signals with Known Pulse Locations

The Gabor samples are a compromise between the exact pointwise samples of $f(t)$ in time and frequency. If $\alpha \ll \beta$ and the signal has only a few active regions in the interval $[-\beta/2, \beta/2]$, as is the case for multipulse signals, then many of the Gabor coefficients are zero. Indeed, if the shift $g(t - ak)$ does not overlap any active region of $f(t)$ then

$$z_{k,l} = \langle f, M_{bl} T_{ak} g \rangle = \int_{-\beta/2}^{\beta/2} f(t) \overline{g(t - ak)} e^{-2\pi i b l t} dt = 0, \quad (29)$$

for all $l \in \mathbb{Z}$. Therefore, when the locations of the active intervals are known, we can reduce the number of channels in Fig. 2(b) from KL to ML , where $M < K$ is the number of ks , $|k| \leq K_0$, for which $z_{k,l} \neq 0$. To reduce M to minimum, one needs to choose a Gabor frame that allows for the sparsest representation of $f(t)$ with respect to the index k .

For signals from $\mathcal{MP}(N, W, \beta, \Omega)$, an optimal choice is an atom $g(t)$ that is supported on $[-W/2, W/2]$ and shift parameters $a = \mu W$, $b = 1/W$ for some $\mu \in (0, 1)$. In that case at most $\lceil 2\mu^{-1} \rceil$ shifts of $g(t)$ by $ak = \mu W k$ overlap one pulse of $f(t)$. Indeed, when $\mu = 1$ then at most two shifts of $g(t)$ overlap one pulse, as depicted in Fig. 4. When $\mu < 1$, then at most $\lceil 2\mu^{-1} \rceil$ shifts of $\text{supp } g$ overlap one pulse of $f(t)$. This can be calculated from

$$\begin{aligned} \frac{W}{2} < -\frac{W}{2} + \mu W K_1 &\Rightarrow K_1 > \mu^{-1} \\ -\frac{W}{2} > \frac{W}{2} + \mu W K_2 &\Rightarrow K_2 < -\mu^{-1} \end{aligned} \quad \Longrightarrow K_1 - K_2 > 2\mu^{-1}. \quad (30)$$

Let \mathbf{Z} denote the $K \times L$ matrix of dominant Gabor coefficients. In terms of Fig. 3 each dot in the rectangular $[-\beta'/2, \beta'/2] \times [-\Omega'/2, \Omega'/2]$ represents one element of \mathbf{Z} . Note that the columns of the rectangular are the rows of the matrix \mathbf{Z} . Therefore, for functions $f \in \mathcal{MP}(N, W, \beta, \Omega)$ each column $\mathbf{Z}[l] = [z_{-K_0, l}, \dots, z_{K_0, l}]^T$ of \mathbf{Z} has at most $\lceil 2\mu^{-1} \rceil$ nonzero entries. Moreover, all columns $\mathbf{Z}[l]$ have nonzero entries at the same places, as modulations $e^{2\pi i b l t}$ applied to $f(t)$ do not change the positions of the pulses. These observations lead to the following proposition, a corollary of Theorem IV.1:

Proposition IV.2. *Let $\mathcal{G}(g, a, b)$ be a Gabor frame with sampling parameters $a = \mu W$ and $b = 1/W$ for some $0 < \mu < 1$. Choose a window $g(t)$ supported on $[-W/2, W/2]$ that is ϵ_B -essentially bandlimited to some $[-B/2, B/2]$ in the S_0 norm. If $f \in \mathcal{MP}(N, W, \beta, \Omega)$ is a multipulse signal and the locations of the pulses are known, then we need only ML channels in Fig. 2(b), with $M = \lceil 2\mu^{-1} \rceil N$ and $L = 2L_0 + 1$, L_0 given in (69), to be able to reconstruct $f(t)$ with $\epsilon = \epsilon_\Omega + \epsilon_B$ accuracy.*

The recovered signal, that is an ϵ approximation of the original signal in the proposition, is obtained by utilizing the Gabor synthesis operator with a dual atom $\gamma(t)$ using only ML coefficients $z_{k, l}$. The total number of necessary samples, when the locations of the pulses are known, is about $2\Omega'WN\mu^{-1}$.

D. Method Comparison

Since time limited functions can be reconstructed only to a certain accuracy, we refer to the minimal number of samples as the minimal number required to reconstruct the signal with a desired accuracy. When working with generalized samples, like the Gabor samples, this number depends on the sampling function $g(t)$ and the redundancy of the frame. For an ϵ accuracy of approximation using the Fourier series and Shannon's interpolation methods, the minimal number of samples is of order $\Omega''\beta$, $\Omega'' > \Omega$, where

$$\int_{F_1^c} |\hat{f}(\omega)| d\omega \leq \epsilon \|f\|_2, \quad (31)$$

and $F_1 = [-\Omega''/2, \Omega''/2]$. For a Gabor frame with redundancy μ , we achieve ϵ approximation with a minimal number of samples of order $\Omega'\beta'\mu^{-1}$ as long as the Gabor window $g(t)$ and its dual $\gamma(t)$ are such that

$$\left(\int_E \int_{F_2^c} |V_g f(x, \omega)|^2 dx d\omega \right)^{1/2} \leq \frac{\epsilon}{C_{a,b} \|\gamma\|_{S_0}} \|f\|_2, \quad (32)$$

where $E = [-\beta'/2, \beta'/2]$ and $F_2 = [-\Omega'/2, \Omega'/2]$. When $g(t)$ generates a tight, or almost tight, frame, and $\hat{f} \in L_1(\mathbb{R})$ or the window $g(t)$ decays fast in frequency, $\Omega' \lesssim \Omega''$.

Table I compares the number of samples necessary for a good approximation of time limited signals in the case of these three methods. As can be seen from the table, the Gabor frame has two main advantages. The first is that it does not require strong decay of $\hat{f}(\omega)$. Second, this approach can be used to efficiently sample multipulse signals with unknown pulse locations, as we will show in the next section. In this case we need approximately $4\Omega'WN\mu^{-1}$ samples which is minimal with respect to the chosen accuracy of the approximation and redundancy of the frame.

	Fourier series	Shannon's interpolation	Gabor series with $\mathcal{G}(g, a, b)$, $ab = \mu$
number of samples	$\approx \Omega''\beta$	$\approx \Omega''\beta$	$\approx \Omega'\beta'\mu^{-1}$
number of samples for multipulse signal with known pulse locations	$\approx \Omega''\beta$	$\approx \Omega''WN$	$\approx 2\Omega'WN\mu^{-1}$
number of samples for multipulse signal with unknown pulse locations	$\approx \Omega''\beta$	$\approx \Omega''\beta$	$\approx 4\Omega'WN\mu^{-1}$
additional conditions	$\hat{f} \in L_1(\mathbb{R})$	$\hat{f} \in L_1(\mathbb{R})$	none
approximation error	(13)	(15)	Theorem IV.1

TABLE I: Comparison of three methods for approximating $L_2(\mathbb{R})$ functions that are time limited to $[-\beta/2, \beta/2]$ and essentially bandlimited to $[-\Omega/2, \Omega/2]$. The second and third lines refer to multipulse signals with N pulses, each of width no more than W . The methods are compared for the same accuracy of approximation. If we assume that $\hat{f} \in L_1(\mathbb{R})$ for the Gabor sampling scheme with a tight Gabor frame, then $\Omega' \approx \Omega''$.

To compare the number of samples in the three methods we need to assume that $\hat{f} \in L_1(\mathbb{R})$ and that $\mathcal{G}(g, a, b)$ is a tight frame. In comparison to the Fourier and pointwise samples, the number of samples necessary to well approximate time limited signals in the Gabor series method is greater by a factor of μ^{-1} , which is the degree of redundancy of the Gabor frame. This comes as no surprise, since the Fourier series is an expansion of a function in an orthogonal basis (hence no oversampling), while Gabor series is an expansion of a signal in a redundant dictionary. If the signal is additionally sparse in time and the locations of the pulses are known, then Gabor series require less samples than the Fourier series but more than Shannon's interpolation method. By Proposition IV.2, we need around $2\Omega'WN\mu^{-1}$ samples in the Gabor frame, to be able to reconstruct the signal well. If $2WN\mu^{-1} < \beta$, then sampling with Gabor atoms outperforms the Fourier series approach in terms of the number of samples. For the reconstruction using Fourier coefficients, we need $\Omega''\beta$ samples since this method does not take sparsity into account. On the other hand, Shannon's interpolation method requires less samples, about $\Omega''WN$, but cannot be extended to the unknown setting.

V. SAMPLING OF MULTIPULSE SIGNALS

We now present a sampling scheme for functions from $\mathcal{MP}(N, W, \beta, \Omega)$ that reduces the number of channels in Fig. 2(b) and does not require knowledge of the pulse locations.

A. Sampling System

Our system, shown in Fig. 6, exploits the sparsity of multipulse signals in time. The signal $f(t)$ enters R channels simultaneously. In the r th channel, $f(t)$ is multiplied by a mixing function $p_r(t)$, followed by an integrator. The design parameters are therefore the number of channels R and the mixing functions $p_r(t)$ for $1 \leq r \leq R$. The role of the mixing functions is to gather together all the information in $f(t)$ over the entire interval $[-\beta/2, \beta/2]$. Namely, $f(t)$ is windowed with shifts of some compactly supported function, and all the windowed versions are summed together with different weights.

