IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received July 9, 2020, accepted July 18, 2020, date of publication July 21, 2020, date of current version July 31, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3011032

Analog-to-Information Conversion for Nonstationary Signals

QIANG WANG, (Graduate Student Member, IEEE), CHEN MENG, AND CHENG WANG
Missil Engineering Department, Army Engineering University of PLA–Shijiazhuang, Shijiazhuang 050003, China

Corresponding author: Chen Meng (15231126568@163.com)

This work was supported by the National Natural Science Foundation of China under Grant 61501493.

ABSTRACT In this paper, we consider the problem of analog-to-information conversion for nonstationary signals, which exhibit time-varying properties with respect to spectral contents. Nowadays, sampling for nonstationary signals is mainly based on Nyquist sampling theorem or signal-dependent techniques. Unfortunately, in the context of the efficient 'blind' sampling, these methods are infeasible. To deal with this problem, we propose a novel analog-to-information conversion architecture to achieve the sub-Nyquist sampling for nonstationary signals. With the proposed scheme, we present a multi-channel sampling system to sample the signals in time-frequency domain. We analyze the sampling process and establish the reconstruction model for recovering the original signals. To guarantee the wide application, we establish the completeness under the frame theory. Besides, we provide the feasible approach to simplify the system construction. The reconstruction error for the proposed system is analyzed. We show that, with the consideration of noises and mismatch, the total error is bounded. The effectiveness of the proposed system is verified in the numerical experiments. It is shown that our proposed scheme outperforms the other sampling methods state-of-the-art.

INDEX TERMS Analog-to-information conversion, compressive sensing, nonstationary signals, sub-Nyquist, time-frequency.

I. INTRODUCTION

As the cornerstone of digital signal processing, Shannon–Nyquist–Whittaker sampling theorem states that, to perfectly reconstruct a band-limited analog signal, the sampling rate must be at least twice the highest frequency of signals [1], [2]. With the expansion of bandwidth of analog signals, the sampling method based on sampling theorem becomes impractical, since it is challenging to build sampling hardware that operates at a sufficient rate [3]. It is, however, well-known that signals with specific structure can be sufficiently sampled well-below the Nyquist rate. For example, the signals can be recovered from incomplete frequency samples-provided that the signals are sparse in Fourier Transform-by minimizing a convex function [4]. The sparse assumption is instructive, since most natural and man-made signals show sparsity in a particular transform domain. In 2006, the conception of incomplete sampling and sparse reconstruction is generalized as the well-known Compressive Sensing (CS) theory [5]–[7], which shows potential uses in a broad range of applications.

The associate editor coordinating the review of this manuscript and approving it for publication was Qiangqiang Yuan.

CS is a novel framework for signal processing. Under this scheme, compressive measurements are conducted to acquire 'just enough' samples that guarantee the perfect recovery of the signal of interest. In essence, CS fuses sampling and compression, instead of sampling signals at the Nyquist rate followed by conventional data compression. Then the original signals can be recovered accurately by exploring the sparsity in transform domain [8]. CS has the potential to acquire signals well-below the Nyquist rate, which may lead to significant reduction in the sampling costs, power consumption, and hardware requirement. As a consequence, CS is commonly believed to be a panacea for wideband signals to achieve the efficient sampling [9], [10].

A. CHALLENGES FOR NONSTATIONARY SIGNALS

Many signals in nature and man-made systems exhibit time-varying properties. Such waveforms are called time-varying or nonstationary signals. They are encountered in various areas such as audio signals, synthetic aperture radar, and machinery [9], [11], [12]. Recent advances dealing with nonstationary signals mainly focus on the time-frequency analysis methods, since the signals are intrinsically sparse on the time-frequency plane. The main goals of suitable methods

are to capture effective representations to describe the signal properties. Fruitful results have been made in this field, such results as Short-time Fourier Transform (STFT) [13], [14], Wavelet Transform (WT) [15], [16], and Wigner Distribution (WD) [17], [18].

Typically, time-frequency analysis for nonstationary signals is conducted by digital signal processing methods, through which the analog signals are firstly converted into digital fashion. Due to the time-varying spectral contents, most nonstationary signals show a wide spectrum (e.g., several GHz). In words, under the Nyquist sampling theorem, achieving such specifications with a single analog-to-digital converter (ADC) is an elusive goal for current semiconductor technology [19]. Nowadays, there are mainly two kinds of approaches to deal with this problem. One is under the framework of Nyquist sampling theorem to design the high-speed sampling system. One of the typical architectures is time-interleaving ADC [20]–[22]. It achieves the high sampling rate with several parallel channels. Then a low working rate is required in one channel. However, for this architecture, maintaining an accurate time shift in different channels is difficult to implement. Another train of thought tries to reduce the sampling rate with respect to a priori information, such as modulating rate or delay time [23]. For example, utilizing the information of modulating rate, Linear Frequency Modulated (LFM) signals can be sparsely represented in fractional Fourier transforms domain. Exploiting this sparsity, some works [24], [25] propose the low-rate sampling methods for LFM signals in fractional Fourier transforms domain. However, in the more general situation, a priori information is hard to obtain. That limits its further application.

B. ANALOG-TO-INFORMATION (AIC) SYSTEM

Considering the limitations of current sampling approaches, CS is viewed as an alternative technique for nonstationary signals to break through the bandwidth barrier [5]. Indeed, one of the main advantages of CS is that it enables the acquisition of larger bandwidth with relaxed sampling-rate requirements, thus enabling less expensive, faster, and potentially more energy-efficient solutions. Therefore, CS-based AIC systems arouse great interest in the literatures for such wideband applications [26]. The next paragraphs summarize the most prominent AIC architectures. For each of these architectures, we briefly discuss the pros and cons from standpoints of signal sampling and hardware design.

Random Demodulator (RD) is one of the typical AIC systems that realizes the compressive sampling in one single channel [27], [28]. RD multiplies the analog input signals by a (pseudo-)random sequence, and correspondingly, the basic tones are smeared across the entire spectrum. Then the samples are captured at a low rate by a low-pass anti-aliasing filter. Since the sampling rate is far below the Nyquist rate, this procedure is also called as sub-Nyquist sampling [29], [30]. For RD system, the (pseudo-)random sequence generator must still run at Nyquist rate. Meanwhile, modulating the signal with a (pseudo-)random sequence is only suitable for very specific signal classes, such as signals that are well-represented in a frequency domain. For nonstationary signals, however, the good representation in frequency cannot be guaranteed.