The functions $p_r(t)$ are constructed from the Gabor frame. Let $\mathcal{G}(g, a, b)$ be a Gabor frame with window $g(t)$ supported on the interval $[-W/2, W/2]$, essentially bandlimited to $[-B/2, B/2]$, and with sampling parameters $a = \mu W$ and $b = 1/W$ for some $0 < \mu < 1$. Then the waveforms $p_r(t)$ are

$$p_r(t) = e^{-2\pi i b l t} \sum_{k=-K_0}^{K_0} c_{mk} \overline{g(t - ak)}, \quad (33)$$

where $r = (m, l)$ is a double index with $l = -L_0, \dots, L_0$, $m = 0, \dots, M - 1$,

$$K_0 = \left\lceil \frac{\beta + W}{2W\mu} \right\rceil - 1 \quad \text{and} \quad L_0 = \left\lceil \frac{(\Omega + B)W}{2} \right\rceil - 1. \quad (34)$$

The waveforms $p_r(t)$ are basically mixtures of K channels of the sampling scheme of Fig. 2(b) corresponding to the same frequency shift bl . An example of such a signal is depicted in Fig. 5 with $g(t) = \cos(\pi t/W)$ on $[-W/2, W/2]$ and zero otherwise, $a = W/2$ and $b = 1/W$. To specify $p_r(t)$ completely, it remains to choose the coefficients c_{mk} . To do so, we first analyze the effect of the sampler on the unknown signal and derive the relation between the samples x_r and the signal $f(t)$.

Consider the r th channel:

$$\begin{aligned} x_r &= \int_{-\beta/2}^{\beta/2} f(t) p_r(t) dt = \sum_{k=-K_0}^{K_0} c_{mk} \langle f, M_{bl} T_{ak} g \rangle \\ &= \sum_{k=-K_0}^{K_0} c_{mk} z_{k,l}. \end{aligned} \quad (35)$$

The above relation ties the known x_r to the unknown Gabor coefficients of $f(t)$ with respect to $\mathcal{G}(g, a, b)$. This relation is key to the recovery of $f(t)$. If we can recover $z_{k,l}$ from the samples x_r , then by Theorem IV.1 we are able to recover $f(t)$ almost perfectly. As can be seen from this relation, the goal of the modulator $p_r(t)$ is to create mixtures of the unknown Gabor coefficients $z_{k,l}$. These mixtures, when chosen appropriately, will allow to recover $z_{k,l}$ from a small number r by exploiting the sparsity of these coefficients and relying on ideas of compressed sensing. Note, that when using the scheme of Fig. 2, each x_r is equal to one value of $z_{k,l}$, so that no combinations are obtained. When $z_{k,l}$ are sparse, with unknown sparsity locations, we will need to acquire all its values using this approach. In contrast, obtaining mixtures of $z_{k,l}$, allows reduction in the number of samples.

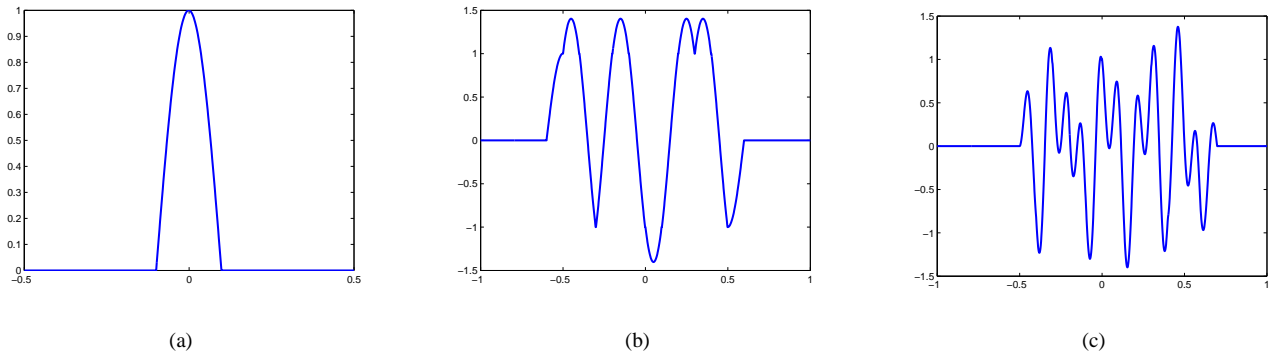


Fig. 5: An example of a waveform $p_r(t)$. (a) cosine window $g(t)$ of width $0.2\mu s$ generating a Gabor frame with shift parameter $a = 0.1\mu s$ used to create $p_r(t)$. (b) a weighted sum of shifted copies of $g(t)$; the weighting coefficients are $c_{mk} = \pm 1$ and $K_0 = 5$. (c) the real part of a full waveform $p_r(t)$, which is a modulation of (a) by $e^{2\pi i b t}$.

B. Signal Recovery

For our purposes, it is convenient to write (35) in matrix form as

$$\mathbf{X} = \mathbf{C}\mathbf{Z}, \quad (36)$$

where \mathbf{X} is a matrix of size $M \times L$ with ml th element equal x_r , for $r = (m, l - L_0)$ and $m = 0, \dots, M - 1$, $l = 0, \dots, 2L_0$. The unknown Gabor coefficients are gathered in the $K \times L$ matrix \mathbf{Z} with columns $\mathbf{Z}[l] = [z_{-K_0, l}, \dots, z_{K_0, l}]^T$, $l = -L_0, \dots, L_0$. The $M \times K$ matrix \mathbf{C} contains the coefficients $\mathbf{C}_{mk} = c_{m, k - K_0}$, $k = 0, \dots, 2K_0$, $m = 0, \dots, M - 1$. The coefficients, or equivalently the matrix \mathbf{C} , has to be chosen such that it is possible to retrieve \mathbf{Z} from the relation (36). Note, that if $M = K$ and \mathbf{C} is an identity matrix, then the system of Fig. 6 reduces to that of Fig. 2(b).

The choice of $a = W\mu$ and window $g(t)$ supported on $[-W/2, W/2]$ results in a $K \times L$ matrix \mathbf{Z} of dominant coefficients from which $f(t)$ can be well reconstructed. Such choice of a frame guarantees that for every l , the column vectors $\mathbf{Z}[l]$ have only $\lceil 2\mu^{-1} \rceil N$ out of K nonzero entries, and the nonzero entries correspond to the locations of the pulses. We conclude that each $\mathbf{Z}[l]$ is $\lceil 2\mu^{-1} \rceil N$ -sparse and all $\mathbf{Z}[l]$ have nonzero entries on the same rows due to the structure of $f(t)$.

The following theorem states the conditions under which one can uniquely reconstruct the Gabor coefficients $z_{k, l}$, $|k| \leq K_0$ and $|l| \leq L_0$, from the outputs x_r .

Theorem V.1. *Let $f \in \mathcal{MP}(N, W, \beta, \Omega)$ be a multipulse signal. Let $\mu \in (0, 1)$ and $\mathcal{G}(g, a, b)$, with $a = W\mu$ and $b = 1/W$, be a Gabor frame with $g(t)$ compactly supported on $[-W/2, W/2]$ and ϵ_B -bandlimited to $[-B/2, B/2]$ in the S_0 norm. Consider the sampling scheme of Fig. 6 with the following parameters:*

- 1) $K_0 = \lceil (\beta + W)/(2W\mu) \rceil - 1$;
- 2) $L_0 = \lceil (\Omega + B)W/2 \rceil - 1$;

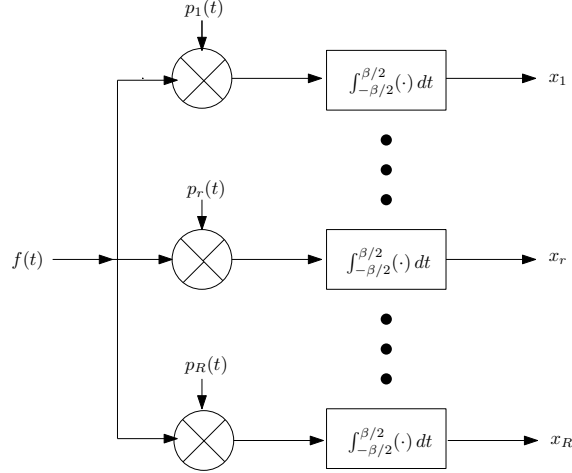


Fig. 6: An efficient sampling system for multipulse signals.

- 3) $p_r(t) = e^{-2\pi i b l t} \sum_{k=-K_0}^{K_0} c_{mk} \overline{g(t - ak)}$ with $r = (m, l)$ where $m = 0, \dots, M - 1$ and $l = -L_0, \dots, L_0$;
 4) $M \geq \lceil 2\mu^{-1} \rceil N$ for non-blind reconstruction or $M \geq 2\lceil 2\mu^{-1} \rceil N$ for blind.

If every set of $2\lceil 2\mu^{-1} \rceil N$ columns of \mathbf{C} is linearly independent, then \mathbf{Z} is a unique sparse solution of (36). Moreover, the function $\sum_{k=-K_0}^{K_0} \sum_{l=-L_0}^{L_0} z_{k,l} M_{bl} T_{ak} \gamma$ reconstructed from the obtained coefficients satisfies (27), with $\gamma \in S_0$ denoting the dual atom of $g(t)$.

In the case of known positions of the pulses, referred to as non-blind, the sampling scheme of Fig. 6 achieves the minimal sampling rate for the desired accuracy of the approximation and a given frame, that is $M = \lceil 2\mu^{-1} \rceil N$, as discussed at the end of Section IV-B. In the blind setting, when the locations of the pulses are not known, the sampling rate increases by a factor of two. Note, that the scheme is efficient only when the pulses occupy less than half of the overall support of f , that is when $2\lceil 2\mu^{-1} \rceil N < K$.