Non-uniform sampling (NUS) is one of the simplest instances of CS-based AIC systems [31]. In principle, NUS achieves sub-Nyquist sampling by using an irregularly spaced time intervals. This architecture is also a one-channel system, mainly consisting of a sample-and-hold (S&H) stage and an ADC. The whole sampling process is controlled by a non-uniform clock. NUS shows advantages in solving the problems of noise folding, aliasing, and limited flexibility [32], [33]. The main challenge of NUS lies in the acquisition of a wideband analog input signal, such as the nonstationary signals. Although the average sampling rate is reduced significantly, high-speed conversion, potentially reaching up the maximal input signal frequency, may still exist between some adjacent samples.

Modulated Wideband Converter (MWC) [34] is, in essence, an extension for RD system. It is comprised of a bank of modulators and low-pass filters. MWC modulates the analog signals by periodic waveforms, and then uses the low-pass filter to achieve the uniform sampling. During reconstruction, MWC establishes the relationship between samples and original signals in spectrum [35]–[37]. Similar with the RD, MWC still has the (pseudo-)random sequence generator running at Nyquist rate. The reconstruction for the wideband signals is firstly conducted in frequency domain. In words, a sparse spectrum is needed for the accurate reconstruction. For the nonstationary signals, however, the spectrum is usually not sparse.

C. CONTRIBUTIONS

As analyzed above, a simple and efficient AIC system for nonstationary signals is still lacking. To fill this gap, in this paper, we address the construction of feasible AIC system to achieve the effective sub-Nyquist sampling and reconstruction for nonstationary signals.

The main contributions are as follows:

1) We propose a novel AIC system for nonstationary signals exploring the sparsity in time-frequency. Under the proposed scheme, the sub-Nyquist samples are obtained by a multi-channel system, which mainly consists of low-pass filters, integrators, and ADCs.

2) We further explore the system properties to achieve the simplification in construction. We propose a simplified scheme for the AIC system, where the same samples can be obtained with reduced inner-channels and (pseudo-)random sequence generators.

3) Under the frame theory, we construct a frame with variable window functions and irregular lattices. We establish the completeness for the system to make sure that the time-frequency sampling is reversible.

4) We analyze the reconstruction model for the AIC system to guarantee the accurate reconstruction. Considering the

noises and mismatch, we establish the upper bound for the total reconstruction error.

It is important to note that the proposed AIC system will make a great sense in some specific applications, such as the uncooperative receiving in cognitive radio system. Uncooperative receiving for wireless communication signals is one of the key techniques in spectrum sensing. It requires us to capture the dynamic spectrum with less user information. With the proposed AIC system, we reduce the sampling rate and sample number by exploiting the CS theory, such that lower working-band and smaller storage is needed for the receiver. The reconstruction process is 'blind' without using a priori information about users. That makes it more suitable for an uncooperative receiver.

D. ORGANIZATION

The remainder of this paper is organized as follows. In Section II, we present the notations and problem formulation. Section III describes some bases for CS theory. In Section IV, the proposed AIC system is presented, and then is further simplified. In Section V, we establish the reconstruction model for AIC system. The completeness is also established based on the frame theory. In Section VI, numerical experiments are conducted to evaluate the effectiveness of the proposed system. Section VII discusses the related works and conclusions about this paper.

II. NOTATION AND PROBLEM FORMULATION

A. NOTATION

Through this paper, we denote matrices and vectors by bold characters, with italic lowercase letters corresponding to vectors and uppercase letters to matrices. The symbols $\mathbb{R}$ and $\mathbb{Z}$ represent the real number field and integral number field respectively. For the square integrable space, we use $L_2(\cdot)$. Symbol supp$(\cdot)$ denotes the support width, $|\cdot|$ is the absolute value of the element, and $\|\cdot\|_2$ represents the l_2 norm. The n-th entry in a vector $\boldsymbol{a}$ is written as $\boldsymbol{a}_n$, whereas $\mathbf{A}_{ij}$ denotes the ij-th entry in matrix $\mathbf{A}$. Superscript $(\cdot)^*$ represents complex conjugation.

B. PROMLEM FORMULATION

Fourier Transform (FT) maps the signal $x(t)$ into frequency domain $X(f)$, which can be depicted as

$$X(f) = \int_{-\infty}^{+\infty} x(t)e^{-j2\pi ft}dt \tag{1}$$

As the main tool for signal analysis and processing, FT is widely used to reveal the essential characteristics of signals. It is, however, limited in the application of nonstationary signal analysis, since FT provides no easily understood timing details about the occurrence of various frequency components. To alleviate this problem, we try to analyze the properties of nonstationary signals using time-frequency analysis methods.

Time-frequency analysis is an effective tool to characterize nonstationary signals [11]–[15]. It reflects the time-varying properties by mapping signals into the joint time-frequency domain. In the literature, time-frequency analysis methods can be divided into two categories: quadratic (bilinear) time-frequency distribution (QTFD) and linear time-frequency transform (LTFT). Compared with QTFD, LTFT has the advantages of simple computation and perfect reconstruction formula, which make it more suitable for the application in AIC system.

Gabor transform (GT) is one of the most widely used LTFT methods in signal processing. It is a class of STFT that employs discrete sampling lattices in time-frequency plane. Given the window function $g(t)$, Gabor coefficient at lattice (ak, bl) can be depicted as

$$G_g x(k,l) := \langle x, M_{bl}T_{ak}g\rangle = \int_{-\infty}^{+\infty} x(t)g(t-ak)^* e^{-j2\pi blt}dt \tag{2}$$

where $M_{bl}T_{ak}g = g(t-ak)e^{j2\pi blt}$ is the time-frequency shifted window function, a and b are shifting intervals in time and frequency respectively.

If the collection $\mathcal{G}(g,a,b) = \{M_{bl}T_{ak}g(t)\}_{k,l\in\mathbb{Z}}$ constructs a Gabor frame for $L_2(\mathbb{R})$, there exists the dual window $\gamma(t)$ such that

$$x(t) = \sum_{k,l\in\mathbb{Z}} G_g x(k,l)\cdot M_{bl}T_{ak}\gamma(t) \tag{3}$$

In words, with sufficient Gabor coefficients, the original nonstationary signals can be recovered perfectly. Meanwhile, considering the inner product in (2), we may obtain the Gabor coefficients through a bank of filters. That makes it more possible to implement the GT in analog circuit.

To illustrate the time-varying properties for nonstationary signals, we conduct the GT for two typical nonstationary signals: LFM and Hopping Frequency (HF) signals. For comparison, spectrum based on FT is also presented. The results are shown in Figure 1 and Figure 2. It is seen that, compared with FT, GT achieves an effective representation for LFM and HF signals in the joint time-frequency domain.

On the other hand, it is seen that the time-varying properties of nonstationary signals bring an inherent sparsity in time-frequency domain. Although only LFM and HF signals are presented in this paper, it is important to note that the sparsity in time-frequency domain is commonly shared by many nonstationary signals [11], [12]. That gives us the inspiration to sample the nonstationary signals in time-frequency domain. Therefore, in this paper, we try to utilize this sparsity to achieve the efficient sub-Nyquist sampling. We wish to design a sampling system for nonstationary signals satisfying the followings

1) The sampling rate should be as low as possible;

2) The system uses less a priori information during sampling and reconstruction;

3) The sample number is supposed to be as small as possible;

4) The sampling system is simple enough to be implemented with existing analog devices and ADCs.

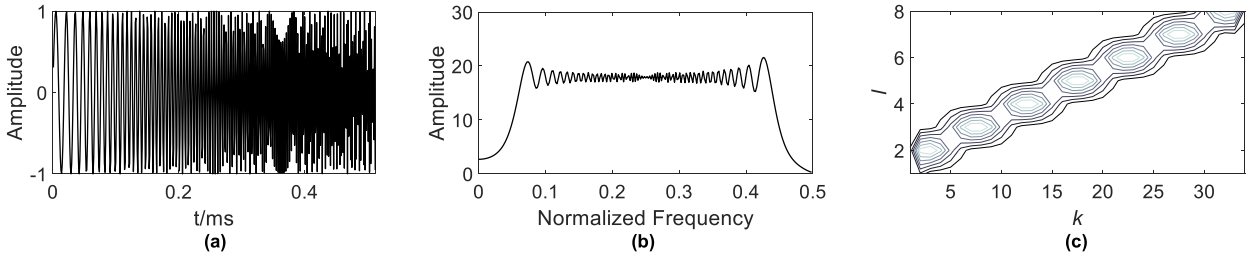

FIGURE 1. LFM signals: (a) waveform (b) spectrum (c) time-frequency representation.

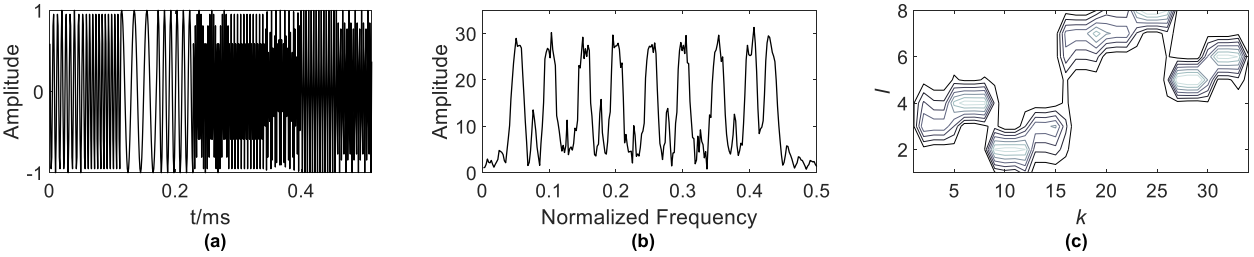

FIGURE 2. HF signals: (a) waveform (b) spectrum (c) time-frequency representation.

III. COMPRESSIVE SENSING

In this paper, we try to achieve the efficient sampling for nonstationary signals by exploring the CS theory. Before bringing the proposed system into practice, we present some bases concerning the CS theory in this section.

Let $\boldsymbol{x} \in \mathbb{R}^N$ be the discrete-time, N-dimensional real-valued signal vector that we wish to acquire. Under the theory of CS, the signals are supposed to be compressible, which means that the signals are sparse in a certain transform domain. Assume that the transform domain is $\boldsymbol{\Psi}$, and then $\boldsymbol{x}$ can be expressed as

$$\boldsymbol{x} = \boldsymbol{\Psi}\boldsymbol{s} = \boldsymbol{s}_1\boldsymbol{\Psi}_1 + \boldsymbol{s}_2\boldsymbol{\Psi}_2 + \ldots \boldsymbol{s}_N\boldsymbol{\Psi}_N \tag{4}$$

where $\boldsymbol{s}$ is sparse vector, $\boldsymbol{s}_i$ $(i = 1\ 2\ \ldots\ N)$ is i-th coefficient in $\boldsymbol{s}$, $\boldsymbol{\Psi}_i (i = 1, 2, \ldots, N)$ is the i-th column in $\boldsymbol{\Psi}$. The signal is so-called S-sparse if there are S non-zero entries in $\boldsymbol{s}$.

In practice, sparsity can be a strong constraint to impose, and we may prefer the weaker concept of compressibility, which is measured by the error of best S-term approximation [5]. For $p > 0$, the l_p-error of best S-term approximation is defined by

$$\sigma_S(\boldsymbol{s})_p := \inf \left\{ \|\boldsymbol{s} - \boldsymbol{z}\|_p , \boldsymbol{z} \in \mathbb{R}^N \ is\ S\text{-}sparse \right\} \tag{5}$$

where nonzero entries in $\boldsymbol{z}$ equal the S largest absolute entries of $\boldsymbol{s}$. We call $\boldsymbol{s}$ is compressible if $\sigma_S(\boldsymbol{s})_p$ decays quickly in S.

For the sparse and compressible signals, compressive sampling can be achieve by a measurement process [5], [7], which can be depicted as

$$\boldsymbol{y} = \boldsymbol{\Phi}\boldsymbol{x} = \boldsymbol{\Phi}\boldsymbol{\Psi}\boldsymbol{s} = \boldsymbol{\Theta}\boldsymbol{s} \tag{6}$$

where $\boldsymbol{\Phi} \in \mathbb{R}^{M\times N}$ is the measurement matrix, $\boldsymbol{\Theta} = \boldsymbol{\Phi}\boldsymbol{\Psi}$ is the sensing matrix, $\boldsymbol{y} \in \mathbb{R}^M$ is the sample vector. If the signal is sparse in time, we have $\boldsymbol{\Psi} = \mathbf{I}$, where $\mathbf{I}$ is identity matrix.

By the measurement matrix $\boldsymbol{\Phi}$, CS achieves the acquisition for signals with a far fewer samples. With respect to the sparsity, we usually have $M \ll N$, which implies that only a small number of samples are generated after sampling.

To recover the original signal, the vector $\boldsymbol{y}$ is supposed to contain all the information concerning $\boldsymbol{x}$. This requirement is guaranteed by the property of measurement matrix, which is named as Restricted Isometry Property(RIP) [39]. To be specific, the sensing matrix $\boldsymbol{\Theta}$ is supposed to satisfy

$$(1 - \delta_S) \|\boldsymbol{s}\|_2^2 \leq \|\boldsymbol{\Theta}_I \boldsymbol{s}\|_2^2 \leq (1 + \delta_S) \|\boldsymbol{s}\|_2^2 \tag{7}$$

where δ_S is the Restricted Isometry Constraint(RIC), $0 < \delta_S < 1$. I is index set with $I \subset \{1, 2, \ldots, N\}$, $\boldsymbol{\Theta}_I$ is the matrix formed by choosing the columns of $\boldsymbol{\Theta}$ whose indices are in I. To guarantee the RIP at the order of S, I is supposed to be any possible index set as $I \subset \{1, 2, \ldots, N\}$, $|I| \leq S$.