Proof: Let $f \in \mathcal{MP}(N, W, \beta, \Omega)$. Then the $K \times L$ matrix \mathbf{Z} of dominant Gabor coefficients is row sparse with only $\lceil 2\mu^{-1} \rceil N$ nonzero rows. In the non-blind setting, when the locations of the nonzero coefficients are known, the conditions on M and the matrix \mathbf{C} ensure that (36) can be inverted on the proper column set, thus providing the uniqueness claim. Let \mathcal{S} denote the index set of nonzero coefficients and $\mathbf{C}_{\mathcal{S}}$ be a submatrix which contains the columns of \mathbf{C} indexed by \mathcal{S} . A closed form expression providing the solution is

$$\mathbf{Z}^{\mathcal{S}}[l] = \mathbf{C}_{\mathcal{S}}^{\dagger} \mathbf{X}[l] \quad (37)$$

where $\mathbf{Z}^{\mathcal{S}}[l]$ contains only entries of $\mathbf{Z}[l]$ indexed by \mathcal{S} and $\mathbf{C}_{\mathcal{S}}^{\dagger} = (\mathbf{C}_{\mathcal{S}}^H \mathbf{C}_{\mathcal{S}})^{-1} \mathbf{C}_{\mathcal{S}}^H$ is the (Moore-Penrose) pseudoinverse of $\mathbf{C}_{\mathcal{S}}$. For k not in \mathcal{S} , $z_{k,l} = 0$.

In blind recovery, the nonzero locations of $\mathbf{Z}[l]$ are unknown. A well known result from the compressed sensing literature is that an S -sparse vector \mathbf{u} is the unique solution of $\mathbf{v} = \mathbf{C}\mathbf{u}$ if every $2S$ columns of \mathbf{C} are linearly independent [21]. This condition translates into $M \geq 2\lceil 2\mu^{-1} \rceil N$ and the condition on \mathbf{C} of the theorem. ■

As we have seen, the matrix \mathbf{Z} that we would like to recover from the measurements \mathbf{X} is row sparse. This problem is referred to as a multiple measurement vector (MMV) problem and has been treated extensively in the literature. Several algorithms have been developed that exploit this structure to recover \mathbf{Z} efficiently from \mathbf{X} in polynomial time when M is increased beyond $M = 2S$ [22], [14], [15], [23], [24], [11], [25], [16], [17], [26]. For example, a popular approach to recovering \mathbf{Z} is by solving the convex problem

$$\min_{\mathbf{Z}} \|\mathbf{Z}\|_{2,1} \quad \text{subject to} \quad \mathbf{X} = \mathbf{CZ}, \quad (38)$$

where $\|\mathbf{Z}\|_{2,1} = \sum_{k=1}^K (\sum_{l=1}^L |z_{k,l}|^2)^{1/2}$.

Consider the system $\mathbf{V} = \mathbf{C}\mathbf{U}$, where \mathbf{U} is an unknown $K \times L$ row-sparse matrix, \mathbf{V} is the measurement matrix of size $M \times L$ and \mathbf{C} is of size $M \times K$. A matrix \mathbf{C} is said to have the restricted isometry property (RIP) of order S , if there exists $0 \leq \delta_S < 1$ such that

$$(1 - \delta_S) \|\mathbf{u}\|_2^2 \leq \|\mathbf{C}\mathbf{u}\|_2^2 \leq (1 + \delta_S) \|\mathbf{u}\|_2^2 \quad (39)$$

for all S -sparse vectors \mathbf{u} [23], [27]. The requirement of Theorem V.1 translates into $\delta_{2S} < 1$. It is well known that Gaussian and Bernoulli random matrices, whose entries are drawn independently with equal probability, have the RIP of order S if $M \geq cS \log(K/S)$, where c is a constant [28], [29]. For random partial Fourier matrices the respective condition is $M \geq cS \log^4(K)$ [30], [31].

C. Equivalent Representation

In terms of the number of channels, the system depicted in Fig. 6 can still be improved. For a fixed Gabor frame $\mathcal{G}(g, a, b)$, as specified in the Theorem V.1, the number of branches can be reduced to L if instead of LM modulations followed by an integrator, we perform L modulations followed by a filter $s(t)$. Consider the system in Fig. 2(a) with $\theta = b$, $\tau = WK$, where $K = 2K_0 + 1$, and K_0 is as in Theorem V.1, and the filter $s(t)$ given by

$$s(t) = \sum_{m=0}^{M-1} s_m(t + WKm) \quad (40)$$

where

$$s_m(t) = \sum_{k=-K_0}^{K_0} c_{mk} \overline{g(t - ak)}. \quad (41)$$

Note, that for all m , $s_m(t)$ is compactly supported in time on $[-W/2 - \mu WK_0, W/2 + \mu WK_0]$, and that its support contains the support $[-\beta/2, \beta/2]$ of $f(t)$. The shifted versions $s_m(t + WKm)$ have non-overlapping supports as the width of $\text{supp } s_m$ is smaller than the shift step WK

$$W(1 + 2\mu K_0) < W(1 + 2K_0) = WK. \quad (42)$$

The support relation between the filter $s(t)$ and the multipulse signal $f(t)$ is depicted in Fig. 7.

Under these assumptions, the output of the ℓ th channel is given by

$$\begin{aligned} x_m[l] &= (e^{-2\pi i b l t} f(t) * s(-t))[WKm] \\ &= \sum_{n=0}^{M-1} \langle M_{-bl} f, T_{WK(m-n)} \overline{s_n} \rangle. \end{aligned} \quad (43)$$

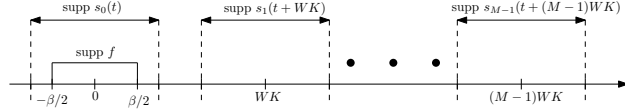


Fig. 7: Relation between the support of the filter $s(t)$, which is the sum of the shifted supports of $s_m(t)$, and the support of the signal f .

The sum is nonzero only when $m - n = 0$, because otherwise the support of $s_n(t)$ shifted by $WK(m - n)$ does not overlap the support of $f(t)$, as depicted in Fig. 7. Therefore it is sufficient to sample only at points $t = WKm$ for $m = 0, \dots, M - 1$, leading to

$$x_m[l] = \langle M_{-bl}f, \overline{s_m} \rangle = \langle f, M_{bl}\overline{s_m} \rangle = \sum_{k=-K_0}^{K_0} c_{mk} z_{k,l} \quad (44)$$

where $z_{k,l} = \langle f, M_{bl}T_{ak}g \rangle$ are the Gabor coefficients. Evidently, if the coefficients c_{km} used to build the blocks $s_m(t)$ of the filter $s(t)$ are the same as coefficients used to create the waveforms $p_r(t)$, then the two systems are equivalent.

VI. NOISY MEASUREMENTS

Up until now we considered signals that were exactly multipulse and such that their samples were noise free. A more realistic situation is when the measurements are noisy and/or the signal $f(t)$ is not exactly multipulse, having some energy leaking outside the pulses. In this section we show that our sampling scheme is robust to bounded noise in both the signal and the samples.

We say that a signal $f(t)$ essentially bandlimited to $[-\Omega/2, \Omega/2]$, is essentially multipulse with N pulses each of width no more than W , if for some $\delta_W < 1$ there exists an $f_p \in \mathcal{MP}(N, W, \beta, \Omega)$ such that

$$\|f - f_p\|_2 \leq \delta_W \|f\|_2. \quad (45)$$

We assume throughout this section that the signals are time limited to the interval $[-\beta/2, \beta/2]$, meaning that the energy leaks only between the pulses, and denote this class of signals by $\mathcal{MP}_{ess}(N, W, \beta, \Omega)$.

Since the energy of $f \in \mathcal{MP}_{ess}(N, W, \beta, \Omega)$ leaks beyond the support of the pulses, the column vectors $\mathbf{Z}[l]$, of the $K \times L$ matrix \mathbf{Z} of dominant coefficients, defined in (36), are no longer sparse. Nonetheless, the following lemma shows that \mathbf{Z} can be approximated by a sparse matrix.

Lemma VI.1. *Let $f \in \mathcal{MP}_{ess}(N, W, \beta, \Omega)$ be δ_W -essentially multipulse and $\mathcal{G}(g, a, b)$ be a Gabor frame with g compactly supported on $[-W/2, W/2]$ and $a = \mu W$, $b = 1/W$ for some $0 < \mu < 1$. Then there exists a subset \mathcal{S} of $\{-K_0, \dots, K_0\}$ such that*

$$\|\mathbf{Z} - \mathbf{Z}^{\mathcal{S}}\|_{2,1} \leq \delta_W \sqrt{K} C_{a,b} \|g\|_{s_0} \|f\|_2, \quad (46)$$

where $\mathbf{Z}^{\mathcal{S}}$ consists of rows of \mathbf{Z} indexed by \mathcal{S} and $K = 2K_0 + 1$ and $\|\mathbf{Z}\|_{2,1} = \sum_{k=-K_0}^{K_0} (\sum_{l=-L_0}^{L_0} |z_{k,l}|^2)^{1/2}$.

Proof: See Appendix B. ■

Lemma VI.1 states that, if the Gabor coefficients $z_{k,l}$ of $f(t)$ decay fast outside the support of the pulses, then the coefficients corresponding to the locations \mathcal{S} of the pulses provide a good approximation. The estimate (46) is rather crude and in practice the errors are much smaller. Since $f(t)$ is multipulse from $\mathcal{MP}_{ess}(N, W, \beta, \Omega)$ and the Gabor frame is as in Theorem V.1, the index set \mathcal{S} has at most $S = \lceil 2\mu^{-1} \rceil$ elements. We refer to $\mathbf{Z}^{\mathcal{S}}$ as the best S -term approximation of \mathbf{Z} . A good S -term approximation of \mathbf{Z} in ℓ_2 norm can be obtained by utilizing compressed sensing algorithms for (36). If the matrix \mathbf{C} has RIP constant $\delta_{2S} \leq \sqrt{2} - 1$, then there exists a unique sparse solution [14], [26], $\tilde{\mathbf{Z}}$ of

$$\min_{\mathbf{Z}} \|\mathbf{Z}\|_{2,1} \quad \text{subject to} \quad \mathbf{X} = \mathbf{C}\mathbf{Z} \quad (47)$$

that satisfies [26]

$$\|\mathbf{Z} - \tilde{\mathbf{Z}}\|_2 \leq C_1 \|\mathbf{Z} - \mathbf{Z}^{\mathcal{S}}\|_{2,1}, \quad (48)$$

where C_1 is a constant depending on δ_{2S} . In particular, if \mathbf{Z} is row sparse, as is the case for $f \in \mathcal{MP}(N, W, \beta, \Omega)$, then $\mathbf{Z} = \mathbf{Z}^{\mathcal{S}}$ and we recover \mathbf{Z} .

Assuming now that the sampling system of Fig. 6 has also some imperfections in the form of noise added to the samples, the input-output relation can be written as

$$x_r = \sum_{k=-K_0}^{K_0} c_{m,k} z_{k,l} + n_r, \quad (49)$$

where n_r stands for the noise added to the samples. In matrix form this becomes

$$\mathbf{X} = \mathbf{C}\mathbf{Z} + \mathbf{N} \quad (50)$$

where \mathbf{Z} is a $K \times L$ matrix of Gabor coefficients and \mathbf{N} is an $M \times L$ noise matrix with ml th element n_r where $r = (m, l - L_0)$, $m = 0, \dots, M - 1$ and $l = 0, \dots, 2L_0$. Under certain conditions on the matrix \mathbf{C} and the noise \mathbf{N} it is possible to find, from the relation (50), a unique sparse matrix $\tilde{\mathbf{Z}}$ such that a function synthesized from it is a good approximation of the original signal.

Theorem VI.2. *Let $f \in \mathcal{MP}_{ess}(N, W, \beta, \Omega)$ be δ_W -essentially multipulse sampled with the sampling system described in Theorem V.1, where the matrix \mathbf{C} has RIP constant $\delta_{2S} \leq \sqrt{2} - 1$ with $S = \lceil 2\mu^{-1} \rceil N$. Let the samples be additionally corrupted by a bounded noise \mathbf{N} . Then there exists a unique S -row sparse matrix $\tilde{\mathbf{Z}}$ approximating \mathbf{Z} , and*

$$\begin{aligned} \left\| f - \sum_{k=-K_0}^{K_0} \sum_{l=-L_0}^{L_0} \tilde{z}_{k,l} M_{bl} T_{ak} \gamma \right\|_2 &\leq \\ &\leq \tilde{C}_0 (\epsilon_B + \epsilon_\Omega) \|f\|_2 + \tilde{C}_1 \|\mathbf{Z} - \mathbf{Z}^{\mathcal{S}}\|_{2,1} + \tilde{C}_2 \|\mathbf{N}\|_2 \end{aligned} \quad (51)$$

where $\tilde{C}_0 = C_{a,b}^2 \|\gamma\|_{S_0} \|g\|_{S_0}$, $\tilde{C}_1 = C_{a,b} \|\gamma\|_{S_0} C_1$ and $\tilde{C}_2 = C_{a,b} \|\gamma\|_{S_0} C_2$.

Proof: The proof is a combination of Theorem V.1 and results relating to MMV systems from [26]. If the noise \mathbf{N} is bounded and matrix \mathbf{C} has a RIP constant $\delta_{2S} \leq \sqrt{2} - 1$, then

$$\min \|\mathbf{Z}\|_{2,1} \quad \text{subject to} \quad \|\mathbf{CZ} - \mathbf{X}\|_2 \leq \|\mathbf{N}\|_2 \quad (52)$$

has a unique S -sparse solution $\tilde{\mathbf{Z}}$. It follows from [26] that the solution to (52) obeys

$$\|\mathbf{Z} - \tilde{\mathbf{Z}}\|_2 \leq C_1 \|\mathbf{Z} - \mathbf{Z}^S\|_{2,1} + C_2 \|\mathbf{N}\|_2, \quad (53)$$

where C_1 and C_2 are constants depending on δ_{2S} , and again \mathbf{Z}^S denotes the best S -term approximation of \mathbf{Z} , i.e., the support of \mathbf{Z}^S consists of the indices corresponding to the S rows with largest ℓ_2 norms.

It remains to show that the error is bounded. From Theorem V.1 and estimates (46) and (53) we have

$$\begin{aligned} & \left\| f - \sum_{k=-K_0}^{K_0} \sum_{l=-L_0}^{L_0} \tilde{z}_{k,l} M_{bl} T_{ak} \gamma \right\|_2 \leq \\ & \leq \left\| f - \sum_{k=-K_0}^{K_0} \sum_{l=-L_0}^{L_0} z_{k,l} M_{bl} T_{ak} \gamma \right\|_2 \\ & + \left\| \sum_{k=-K_0}^{K_0} \sum_{l=-L_0}^{L_0} (z_{k,l} - \tilde{z}_{k,l}) M_{bl} T_{ak} \gamma \right\|_2 \\ & \leq \tilde{C}_0 (\epsilon_B + \epsilon_\Omega) \|f\|_2 + C_{a,b} \|\gamma\|_{S_0} \|\mathbf{Z} - \tilde{\mathbf{Z}}\|_2. \end{aligned} \quad (54)$$

Inserting the relation (53) into the last expression completes the proof. \blacksquare

In particular, if \mathbf{Z} is row sparse, as is the case for $f \in \mathcal{MP}(N, W, \beta, \Omega)$, then $\mathbf{Z} = \mathbf{Z}^S$ and the error of the approximation depends only on the noise added to the samples. When the signal is essentially multipulse, then the error bound depends on the decay of the coefficients as stated in Lemma VI.1. If that quantity is small, then a good approximation of $f(t)$ is achieved by synthesizing a signal from the solution $\tilde{\mathbf{Z}}$ of (52). Note here, that if the dual window $\gamma(t)$ is compactly supported, then a function reconstructed from the coefficients $\tilde{\mathbf{Z}}$ is multipulse.

VII. RELATED WORK

Recently, the ideas of compressed sensing have been extended to the analog domain to allow for sub-Nyquist sampling of analog signals [12], [32], [26], [11], [6], [7], [33]. These works follow the Xampling paradigm, which provides a framework for incorporating and exploiting structure in analog signals without the need for discretization [9], [10]. Two of these sub-Nyquist solutions are closely related to our scheme: the first is a sub-Nyquist sampling architecture for multiband signals introduced in [11], while the second is a sampling system for multipulse signals with known pulse shape introduced in [6]. We now briefly comment on the connection of our results to these works. The observations made here will be expanded in the second part of this series, in which we generalize our sampling scheme by certain mixing of the channels. We will show that by proper mixing of the channels, we can sample efficiently both multipulse signals with known pulse shape, and time limited signals that are essentially multiband, connecting our results more explicitly to prior sampling architectures.

A. The Modulated Wideband Converter

The concept of using modulation waveforms is based on ideas presented in [11] for a multiband model, which is Fourier dual to ours: the signals in [11] are assumed to be sparse in frequency, while multipulse signals are sparse in time. More specifically, [11] considers multiband signals whose Fourier transform is concentrated on N frequency bands, and the width of each band is no greater than B . The locations of the bands are unknown in advance. A low rate sampling scheme, called the modulated wideband converter (MWC), allowing recovery of such signals at the rate of $4NB$ was proposed in [11]; a hardware prototype appears in [10]. This scheme consists of parallel channels where in each channel the input is modulated with a periodic waveform followed by a low-pass filter and low-rate uniform sampling. The main idea is that in each channel the spectrum of the signal is scrambled, such that a portion of the energy of all bands appears at baseband. Therefore, the input to the sampler contains a mixture of all the bands. Mixing of the frequency bands in [11] is analogous to mixing the Gabor coefficients in our scheme.

Despite the similarity, there are some important differences between the systems. First, in the mixing stage we use time-limited and non-periodic waveforms, while the MWC relies on periodic functions. Second, following the mixing stage, we use an integrator in contrast to the low-pass filter in [11]. These differences are due to the fact that we are interested in different quantities: content of the signal on time intervals in our work as opposed to frequency bands in [11]. However, in both systems the mixing is used for the same purpose: to reduce the sampling rate relative to the Nyquist rate.

We will show in the second part of this series that by choosing slightly different waveforms $p_r(t)$ and mixing them in a particular way, we can use our sampling scheme to sample time limited, essentially multiband signals at low rates.

B. Multipulse Signals with Known Pulse Shape

Another related signal model is that of multipulse signals with known pulse shapes [7], [6], [8]:

$$f(t) = \sum_{s=1}^S \sigma_s h(t - t_s) \quad (55)$$

where $h(t)$ is known. This problem reduces to finding the amplitudes σ_s and time delays t_s . Under certain assumptions on the pulse $h(t)$ it is possible to recover the amplitudes and shifts from a finite number of Fourier coefficients of $f(t)$, and therefore to reconstruct $f(t)$ perfectly. The recovery process is a two step method. First the time-delays can be estimated using nonlinear techniques e.g. the annihilating filter method [8], as long as the number of measurements L satisfies $L \geq 2S$ and the time-delays are distinct. Once the time delays are known, the amplitudes can be found via a least squares approach. The number of channels is motivated by the number of unknown parameters (σ_s, t_s) which equals $2S$.

The Fourier coefficients can be determined from samples of $f(t)$ using a scheme similar to that of Fig. 6 with L channels and modulators $p_l(t) = e^{-2\pi i b l t}$ with $b = 1/\beta$. In this case, the input-output relation becomes $\mathbf{x} = \mathbf{f}$, where \mathbf{x} is a vector of length L and \mathbf{f} is a vector of Fourier coefficients of $f(t)$ of length L . In [6] the authors

proposed a more general scheme based on mixing the modulations $e^{-2\pi itl/\beta}$ with proper coefficients, resulting in periodic waveforms, before applying them to the signal. The corresponding samples are weighted superpositions of the Fourier coefficients \mathbf{f} . When the weights are properly chosen, \mathbf{f} can be recovered and therefore the time-delays and amplitudes as well. In the second part of this series, we incorporate the idea of mixing the channels into our sampling system and show that under certain conditions on the Gabor frame, our generalized system can be used to sample signals of the form (55). We note here, that the system of [6] is inefficient for our signal model, since it reduces to the Fourier series method, which does not take sparsity in time into account.