With the sensing matrix $\boldsymbol{\Theta}$ satisfying the RIP, it is possible to recover the original signal with sufficient samples, typically scaling as $M \geq \mathrm{O}(S \times \log(N/S))$. This process is usually conducted by a sparse recovery algorithm that achieves robust estimates for the sparse vector $\boldsymbol{s}$, and hence, enables the recovery for the signal $\boldsymbol{x}$. For more details on CS, please refer to [40], [41].

IV. AIC SYSTEM FOR NONSTATIONARY SIGNALS

Under the CS theory, we design a novel AIC sampling system for nonstationary signals. The system is schematically drawn in Figure 3. Corresponding parameters are shown in Tabl. 1.

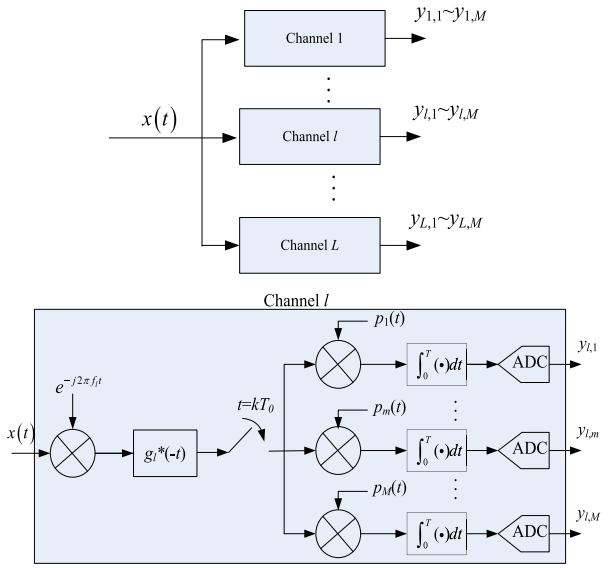

FIGURE 3. The proposed CS-based AIC system.

TABLE 1. Parameters for AIC system.

Symbol	Meaning
L	number of channels
M	number of inner-channels
T	time of integration
T_0	time-interval between filter outputs
f_l	modulating frequency in l-th channel

In this section, we provide description and analysis for the proposed AIC system.

A. SYSTEM DESCRIPTION

Our system achieves the sub-Nyquist sampling for nonstationary signals by exploring the sparsity in time-frequency domain. The sampling process is conducted in several channels, implementing different modulation, so that a sufficiently large number of channels guarantee the accurate reconstruction for original nonstationary signals.

The proposed AIC systems can be divided into two parts. One is used to map the nonstationary signals into the time-frequency domain. The other is used to achieve the undersampling for time-frequency coefficients. In the first part, the input signals enter L channels simultaneously. In the l-th channel, the signals are firstly modulated by the function $e^{-j2\pi f_l t}$. Then a low-pass filter is used to prevent aliasing. The filter output is sampled at the rate $1/T_0$. Under the conception of linear time-frequency transform, the sampling period T_0 is the shifting interval in time. Note that, different from the conventional discretization process, here we use the sampler to achieve the sampling in time, without quantification. This is mainly because that the filter output is further processed in analog circuit. The impulse response for low-pass filter is $g_l^*(-t)$, which is conjugated and time-reversed window function $g(t)$. Combined with the modulating function $e^{-j2\pi f_l t}$, the filter, in essence, maps the nonstationary signals into the time-frequency domain. The definite relationship between time-frequency coefficient and filter output is presented in Section IV.B.

The second part is used to achieve the undersampling for time-frequency coefficients, which is essentially the implementation of CS theory. In channel l, the undersampling process is conducted in different inner-channels. With M inner-channels, the filter output is multiplied by M mixing functions $p_m(t)$. As depicted in Figure 4, the mixing function is generated by the (pseudo-) random sequence generator, which produces a discrete (pseudo-) random sequence $\varepsilon_{m,0}, \varepsilon_{m,1}, \varepsilon_{m,2}, \ldots$ of numbers that take values $\pm 1/\sqrt{K}$ with equal probability, to yield

$$\varepsilon_{m,k} = \begin{cases} +1/\sqrt{K}, & P = 0.5 \\ -1/\sqrt{K}, & P = 0.5 \end{cases} \tag{8}$$

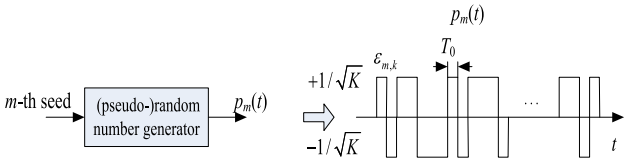

FIGURE 4. The generation of mixing function $p_m(t)$.

Then the (pseudo-)random sequence is used to create the continuous-time mixing function $p_m(t)$ via the equation

$$p_m(t) = \varepsilon_{m,k}, \quad t \in \left[k\frac{T}{K}, (k+1)\frac{T}{K}\right), \ 0 \le k \le K-1 \tag{9}$$

where $T = KT_0$. Under the frame theory, parameters K and L are supposed to make a dense division for a given time-frequency interval. The theoretical bound for K and L will be discussed in Section V B. It is important to note that, although the filter output is discrete in time, we still use a mixing function continuous in time. Referring to some current CS-based AIC systems (such as RD and MWC), continuous-time mixing function is available with simple circuit components. Then, to have the discrete mixing function, we may need M more samplers in each channel that work in coordination with the sampler for the filter. That will obviously increase the system complexity. Meanwhile, considering the practical application, the non-ideal sampler will inevitably interfere with the sampling process. Using more samplers will generate stronger interference. Therefore, in this paper, we just use one sampler in each channel to sample the filter output. With the deduction for system output, we will show that it is enough to achieve the aim of compressive measurement for time-frequency coefficients.

Actually, (pseudo-)random sequence is widely used in some current CS-based AIC systems. From a hardware perspective, the (pseudo-)random sequence generator used in these systems must still support the working bandwidth up to Nyquist rate, which makes it difficult to be implemented

for wideband signals. In this paper, we use the (pseudo-)random sequence generator operating at the rate $1/T_0$. Under the theory of frame work, time-shifting parameter T_0 can be much larger than the Nyquist period T_{Nyq}. That means, with the proposed system, the working bandwidth of the (pseudo-)random sequence generator is reduced significantly. That will make it easier to be implemented in application.