VIII. GABOR WINDOWS

The sampling scheme presented in this paper is based on Gabor frames. As discussed in previous sections, a Gabor frame that is not too redundant, meaning $\mu \in [\frac{1}{2}, 1)$, with a compactly supported window guarantees a small number of time samples. The number of samples in frequency is dictated by the localization of the window in frequency and whether it forms a stable frame (lower frame constant away from zero). In part II of these series we will show that, if the window forms a partition of unity, then the system can also be used to sample finite rate of innovation signals of the form (55). We recall here some methods to construct Gabor frames with well localized windows for a chosen redundancy μ . This material is an accumulation of the main results from [19] and [20].

Daubechies *at al.* [19] developed a method to construct tight Gabor frames that are compactly supported in time and with the desired decay rate in frequency. Their method works for all ranges of μ , however, since we are interested in less redundant frames, we focus here only on the technique for $\mu \geq \frac{1}{2}$. A window $g(t)$ that is supported on $[-W/2, W/2]$ and forms a frame with $a = \mu W$ and $b = 1/W$ can be constructed from an everywhere increasing function $h(t)$ such that $h(t) = 0$ for $t \leq 0$, and $h(t) = 1$ for $t \geq 1$ by

$$g(t) = \begin{cases} 0, & t \leq -\frac{W}{2}, \\ \left[h\left(\frac{t/W+1/2}{1-\mu}\right) \right]^{1/2}, & t \in \left[-\frac{W}{2}, -\frac{W\lambda}{2}\right], \\ 1, & |t| \leq \frac{W\lambda}{2}, \\ \left[1 - h\left(\frac{t/W-\lambda/2}{1-\mu}\right) \right]^{1/2}, & t \in \left[\frac{W\lambda}{2}, \frac{W}{2}\right], \\ 0, & t \geq \frac{W}{2}, \end{cases} \quad (56)$$

where $\lambda = 2\mu - 1$. The function $g(t)$ is non-negative, has the desired support and equals 1 on $[-W\lambda/2, W\lambda/2]$. If $h(t)$ is taken to be $2k$ continuously differentiable, than $g(t)$ is k times continuously differentiable, which implies that $\hat{g}(\omega)$ decays like $(1 + |\omega|)^{-k-\epsilon}$, $\epsilon > 1$. The points $t = \pm W\lambda/2$, where $g(t)$ becomes constant, have been chosen so that their distance to the furthest edge of $\text{supp } g$ is exactly μW . The frame bounds of such a constructed frame equal $A_1 = A_2 = 1$ [19], since

$$\sum_{k \in \mathbb{Z}} |g(t + k\mu W)|^2 = 1 \quad (57)$$

An example of a window illustrating the above construction is

$$g(t) = \begin{cases} 0 & |t| \geq W/2, \\ \cos(\pi t/W) & |t| \leq W/2, \end{cases} \quad (58)$$

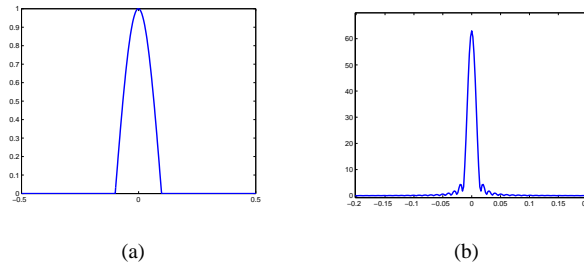


Fig. 8: Tight Gabor window (a) and its Fourier transform (b) created using the method from [19] with $\mu = \frac{1}{2}$.

where $\mu = \frac{1}{2}$ and a corresponding function $h(t) = \sin(\pi t/2)$ on $[0, 1]$. The window $g(t)$ is depicted in Fig. 8 together with its Fourier transform.

Other tight Gabor frames with $a = \mu W$ and $b = 1/W$ can be constructed from any nonnegative and bounded window $h(t)$ supported on $[-W/2, W/2]$ such that $\sum_{k \in \mathbb{Z}} h(t - \mu W k) = 1$, by setting $g(t) = \sqrt{h(t)}$. However, due to the square root, the behavior of $g(t)$ in the frequency domain is not apparent. A square root reduces smoothness at the edge of the support. To compensate for this loss of smoothness, $h(t)$ must be constructed to have a higher order of vanishing derivatives at the endpoints.

An alternative construction for $\mu \geq 1/2$ was developed in [34], [35]. The method results in a spline type windows $g(t)$ of any order of smoothness that satisfy the partition of unity criterion. The windows and their duals are constructed so that they are supported on $[-1, 1]$. Using the dilation operator, the windows can be changed to be supported on any symmetric interval. The constructions are made by counting the number of constraints (in the Ron-Shen duality condition [36], and on the points where continuity/differentiability is required) and then searching for polynomials on $[-1, 0]$ and on $[0, 1]$ of a matching degree. The coefficients in the polynomials are found by Mathematica. The author provides many examples of pairs of dual Gabor frames. One example is $\tilde{g}(t)$ supported on $[-2/3, 2/3]$ and given by

$$\tilde{g}(t) = \begin{cases} 2 + 3t & t \in [-2/3, -1/3], \\ 1 & |t| \leq 1/3, \\ 2 - 3t & t \in [1/3, 2/3], \end{cases} \quad (59)$$

that forms a frame with $a = 1$ and $b = 3/4$. It forms a partition of unity with a shift parameter $a = 1$, $\sum_{k \in \mathbb{Z}} \tilde{g}(t - k) = 1$. The dual window is also supported on $[-2/3, 2/3]$ and is given by

$$\tilde{\gamma}(t) = \begin{cases} -18t^2 - 15t - 2 & t \in [-2/3, -1/3], \\ 1 & |t| \leq 1/3, \\ -18t^2 + 15t - 2 & t \in [1/3, 2/3]. \end{cases} \quad (60)$$

The windows are plotted in Fig. 9(a) and (b), respectively. The dual window $\tilde{\gamma}$, although not canonical, is an element of S_0 , and the bounds developed in the previous sections hold true for this pair. Applying dilation by $(\mu W)^{-1}$,

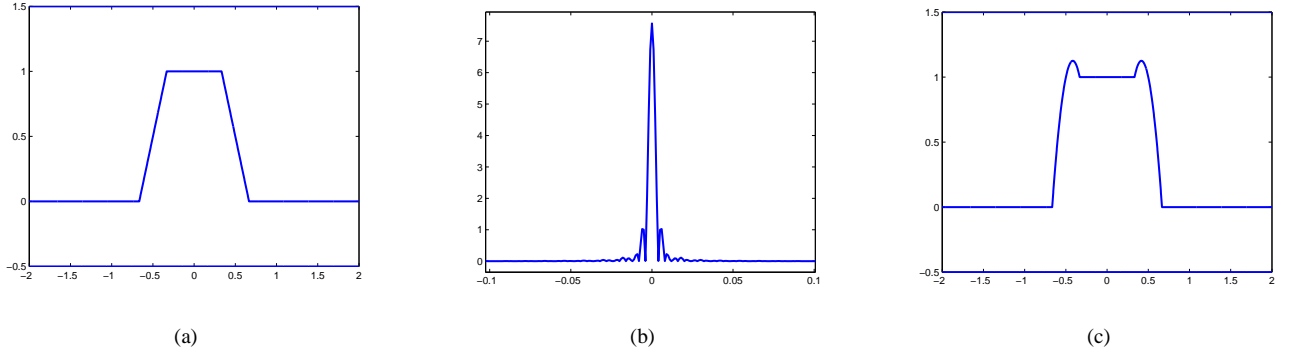


Fig. 9: A pair of dual Gabor windows generated by a method described in [34] for $\mu = 3/4$. Plots (a) and (b) depict the window and its Fourier transform, while (c) is the dual window given by (60).

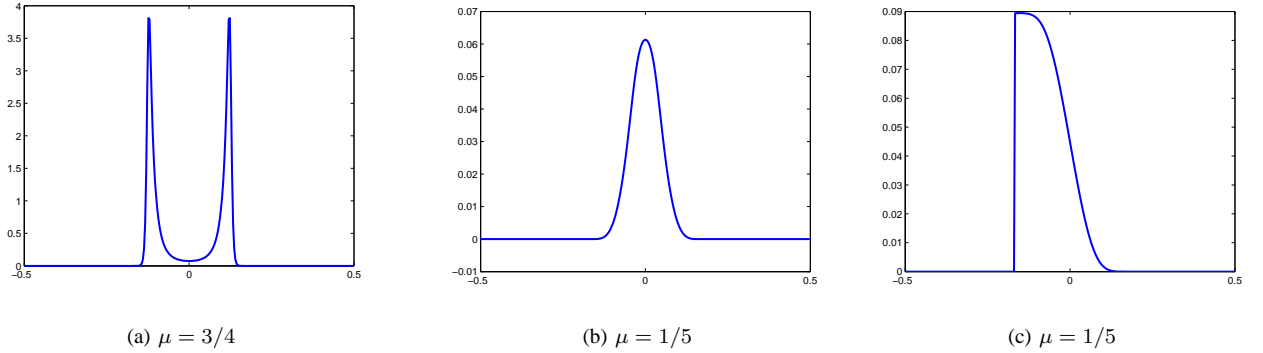


Fig. 10: The dual windows for B-spline $B_5(t)$ of order 5 for different values of the oversampling parameter μ . For $\mu = 3/4$ the canonical dual is depicted in (a) and for $\mu = 1/5$ the canonical dual is plotted in (b). When $\mu = 1/5$ another dual can be computed using (63) (c).

with $\mu = 3/4$, both to $\tilde{g}(t)$ and $\tilde{\gamma}(t)$ we obtain a dual pair of windows $g(t)$ and $\gamma(t)$

$$g(t) = \tilde{g}(t/(\mu W)) \quad \gamma(t) = \tilde{\gamma}(t/(\mu W)) \quad (61)$$

that are supported on $[-W/2, W/2]$, and such that $\mathcal{G}(g, \mu W, 1/W)$ forms a frame with frame bounds $A_1 = 1/2$ and $A_2 = 1$. Moreover, $g(t)$ forms a partition of unity with shift parameter $a = \mu W$, $\sum_{k \in \mathbb{Z}} g(t - \mu W k) = 1$.