After multiplying with the mixing function, a discrete system output is obtained by an integrator and ADC, working at the rate $f_s = 1/T$. This working rate matches the sampling rate of the system, since both the integration and quantification are conducted at the time period T.

Here, we make an overall description for the proposed AIC system. The system has the sampling rate far below the Nyquist rate ($f_s = 1/T$, $T = KT_0 \gg T_{Nyq}$). In practice, this sampling rate allows flexible choice of an ADC from a variety of commercial devices in the low rate regime. The total sample number is LM, where M is determined by the sparsity of signals. Since the nonstationary signals have good sparsity in time-frequency domain, we can use few samples to achieve the accurate reconstruction. The sampling process uses less a priori information, such that it has wide applicability. Besides, some other advantages in practical implementation are shown as follows.

1) The analog mixer used in the proposed AIC system is a provable technology in wideband regime. In words, it is easy to design the mixer adapted to the working bandwidth of input signals.

2) The special designing of continuous-time random mixing function allows us to use the shift register, which is able to work at the rate up to 80 GHz. Then the complexity of system construction is further simplified.

3) Although the sampling process in different channels is conducted simultaneously, there is not a strict requirement for time synchronization. In words, we do not need to consider the problem of timing jitter for practical designing.

Note that the proposed AIC system is an analog front-end, where input signals, mixers, and filters work in analog forms. The system output is digital and discrete, such that it is convenient for further processing in digital units. Meanwhile, the system is only used for sampling, since the reconstruction is carried out in back-end server. That will speed up the computation and make full use of computing resources.

B. SYSTEM ANALYSIS

Based on the proposed AIC system, we make an analysis for the whole sampling process. As shown in Section II.B, the sparsity in time-frequency is the inherent property for nonstationary signals. This is also the inspiration for us to design the AIC system. Here, we reveal the relationship between time-frequency coefficients and the system output.

As analyzed above, the impulse response for low-pass filter is conjugated and time-reversed window function $g_l(t)$, which can be expressed as $g_l{}^*(-t)$. Then the filter output in the l-th channel can be expressed as

$$\begin{aligned} f_l(t) &= \int_{-\infty}^{+\infty} x(t')e^{-j2\pi f_l t'} \cdot g_l{}^*(t'-t)dt' \\ &= \int_{-\infty}^{+\infty} x(t') \cdot g_l{}^*(t'-t)e^{-j2\pi f_l t'}dt' \end{aligned} \tag{10}$$

Followed by the sampler operating at the rate $1/T_0$, the filter output is sampled in time. Since the filter output is further processed in analog circuit, for convenience, we introduce the Dirac delta function $\delta(t)$ to depict the sampled filter output $f_l^d(t)$, to yield

$$f_l^d(t) = \sum_{k=0}^{K-1} f_l(t)\delta(t - kT_0) \tag{11}$$

For $t = kT_0$, we have

$$f_l^d(kT_0) = \int_{-\infty}^{+\infty} x(t') \cdot g_l{}^*(t' - kT_0)e^{-j2\pi f_l t'}dt' \tag{12}$$

It is seen from (12) that $f_l^d(kT_0)$ is, in essence, the time-frequency coefficient at the lattice (kT_0, f_l). In words, by the modulating function $e^{-j2\pi f_l t}$ and low-pass filter $g_l{}^*(-t)$, we realize the time-frequency representation for nonstationary signals. For convenience, in the following content, we use the symbol $f_{l,k}$ to represent the filter output $f_l^d(t)$ at time $t = kT_0$.

Under the conception of time-frequency analysis, the effectiveness of time-frequency representation depends heavily on window function $g_l(t)$. The designing of window function aims to have a good time-frequency resolution. To have a good time resolution, $g_l(t)$ is required to be short, whereas, to have a good frequency resolution, $g_l(t)$ is supposed to be wide. In words, it is contradictory to improve the time and frequency resolutions simultaneously. The common countermeasure is utilizing the Gaussian function to create a $g_l(t)$ with a minimum product of time and frequency widths [13], [42], such that

$$\Delta t \cdot \Delta f = \frac{1}{4\pi} \tag{13}$$

where

$$(\Delta t)^2 = \frac{\int_{-\infty}^{+\infty} t^2 |g_l(t)|^2 dt}{\int_{-\infty}^{+\infty} |g_l(t)|^2 dt}, \quad (\Delta f)^2 = \frac{\int_{-\infty}^{+\infty} f^2 |G_l(f)|^2 df}{\int_{-\infty}^{+\infty} |G_l(f)|^2 df} \tag{14}$$

And $G_l(f)$ is the FT for $g_l(t)$. With the Gaussian function, $g_l(t)$ can be given by

$$g_l(t) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{t^2}{2\sigma^2}} \tag{15}$$

where σ^2 is the variance for Gaussian window.

It is important to note that we may use different window functions with variable window widths in different channels. It will make sense in some special applications. For example, when there is a transient component involved in the nonstationary signals, it is more suitable to use a

frequency-dependent window function to capture enough information about the transient property [43]. Then in the l-th channel, window function $g_l(t)$ can be expressed as

$$g_l(t) = \frac{|f_l|}{\sqrt{2\pi}} e^{-\frac{f_l^2 t^2}{2}} \tag{16}$$

In words, the time-frequency transform in this paper, is different from some conventional LTFT methods, such as STFT or GT. Therefore, to guarantee the effectiveness of the time-frequency transform, we establish the completeness of time-frequency representation using frequency-dependent window function. The details are shown in Section V B.

In the l-th channel, M samples are obtained by M inner-channels. The m-th sample in l-th channel can be described as

$$y_{l,m} = \int_0^T f_l^d(t) \cdot p_m(t)\, dt \tag{17}$$

With the expression of $p_m(t)$ and $f_l^d(t)$ given in (9) and (11) respectively, $y_{l,m}$ can be rewritten as

$$y_{l,m} = \sum_{k=0}^{K-1} f_{l,k} \varepsilon_{m,k} \tag{18}$$

where $M \ll K$. That means $y_{l,m}$ is the linear measurement for the time-frequency coefficients. As analyzed above, nonstationary signals show a good sparsity in the time-frequency domain, such that most coefficients in $\{f_{l,k}\}_{k=0,1,\ldots,K-1,l=1,2,\ldots,L}$ are zero or small enough. That makes it possible to recover the time-frequency coefficients from the undersampled measurements. To guarantee the successful reconstruction of $f_{l,k}$, a sufficiently large M is required. Under the CS theory, the theoretical bound for M is given by $M \geq \mathrm{O}(S \times \log(K/S))$, where S is the sparsity for the set $\{f_{l,k}\}_{k=0,1,\ldots,K-1}$.