Another construction that results in functions $g(t)$ satisfying the partition of unity criterion was developed in [20]. The advantage of this method is the simple computation of a dual window, however, it comes with a cost of higher redundancy of the frame, meaning smaller μ . The method starts with a bounded, real valued function $\tilde{g}(t)$ supported on an interval $[-N/2, N/2]$, satisfying

$$\sum_{n \in \mathbb{Z}} \tilde{g}(t - n) = 1. \quad (62)$$

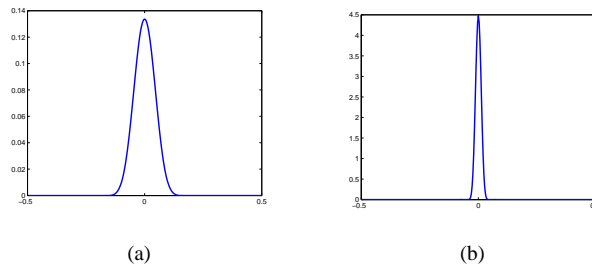


Fig. 11: B-spline $B_5(t)$ of order 5 (a) and its Fourier transform (b).

Let $a, b > 0$ be given such that $\mu = ab \leq \frac{1}{N}$. Take

$$\tilde{\gamma}(t) = \left(\mu \tilde{g}(t) + 2\mu \sum_{n=1}^{N-1} \tilde{g}(t+n) \right) \chi_{N/2}(t). \quad (63)$$

Then the functions $g(t) = \tilde{g}(t/a)$ and $\gamma(t) = \tilde{\gamma}(t/a)$ generate dual Gabor frames with parameters a and b , [20]. Moreover, $g(t)$ forms a partition of unity with shifts ak . A cut by a rectangular function in (63) results in a dual window that is not continuous. However, if $\mu \leq \frac{1}{2N-1}$ the cut-off is not necessary and we obtain a pair of dual windows that have the same smoothness properties.

In [20] the author chooses $\tilde{g}(t)$ to be a B-spline. Let $B_N(t)$ be a spline of order N ,

$$B_1(t) = \chi_{1/2}(t), \quad B_{N+1}(t) = (B_N * B_1)(t). \quad (64)$$

Then $B_N(t)$ is supported on $[-N/2, N/2]$ and forms a partition of unity with shift parameter $a = 1$. To generate a Gabor frame from $B_N(t)$ with a window supported on $[-W/2, W/2]$ and lattice parameters $a = \mu W$, $b = 1/W$, such that the window forms a partition of unity with shift μW , we need to choose $\mu = 1/N$ [37]. Then $g(t) = B_N(tN/W)$ is supported on the desired interval and decays like $(1 + |\omega|)^{-N-\epsilon}$ in the frequency domain. Note that μ decreases as the order N of smoothness of the B-spline is increased. Thus smoother windows can be obtained only at the cost of a smaller μ . However, already for $N = 3$ we get good concentration properties of $g(t)$. The canonical dual window for the window $g(t)$ generated from a B-spline of order 5 is plotted in Fig. 10(b). The oversampling parameter μ is $1/5$ and the dual can be computed by inverting the Gabor frame operator. On the other hand, the dual generator built using (63), although not continuous, is a finite linear combination of $g(t)$ and therefore easy to construct. It is depicted in Fig. 11(c).

B-splines of positive order give rise also to other less redundant Gabor frames. In [38], [39] it was shown that $B_N(t)$, which is supported on $[-N/2, N/2]$, forms a frame for $L_2(\mathbb{R})$ whenever $a < N$ and $b \leq 1/N$. Let $\mu < 1$ and choose $a = \mu N$ and $b = 1/N$. To generate a Gabor frame from $B_N(t)$ with a window supported on $[-W/2, W/2]$, for any W , and lattice parameters $a = \mu W$, $b = 1/W$, we need to apply a dilation operator to $B_N(t)$. The resulting window is $g(t) = B_N(tN/W)$. The drawback of choosing $\mu \geq 1/2$ is that the frame becomes unstable, as the lower frame constant approaches zero, and the shifts of the window will not properly cover the signal. That means, that

even though the window is well localized in frequency, we would need many frequency samples to achieve a good approximation of the signal by a truncated Gabor series. To obtain the dual window necessary for the reconstruction of the signal we invert the Gabor frame operator. Fig. 11(a) depicts a Gabor window generated from a B-spline of order 5. It can be seen that it is well concentrated in the frequency domain (b). The canonical dual window for $\mu = 3/4$ is plotted in Fig. 10(a).

IX. SIMULATIONS

We now present some numerical experiments illustrating the recovery of multipulse signals from the samples obtained by our sampling scheme.

We consider a multipulse signal with three pulses, each of width $W = 0.13s$, randomly distributed in the interval $[-4s, 4s]$. The pulses are a B-spline of order 2, B-spline of order 4 and a cosine pulse. The signal is essentially bandlimited to $[-10Hz, 10Hz]$. We examine the performance of the system of Fig. 6 for five different frame windows constructed in Section VIII:

- 1) Gabor frame with trapezoidal window $g(t)$ defined in (59) of width $0.13s$ with lattice parameters $a = 0.098s$, $b = 7.6Hz$, which results in oversampling of $\mu = 3/4$. The frame bounds are $A_1 = 1/2$ and $A_2 = 1$. The window, its Fourier transform, and dual window are depicted in Fig. 9;
- 2) Gabor frame with fifth order B-spline window $g(t) = B_5(t)$ supported on an interval of width $0.13s$ with the same lattice parameters as the previous frame. The frame is very unstable, with lower frame bound $A_1 = 0.0003$, and upper frame bound $A_2 = 1$. The window, its Fourier transform and dual window are depicted in Fig. 11;
- 3) tight Gabor frame with cosine window $g(t)$ defined in (58) of width $0.13s$ and lattice parameters $a = 0.065s$ and $b = 7.6Hz$, which results in $\mu = 1/2$. The frame bounds are $A_1 = A_2 = 1$. The window together with its Fourier transform is depicted in Fig. 8;
- 4) Gabor frame with window $g(t) = B_2(t)$ being a second order B-spline supported on the interval of width $0.13s$. The frame bounds are $A_1 = 1/2$ and $A_2 = 1$. The lattice parameters are the same as for a cosine window. The window, together with its Fourier transform and a canonical dual are plotted in Fig. 12;
- 5) Gabor frame with fifth order B-spline window $g(t) = B_5(t)$ supported on the interval of width $0.13s$. The lattice parameters are $a = 0.026s$ and $b = 7.6Hz$, which results in $\mu = 1/5$. The frame bounds are $A_1 = 1.1$ and $A_2 = 1.2$. The window, its Fourier transform and a canonical dual are depicted in Fig. 11 and Fig. 10.

The signal was first sampled with the sampling scheme of Fig. 6. The coefficient matrix \mathbf{C} that defines the waveforms $p_r(t)$ has entries ± 1 chosen from a Bernoulli process. A few examples are depicted in Fig. 13 for different window choices $g(t)$. Table II compares the performance of our sampling system for the five different Gabor frames specified earlier. All five windows have approximately the same essential bandwidth $[-8Hz, 8Hz]$, with different degree of decay, for a good approximation we have to take at least $L = 5$ frequency coefficients. This follows from equation (34) where L_0 becomes $L_0 = 2$. For frames with redundancy factor $\mu = 3/4$, the number of samples in time is $K = 83$, with $K_0 = 41$. When sparsity is taken into account, we can reduce that number

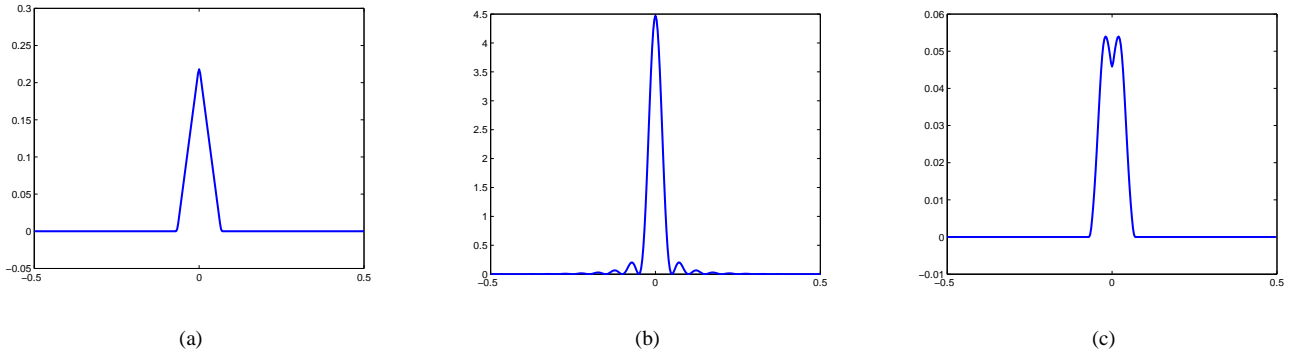


Fig. 12: Gabor frame windows $B_2(t)$ (a) together with its Fourier transforms (b) and a canonical dual (c).

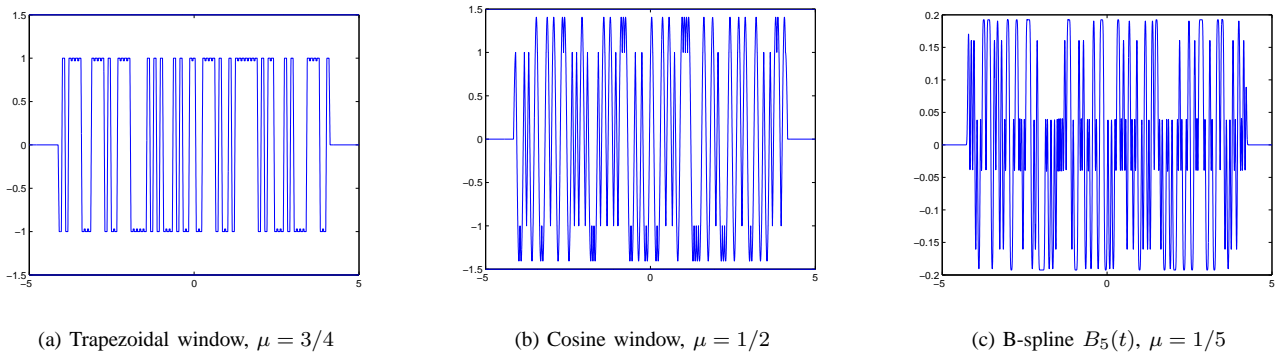


Fig. 13: The waveforms $p_r(t)$ corresponding to different frames. Plotted are only waves $p_r(t)$ for which $\ell = 0$.

to $M = 18$ and still achieve the same quality of reconstruction because the excluded Gabor samples are zero, and hence do not contribute to the quality of the reconstruction, by Theorem IV.1 and Theorem V.1. In our simulations we took M even smaller, $M = 12$, which resulted in $ML = 60$ samples. The decrease in the number of samples is 7-fold. For the frame generated by a B-spline of order five with $\mu = 3/4$ the number of time samples is the same as when using the trapezoidal window, however for a good reconstruction many frequency samples are necessary even though $B_5(t)$ decays fast in frequency. The reason for this is that the window $B_5(t)$ with $\mu = 3/4$ forms an unstable frame, with a lower frame bound equal $A_1 = 0.0003$. To reduce the reconstruction error we would need to take many more frequency samples for Gabor coefficients to preserve the energy of the signal. On the other hand, a Gabor frame generated by $B_5(t)$ but with oversampling factor $\mu = 1/5$ is a stable frame and the reconstruction is good with just $L = 5$ frequency coefficients. However, since the oversampling is high, the price we pay is a high number of time samples $K = 313$. Knowing that the signal is sparse, the number of samples can be significantly reduced to $M = 60$, resulting in the overall number of $ML = 300$ samples. For the frames with redundancy $\mu = 1/2$, the number M increases from 18 to 24, however here, we also took smaller number of

window	number of samples without sparsity, KL	number of samples with sparsity, ML	recovery error
trapezoidal, $\mu = 3/4$	415	60	0.0803
$B_5(t)$, $\mu = 3/4$	415	60	3.3896
cosine, $\mu = 1/2$	625	90	0.0540
$B_2(t)$, $\mu = 1/2$	625	90	0.0631
$B_5(t)$, $\mu = 1/5$	1565	300	0.0514

TABLE II: Comparison of the performance of the sampling scheme for different Gabor frames. The number of frequency samples was taken everywhere the same, with $L = 5$. The recovery error is the same if we take sparsity into account or not by Theorem IV.1 and Theorem V.1.

$M = 18$. Then, the number of necessary samples reduces from $KL = 125 \cdot 5 = 625$, when we do not take sparsity into consideration, to $ML = 18 \cdot 5 = 90$ when we know that there are $N = 3$ pulses. We can see from Table II that the reconstruction error using cosine window is slightly smaller than one using B-spline of order 2. This is due to the fact that cosine windows decays slightly better in frequency than the B-spline.

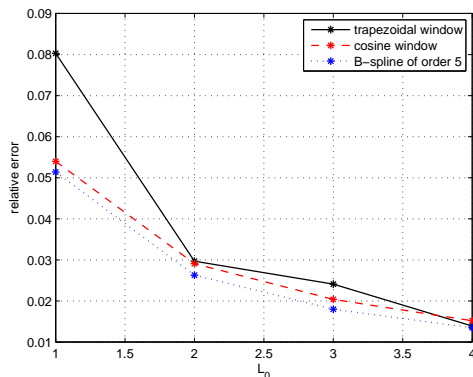


Fig. 14: Relative error of the reconstruction using different windows.

In Fig. 14 we analyze the dependency of the accuracy of the reconstruction on the number L of oscillators with respect to the different windows. The experiments confirm the theory, namely that increasing L reduces the relative error $\|f - \tilde{f}\|_2 / \|f\|_2$ and therefore improves the accuracy of the approximation. Moreover, the best performance with respect to the relative error is due to the B-spline of order five and oversampling factor $\mu = 1/5$. This comes as no surprise, as $B_5(t)$ decays much faster in frequency than other three windows, and with $\mu = 1/5$ it forms a partition of unity. However, while in theory for trapezoidal window we need $M = 18L$ samples and for cosine and $B_2(t)$ windows we need $ML = 24L$ samples, the number of samples for $B_5(t)$ with $\mu = 1/5$ increases to $ML = 60L$.

X. CONCLUSIONS

We presented an efficient sampling scheme for multipulse signals, which is designed independently of the time support of the input signal. Our system allows to sample multipulse signals at the minimal rate, far below Nyquist, without any knowledge of the pulse shapes or its locations. The scheme fits into the broad context of Xampling - a recent sub-Nyquist sampling paradigm for analog signals. Our architecture relies on Gabor frames which lead to sparse expansions of multipulse signals, and consists of modulating the signal with several waveforms followed by integration. We showed that the Gabor coefficients, necessary for reconstruction, can be recovered from the samples of the system by utilizing compressed sensing techniques. The number of necessary samples depends on the desired accuracy of the approximation, essential bandwidth of the signal, and redundancy factor μ related to the Gabor frame, and equals $4\Omega'NW\mu^{-1}$. This is greater by a factor of two from the situation when the pulse locations are known in advance. The increase in the number of samples is a result of not knowing the locations. The proposed sampling and recovery technique is stable with respect to noise and mismodeling.

In Part II of this series we generalize our sampling scheme in several directions. In particular, we consider practical implementations from a hardware point of view that also, with an appropriately chosen Gabor frame, allow to efficiently sample signals with known pulse shapes at the rate of innovation. In addition, this generalization leads to further reduction in sampling rate when the multipulse signals are also sparse in frequency, in particular radar-like functions.

APPENDIX A

PROOF OF THEOREM IV.1

The proof is rooted in that of Theorem 3.6.15 in [18] with appropriate adjustments. Since $\mathcal{G}(g, a, b)$ is a Gabor frame, $f(t)$ admits a decomposition

$$f = \sum_{k=-K_0}^{K_0} \sum_{l \in \mathbb{Z}} z_{k,l} M_{bl} T_{ak} \gamma. \quad (65)$$

Let $\epsilon_B > 0$. The bandlimited S_0 functions are dense in S_0 , therefore, there exists $g_c \in S_0$ bandlimited to some $[-B/2, B/2]$, such that

$$\|g - g_c\|_{S_0} \leq \epsilon_B \|g\|_{S_0}. \quad (66)$$

Since $f(t)$ is an essentially bandlimited function, there exists a function $f_c(t)$ bandlimited to $[-\Omega/2, \Omega/2]$, such that

$$\|f - f_c\|_2 \leq \epsilon_\Omega \|f\|_2. \quad (67)$$

Consequently, $|z_{k,l}| = |\langle \widehat{f}_c, M_{-ak} T_{bl} \widehat{g}_c \rangle| \neq 0$ only for those ℓ such that $\text{supp } \widehat{f}_c \cap (\text{supp } \widehat{g}_c + bl) \neq \emptyset$, that is

$$[-\Omega/2, \Omega/2] \cap [bl - B/2, bl + B/2] \neq \emptyset. \quad (68)$$

The fact that $f_c(t)$ and $g_c(t)$ are bandlimited implies that there are only a finite number of values ℓ for which $V_{g_c} f_c(ak, bl) \neq 0$. Let L_0 be the smallest integer such that $|V_{g_c} f_c(ak, bl)| = 0$ for $|l| > L_0$. The exact value of L_0

can be calculated as

$$L_0 = \left\lceil \frac{\Omega + B}{2b} \right\rceil - 1. \quad (69)$$

Define a sequence $d_{k,l}$ as

$$d_{k,l} = \begin{cases} z_{k,l}, & |k| \leq K_0, |l| > L_0 \\ 0, & \text{else.} \end{cases} \quad (70)$$

Then $|d_{k,l}| \leq |V_{g-g_c} f(ak, bl) + V_{g_c}(f - f_c)(ak, bl)|$ for all $k, l \in \mathbb{Z}$, and

$$\begin{aligned} & \left\| f - \sum_{k=-K_0}^{K_0} \sum_{l=-L_0}^{L_0} z_{k,l} M_{bl} T_{ak} \gamma \right\|_2 = \\ & = \left\| \sum_{k \in \mathbb{Z}} \sum_{l \in \mathbb{Z}} d_{k,l} M_{bl} T_{ak} \gamma \right\|_2 \leq C_{a,b} \|\gamma\|_{S_0} \|d\|_{\ell_2} \\ & \leq C_{a,b} \|\gamma\|_{S_0} (\|V_{g-g_c} f\|_{\ell_2} + \|V_{g_c}(f - f_c)\|_{\ell_2}) \\ & \leq C_{a,b}^2 \|\gamma\|_{S_0} \|g\|_{S_0} (\epsilon_B + \epsilon_\Omega) \|f\|_2 \end{aligned} \quad (71)$$

where we first used the boundedness of the analysis operator A_g^* and then the synthesis operator A_γ whenever g and γ are in S_0 .