C. SYSTEM SIMPLIFICATION

As depicted in Figure 3, the proposed AIC system acquires M samples in M inner-channels. This designing, however, has a drawback of low utilization for both inner-channel and (pseudo-)random sequence generator. To deal with this problem, we present a simplified sampling scheme.

The simplified AIC system is achieved by exploring the properties of low-pass filter and mixing function. Assume that signal $x(t)$ is compactly supported on the interval $[0, T_w]$, and $g_l(t)$ is limited to $[-T_l/2, T_l/2]$, satisfying

$$T_l \leq 2KT_0 - 2T_w \tag{19}$$

Then we design the new filter with the impulse response expressed as

$$\tilde{g}_l{}^*(-t) = \sum_{m=1}^{M} g_l{}^*(-t + (m-1)T) \tag{20}$$

where $KT_0 = T$. The filter output is pseudo-periodic, which can be expressed as

$$f_l[t+(m-1)T] = \int_{-\infty}^{+\infty} x(t') \cdot \tilde{g}_l{}^*[t'-t-(m-1)T] e^{-j2\pi f_l t'} dt'$$

$$= \sum_{m'=1}^{M} \int_{-\infty}^{+\infty} x(t') \cdot g_l{}^*[t' + (m-1)T - t - (m'-1)T] e^{-j2\pi f_l t'} dt' \tag{21}$$

where $1 \leq m \leq M$. Considering the support time for $x(t)$ and $g_l{}^*(t)$, $f_l[t + (m-1)T]$ is non-zero only when $m' = m$, to yield

$$f_l[t+(m-1)T] = \int_{-\infty}^{+\infty} x(t') \cdot g_l{}^*(t'-t) e^{-j2\pi f_l t'} dt' \tag{22}$$

That means $f_l[t + (m-1)T] = f_l(t), 1 \leq m \leq M$. Meanwhile, we extend the mixing function as

$$p'(t) = \sum_{m=1}^{M} p_m[t - (m-1)T] \tag{23}$$

Correspondingly, the integration and system construction is modified. For the l-th channel, the simplified AIC system can be depicted in Figure 5.

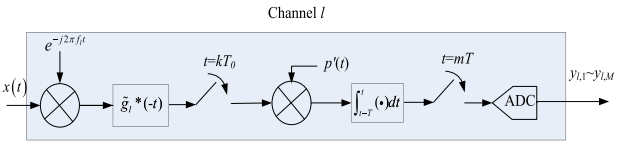

FIGURE 5. The simplified AIC system.

The system output at the time $t = mT$ can be expressed as

$$y_l(mT) = \int_{t-T}^{t} f_l^d(t) \cdot p'(t) dt = \int_{(m-1)T}^{mT} f_l^d(t) \cdot p'(t) dt$$

$$= \int_0^T f_l^d[t + (m-1)T] \cdot p'[t + (m-1)T] dt \tag{24}$$

Since $f_l[t + (m-1)T] = f_l(t)$ $(1 \leq m \leq M)$, with the sampler operating at rate $1/T_0$, we have $f_l^d[t + (m-1)T] = f_l^d(t)$, to yield

$$\int_0^T f_l^d[t + (m-1)T] \cdot p'(t + (m-1)T) dt$$

$$= \sum_{m'=1}^{M} \int_0^T f_l^d(t) \cdot p_m[t + (m-1)T - (m'-1)T] dt$$

$$= \int_0^T f_l^d(t) \cdot p_m(t) dt = y_{l,m} \tag{25}$$

That means $y_l(mT) = y_{l,m}$. In words, in the l-th channel, the same samples can be obtained by the simplified AIC system.

It is seen that the simplification is achieved by incorporating M inner-channels. Therefore, M samples are obtained in one channel. As analyzed in (25), the simplification does not change the sampling rate and system output. Moreover, the (pseudo-)random sequence generators are also reduced. As a result, to complete the whole sampling process, only L channels and (pseudo-)random sequence generators are needed.

V. RECONSTRUCTION FOR NONSTATIONARY SIGNALS

After sub-Nyquist sampling for nonstationary signals, another task is to recover the original signals. Based on the CS theory, we prefer to apply a 'blind' process to achieve the accurate reconstruction.

A. RECONSTRUCTION MODEL

The proposed AIC system maps the time-frequency coefficients into another low-dimensional space. That will make sense for data saving and transport. However, it also brings troubles in reconstruction. In the l-th channel, we define that

$$\boldsymbol{y}_l = [y_{l,1}, y_{l,2}, \ldots, y_{l,M}]^T \tag{26}$$

$$\boldsymbol{\Phi} = \begin{pmatrix} \varepsilon_{1,0} & \cdots & \varepsilon_{1,K-1} \\ \vdots & \ddots & \vdots \\ \varepsilon_{M,0} & \cdots & \varepsilon_{M,K-1} \end{pmatrix} \tag{27}$$

$$\boldsymbol{f}_l = [f_{l,0}, f_{l,1}, \ldots f_{l,K-1}]^T \tag{28}$$

The equation (19) can be rewritten as

$$\boldsymbol{y}_l = \boldsymbol{\Phi}\boldsymbol{f}_l \tag{29}$$

where $\boldsymbol{y}_l \in \mathbb{R}^M$, $\boldsymbol{\Phi} \in \mathbb{R}^{M\times K}$, $\boldsymbol{f}_l \in \mathbb{R}^K$. Then the reconstruction is to recover the vector $\boldsymbol{f}_l$ from the sample $\boldsymbol{y}_l$. Since $M < N$, the reconstruction model is ill-conditioned. However, when the vector $\boldsymbol{f}_l$ is sparse, we can solve this problem by optimization algorithm.

Under the theory of CS, the successful reconstruction for the sparse vector $\boldsymbol{f}_l$ requires the matrix $\boldsymbol{\Phi}$ to satisfy the RIP. From the expression of (8) and (27), we can find that the matrix used in this paper is a kind of random measurement matrix, which is named as Bernoulli random matrix. The RIP of Bernoulli random matrix has been studied in some previous works [5], [6]. The main results are presented in **Lemma I.**

Lemma I: Let $\boldsymbol{\Phi}$ be a random $M \times K$ matrix whose entries $\varepsilon_{m,k}$ are drawn according to the distribution in (8). Given the sparsity S, $\delta_{3S} \in (0, 1)$, $c_0 > 0$ and measurement number $M \geq \mathrm{O}(S \times \log(K/S))$, there exist constant $c_1 > 0$ depending on the δ_{3S} and $\boldsymbol{\Phi}$ satisfying $3S$-order RIP with probability

$$\geq 1 - 2e^{-c_1 M} \tag{30}$$

where

$$c_1 \leq \frac{3\delta_{3S}^2 - \delta_{3S}^3}{48} - c_0[1 + (1 + \log(\frac{12}{\delta_{3S}}))/\log(\frac{K}{S})] \tag{31}$$

Then with the Bernoulli random matrix $\boldsymbol{\Phi}$, it is possible to recover the sparse vector $\boldsymbol{f}_l$ from the undersampled measurements. The reconstruction model [8] can be expressed as

$$\hat{\boldsymbol{f}}_l = \arg\min \left\|\boldsymbol{f}_l\right\|_0 \quad subject\ to\ \boldsymbol{y}_l = \boldsymbol{\Phi}\boldsymbol{f}_l \tag{32}$$

In words, we try to seek the sparsest vector that satisfies the condition (29). The introduction of l_0 norm gives an unfavorable complexity in computation. To deal with this challenge, a fundamental method is relaxing the l_0 norm to l_1 norm [4], to yield

$$\hat{\boldsymbol{f}}_l = \arg\min \left\|\boldsymbol{f}_l\right\|_1 \quad subject\ to\ \boldsymbol{y}_l = \boldsymbol{\Phi}\boldsymbol{f}_l \tag{33}$$

Both (32) and (33) are basic reconstruction models in CS theory, which can be solved by some optimization algorithms such as Sparse Bayesian Learning (SBL) [44], or Iteratively Reweighted Least Squares (IRLS) [45].

As analyzed above, the good sparsity of $\boldsymbol{f}_l$ is one of the key points to guarantee the successful reconstruction from undersampled measurements. In this paper, the sparsity for $\boldsymbol{f}_l$ is guaranteed by the time-varying properties, which are commonly shared by many nonstationary signals. However, it is also important to note that the proposed AIC system may be unsatisfactory in some cases. Firstly, multiple components in nonstationary signals will increase the sparsity in time-frequency, and inevitably, increase the sample number. The proposed system will be meaningless if the sampling number is too large due to the poor sparsity caused by multiple components. Secondly, there will be obvious reconstruction errors if the signals contain stationary components. From the construction of $\boldsymbol{f}_l$, we know that $\boldsymbol{f}_l$ will be non-sparse if the nonstationary signals contains stationary frequency component f_l. That will cause failure for the reconstruction of $\boldsymbol{f}_l$.

B. COMPLETENESS

After reconstructing all vectors $\{\boldsymbol{f}_l\}_{l=1,2,\ldots,L}$, another main task is to reconstruct the original signals. Since the entries in $\boldsymbol{f}_l$ are the time-frequency coefficients, it is a natural choice to introduce the frame theory to provide theoretical support for the reconstruction process [12], [50].

For convenience, we rewrite the expression of time-frequency coefficient $f_{l,k}$ as

$$\begin{aligned} f_{l,k} &= \int_{-\infty}^{+\infty} x(t') \cdot {g_l}^*(t' - kT_0) e^{-j2\pi f_l t'} dt' \\ &= \left\langle x(t), g_l(t - kT_0) e^{j2\pi f_l t} \right\rangle \end{aligned} \tag{34}$$

where the collection $\left\{g_l(t - kT_0)e^{j2\pi f_l t}\right\}_{l,k\in\mathbb{Z}}$ is the basis for time-frequency transform. We introduce the translation and modulation operators as

$$T_x g(t) := g(t - x) \tag{35}$$

$$M_\omega g(t) := g(t) e^{j2\pi\omega t} \tag{36}$$

Then the collection can be simplified as $\{M_{f_l} T_{kT_0} g_l(t)\}_{l,k\in\mathbb{Z}}$. To guarantee the effectiveness of the proposed AIC system, we should make sure that for any nonstationary signal $x(t) \in L_2(\mathbb{R})$, the coefficient collection $\left\{f_{l,k}\right\}_{l,k\in\mathbb{Z}}$ is sufficient to represent the signals. Such that the basis $\left\{M_{f_l} T_{kT_0} g_l(t)\right\}_{l,k\in\mathbb{Z}}$ should be complete to recover the nonstationary signal $x(t)$. Therefore, we introduce the frame theory to establish completeness. The definition for a frame is shown in **Definition I**.

Definition I: A collection $\left\{M_{f_l} T_{kT_0} g_l(t)\right\}_{l,k\in\mathbb{Z}}$ is a frame in $L_2(\mathbb{R})$ if there exist constants $0 < A \leq B < \infty$, such that for any $x(t) \in L_2(\mathbb{R})$ we have

$$A \left\|x\right\|^2 \leq \sum_{k,l\in\mathbb{Z}} \left|\left\langle x, M_{f_l} T_{kT_0} g_l \right\rangle\right|^2 \leq B \left\|x\right\|^2 \tag{37}$$

where A and B are frame bounds.

If $\{M_{f_l} T_{kT_0} g_l(t)\}_{l,k\in\mathbb{Z}}$ constructs a frame, there exists dual window $\gamma_l(t)$ such that

$$x(t) = \sum_{k,l\in\mathbb{Z}} f_{l,k}\gamma_l(t - kT_0)e^{j2\pi f_l t} \tag{38}$$

where $\gamma_l(t) = S^{-1}g_l(t)$. S is the frame operator defined by

$$Sx = \sum_{k,l\in\mathbb{Z}} \left\langle x(t), g_l(t - kT_0)e^{j2\pi f_l t}\right\rangle g_l(t - kT_0)e^{j2\pi f_l t} \tag{39}$$

Under the scheme of frame theory, completeness is highly impacted by the density of the lattice (kT_0, f_l). In this paper, the density is defined by $d = 1/(T_0 f_0)$, where $f_0 = \max_l |f_{l+1} - f_l|$, $l \in \mathbb{Z}$. Then the time-frequency representation, also called time-frequency sampling, can be divided into three categories as:

- Undersampling—$d < 1$.
- Critical sampling—$d = 1$.
- Oversampling—$d > 1$.

In the case of undersampling, the representation is proven to be incomplete. So the necessary condition for completeness is $T_o f_0 \leq 1$, where critical sampling occurs when $T_o f_0 = 1$.