APPENDIX B

PROOF OF LEMMA VI.1

Let $f_p \in \mathcal{MP}(N, W, \beta, \Omega)$ be a multipulse δ_W -approximation of f . Then $V_g f_p(ak, bl) = 0$ for all $|k| > K_0$, and the column vectors $[V_g f_p(-aK_0, bl), \dots, V_g f_p(aK_0, bl)]^T$, $|l| \leq L_0$, are all jointly sparse with $\lceil 2\mu^{-1} \rceil N$ nonzero coefficients. Let \mathcal{S} denote the index set of nonzero coefficients. For $|\ell| \leq L_0$, let $\mathbf{Z}^{\mathcal{S}}[\ell]$ be vectors with coefficients $z_{k,l}^{\mathcal{S}}$ defined by

$$z_{k,l}^{\mathcal{S}} = \begin{cases} z_{k,l} & k \in \mathcal{S} \\ 0 & k \notin \mathcal{S} \end{cases} \quad (72)$$

Then $\mathbf{Z}^{\mathcal{S}}[\ell]$ is the best $\lceil 2\mu^{-1} \rceil N$ -term approximation of $\mathbf{Z}[\ell]$, for each ℓ . Note that $|z_{k,l} - z_{k,l}^{\mathcal{S}}| \leq |V_g(f - f_p)(ak, bl)|$ for all k and ℓ , so that

$$\begin{aligned} \|\mathbf{Z} - \mathbf{Z}^{\mathcal{S}}\|_{2,1} &= \sum_{k=-K_0}^{K_0} \left(\sum_{l=-L_0}^{L_0} |z_{k,l} - z_{k,l}^{\mathcal{S}}|^2 \right)^{1/2} \\ &\leq \sum_{k=-K_0}^{K_0} \left(\sum_{l=-L_0}^{L_0} |V_g(f - f_p)(ak, bl)|^2 \right)^{1/2} \\ &\leq \sum_{k=-K_0}^{K_0} \|V_g(f - f_p)(ak, \cdot)\|_{\ell_2} \\ &\leq \sqrt{K} \|V_g(f - f_p)\|_{\ell_2} \\ &\leq \sqrt{K} C_{a,b} \|g\|_{S_0} \|f - f_p\|_2 \\ &\leq \delta_W \sqrt{K} C_{a,b} \|g\|_{S_0} \|f\|_2, \end{aligned} \quad (73)$$

completing the proof.

REFERENCES

- [1] G. B. Folland and A. Sitaram, “The uncertainty principle: a mathematical survey,” *J. Fourier Anal. Appl.*
- [2] P. L. Butzer and W. Splettstösser, “A sampling theorem for duration-limited functions with error estimates,” *Information and Control*, vol. 34, 1977.
- [3] P. L. Butzer and R. L. Stens, “Sampling theory for not necessarily band-limited functions: A historical overview,” *SIAM Review*, vol. 34, no. 1, 1992.
- [4] K. Gröchenig, *Foundations of Time-Frequency Analysis*. Birkhäuser, Boston, 2001.
- [5] J. J. Benedetto, C. Heil, and D. F. Walnut, “Gabor systems and the Balian-Low theorem,” in *Gabor Analysis and Algorithms: Theory and Applications*, H. Feichtinger and T. Strohmer, Eds. Birkhäuser, Boston, MA, 1998.
- [6] K. Gedalyahu, R. Tur, and Y. C. Eldar, “Multichannel sampling of pulse streams at the rate of innovation,” *submitted to IEEE Trans. on Signal Processing*, Apr. 2010.
- [7] R. Tur, Y. C. Eldar, and Z. Friedman, “Low rate sampling of pulse streams with application to ultrasound imaging,” *submitted to IEEE Trans. on Signal Processing*, Mar. 2010.
- [8] M. Vettereli, P. Marziliano, and T. Blu, “Sampling signals with finite rate of innovation,” *IEEE Trans. Signal Process.*, vol. 50, no. 6, pp. 1417–1428, June 2002.
- [9] M. Mishali, Y. C. Eldar, and A. Elron, “Xampling: Signal acquisition and processing in union of subspaces,” *CIT Report 747, EE Dept., Technion-Israel Institute of Technology*, vol. 1704, Oct. 2009.
- [10] M. Mishali, Y. C. Eldar, O. Dounaevsky, and E. Shoshan, “Xampling: Analog to digital at sub-Nyquist rates,” *to appear in IET Journal of Circuits, Devices and Systems*.
- [11] M. Mishali and Y. C. Eldar, “From theory to practice: Sub-Nyquist sampling of sparse wideband analog signals,” *IEEE Journal of Selected Topics on Signal Processing*, vol. 4, no. 2, pp. 375–391, Apr. 2010.
- [12] —, “Blind multi-band signal reconstruction: Compressed sensing for analog signals,” *IEEE Trans. Signal Processing*, vol. 57, no. 3, pp. 993–1009, Mar. 2009.
- [13] E. J. Candés, Y. C. Eldar, and D. Needell, “Compressed sensing with coherent and redundant dictionaries,” *submitted*, 2010.
- [14] J. A. Tropp, A. C. Gilbert, and M. J. Strauss, “Algorithms for simultaneous sparse approximation. Part I: Greedy pursuit,” *Signal Processing*, vol. 86, 2006.
- [15] J. A. Tropp, “Algorithms for simultaneous sparse approximation. Part II: Convex relaxation,” *Signal Processing*, vol. 86, 2006.
- [16] S. F. Cotter, B. D. Rao, K. Engan, and K. Kreutz-Delgado, “Sparse solutions to linear inverse problems with multiple measurement vectors,” *IEEE Trans. Signal Process.*, vol. 53, no. 7, pp. 2477–2488, June 2005.
- [17] M. Mishali and Y. C. Eldar, “Reduce and boost: recovering arbitrary sets of jointly sparse vectors,” *IEEE Trans. Signal Process.*, vol. 56, no. 10, pp. 4692–4702, Oct. 2008.
- [18] H. Feichtinger and G. Zimmermann, “A Banach space of test functions for Gabor analysis,” in *Gabor Analysis and Algorithms: Theory and Applications*, H. Feichtinger and T. Strohmer, Eds. Birkhäuser, Boston, MA, 1998.
- [19] I. Daubechies, A. Grossmann, and Y. Meyer, “Painless nonorthogonal expansions,” *J. Math. Phys.*, vol. 27, no. 5, 1986.
- [20] O. Christensen, “Pairs of dual Gabor frame generators with compact support and desired frequency localization,” *Applied and Computational Harmonic Analysis*, vol. 20, no. 3, 2006.
- [21] D. L. Donoho and M. Elad, “Maximal sparsity representation via ℓ_1 minimization,” *Proc. Nat. Acad. Sci.*, vol. 100, 2003.
- [22] J. Chen and X. Huo, “Theoretical results on sparse representations of multiple-measurement vectors,” *IEEE Trans. Signal Process.*, vol. 54, no. 12, pp. 4634–4643, Nov. 2006.
- [23] E. J. Candés, J. Romberg, and T. Tao, “Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information,” *IEEE Trans. Inform. Theory*, vol. 52, no. 2, pp. 489–509, Jan. 2006.
- [24] S. S. Chen, D. L. Donoho, and M. A. Saunders, “Atomic decomposition by basis pursuit,” *SIAM J. Scientific Computing*, vol. 20, no. 1, 1999.
- [25] M. E. Davies and Y. C. Eldar, “Rank awareness in joint sparse recovery,” *submitted to IEEE Trans. on Info. Theory*, Apr. 2010.

- [26] Y. C. Eldar and M. Mishali, "Robust recovery of signals from a structured union of subspaces," *IEEE Trans. Inform. Theory*, vol. 55, no. 11, pp. 5302–5316, Nov. 2009.
- [27] E. J. Candés, "The restricted isometry property and its implications for compressed sensing," *C. R. Acad. Sci. Paris, Ser. I*, vol. 346, 2008.
- [28] R. G. Baraniuk, M. Davenport, R. DeVore, and M. Wakin, "A simple proof of restricted isometry property for random matrices," *Const. Approx.*, 2007.
- [29] S. Mendelson, A. Pajor, and N. Tomczak-Jaegermann, "Uniform uncertainty principle for Bernoulli and subgaussian ensembles," *Constr. Approx.*, vol. 28, no. 3, 2008.
- [30] E. J. Candés and T. Tao, "Near optimal signal recovery from random projections: Universal encoding strategies?" *IEEE Inf. Theory*, vol. 52, no. 12, pp. 5406–5425, Nov. 2006.
- [31] M. Rudelson and R. Vershynin, "On sparse reconstruction from Fourier and Gaussian measurements," *Commun. Pure Appl. Math.*, vol. 61, no. 8, 2008.
- [32] Y. C. Eldar, "Compressed sensing of analog signals in shift-invariant spaces," *IEEE Trans. Signal Processing*, vol. 57, no. 8, pp. 2986–2997, Aug. 2009.
- [33] K. Gedalyahu and Y. C. Eldar, "Time-delay estimation from low-rate samples: A union of subspaces approach," *IEEE Trans. on Signal Processing*, vol. 58, no. 6, pp. 3017–3031, June 2010.
- [34] R. S. Laugesen, "Gabor dual spline windows," *Appl. Comput. Harmon. Anal.*, vol. 27, no. 2, 2009.
- [35] O. Christensen, H. O. Kim, and R. Y. Kim, "Gabor windows supported on $[-1, 1]$ and compactly supported dual windows," *Appl. Comp. Harm. Anal.*, vol. 28, 2010.
- [36] A. Ron and Z. Shen, "Frames and stable bases for shift-invariant subspaces of $L^2(\mathbb{R}^d)$," *Canad. J. Math*, vol. 47, no. 5, 1995.
- [37] V. D. Prete, "Estimates, decay properties, and computation of the dual function for Gabor frames," *J. Fourier Anal. Appl.*, vol. 5, no. 6, 1999.
- [38] K. Gröchenig, A. J. E. M. Janssen, N. Kaiblinger, and G. E. Pfander, "Note on B-splines, wavelet scaling functions, and Gabor frames," *IEEE Trans. Inform. Theory*, vol. 49, no. 12, pp. 3318–3320, Jan. 2003.
- [39] V. D. Prete, "On a necessary condition for B-spline Gabor frames," *Ricerche mat.*, vol. 59, 2010.