It is presented in Section II that the collection of window functions with invariant shifts in time and frequency is able to construct a Gabor frame. However, in this paper, the collection $\{M_{f_l} T_{kT_0} g_l(t)\}_{l,k\in\mathbb{Z}}$ is not the typical shift-invariant system, since we may use variable window functions and irregular lattices in some cases. Considering the differences above, we present the proof of the existence of the frame with variable window functions and irregular lattices. The result is shown in **Theorem I.**

Theorem I: Assume that $g_l(t) \in L_2(\mathbb{R})$ is compactly supported on both time and frequency. There exist $0 < T_0, 0 < f_l$, such that

1) $C \leq \sum_{l=1}^{L} |G_l(f - f_l)|^2 \leq D$
2) $\max_l |\text{supp}(G_l(f))| \leq \frac{1}{T_0}$

Then for any $x(t) \in L_2(\mathbb{R})$, we have

$$\frac{C}{T_0}\|x\|^2 \leq \sum_{k,l\in\mathbb{Z}} \left|\langle x, M_{f_l} T_{kT_0} g_l\rangle\right|^2 \leq \frac{D}{T_0}\|x\|^2 \tag{40}$$

In words, $\{M_{f_l} T_{kT_0} g_l(t)\}_{l,k\in\mathbb{Z}}$ is a frame with the bounds $\frac{C}{T_0}$ and $\frac{D}{T_0}$.

Proof: see Appendix A.

In particular, the nonstationary signals, considered in this paper, are compactly supported on the interval $[0, T_w]$. Meanwhile, for the practical implementation, the frequency content is also confined to a finite interval $[-f_\Omega/2, f_\Omega/2]$. Therefore, $\{f_l\}_{l\in\mathbb{Z}}$ and $\{kT_0\}_{k\in\mathbb{Z}}$ are finite and supposed to make a division for the joint time-frequency interval. The values of K and L are chosen to make sure that $\{M_{f_l} T_{kT} g_l(t)\}_{1\leq l\leq L, 0\leq k\leq K-1}$ can spread across the intervals $[0, T_w]$ and $[-f_\Omega/2, f_\Omega/2]$. Assume that $g_{l\,\min}(t)$ and $g_{l\,\max}(t)$ are the window functions with minimum support width and maximum support width in time, respectively. According to the Uncertainty principle, minimum support in time will result in the maximum support in frequency. Therefore, we consider the time interval $[-T_{\max}/2, T_{\max}/2]$ for $g_{l\,\max}(t)$ and frequency interval $[-f_{\max}/2, f_{\max}/2]$ for $G_{l\,\min}(f)$. The theoretical bounds for K and L are given by

$$K \geq \frac{2T_w + T_{\max}}{2T_0} \tag{41}$$

$$L \geq \frac{f_\Omega + f_{\max}}{2f_0} - 1 \tag{42}$$

where symmetry in frequency is considered for real-valued nonstationary signals.

To support a higher f_Ω, the direct method is to increase the value of L, such that more channels are used in the system. On the other hand, we can also increase the bandwidth of low-pass filter. Then the time-shifting parameter T_0 must be reduced, resulting in higher inner-channel processing load and working bandwidth for the (pseudo-)random sequence generator.

C. NOISES AND MISMATCH

As analyzed in Section V.A and V.B, we can achieve accurate reconstruction for the nonstationary signals from the sparse time-frequency coefficients. However, in the practical implementation, the AIC system will inevitably be interfered by noises and mismatch. More specifically, the sampling process may be interfered by some uncertain factors such as circuit crosstalk, grounding and measurement instability. Then noises will be mixed in the samples. On the other hand, there is also a mismatch for the sampling and reconstruction process, such as the finite samples and approximate duality. Therefore, we make an analysis for the reconstruction error at the existence of noises and mismatch.

Assume that $x(t) = x_0(t) + e(t)$, $x_0(t)$ is the original signal without noises, $e(t)$ is the Gaussian noise. Then in the l-th channel, we have

$$\begin{aligned} f_l(t) &= \int_{-\infty}^{+\infty} [x_0(t') + e(t')] \cdot {g_l}^*(t' - t)e^{-j2\pi f_l t'}dt' \\ &= \int_{-\infty}^{+\infty} x_0(t') \cdot {g_l}^*(t' - t)e^{-j2\pi f_l t'}dt' \\ &\quad + \int_{-\infty}^{+\infty} e(t') \cdot {g_l}^*(t' - t)e^{-j2\pi f_l t'}dt' \\ &= f_l^0(t) + f_l^e(t) \end{aligned} \tag{43}$$

where

$$f_l^0(t) = \int_{-\infty}^{+\infty} x_0(t') \cdot {g_l}^*(t' - t)e^{-j2\pi f_l t'}dt' \tag{44}$$

$$f_l^e(t) = \int_{-\infty}^{+\infty} e(t') \cdot {g_l}^*(t' - t)e^{-j2\pi f_l t'}dt' \tag{45}$$

Then the filter output $f_l^d(t)$ at time $t = kT_0$ is expressed as $f_{l,k} = f_l^0(kT_0) + f_l^e(kT) = f_{l,k}^0 + f_{l,k}^e$. Besides, we introduce $w_{l,m}$ as the Gaussian noise caused by uncertain factors during

sampling. Then the system output can be depicted as

$$y_{l,m}=\sum_{k=0}^{K-1} f_{l,k}\varepsilon_{m,k}+w_{l,m}=\sum_{k=0}^{K-1}(f^0_{l,k}\varepsilon_{m,k}+f^e_{l,k}\varepsilon_{m,k})+w_{l,m} \tag{46}$$

In matrix form, we have

$$\boldsymbol{y}_l=\boldsymbol{\Phi}(\boldsymbol{f}^0_l+\boldsymbol{f}^e_l)+\boldsymbol{w}_l \tag{47}$$

where

$$\boldsymbol{f}^0_l=[f^0_{l,0},f^0_{l,1},\ldots f^0_{l,K-1}]^T \tag{48}$$
$$\boldsymbol{f}^e_l=[f^e_{l,0},f^e_{l,1},\ldots f^e_{l,K-1}]^T \tag{49}$$
$$\boldsymbol{w}_l=[w_{l,0},w_{l,1},\ldots w_{l,K-1}]^T \tag{50}$$

Assume that $\boldsymbol{z}_l=\boldsymbol{\Phi}\boldsymbol{f}^e_l+\boldsymbol{w}_l$, we have

$$\boldsymbol{y}_l=\boldsymbol{\Phi}\boldsymbol{f}^0_l+\boldsymbol{z}_l \tag{51}$$

Since the noises are independent, entries in $\boldsymbol{z}_l$ also follow Gaussian distribution. Assume that the variance for $\boldsymbol{z}_l$ is σ_z^2, we modify the reconstruction model as

$$\hat{\boldsymbol{f}}^0_l=\arg\min\left\|\boldsymbol{f}^0_l\right\|_0 \quad subject\ to\ \left\|\boldsymbol{y}_l-\boldsymbol{\Phi}\boldsymbol{f}^0_l\right\|_2^2\le\sigma_z^2 \tag{52}$$

In words, $\hat{\boldsymbol{f}}^0_l$ is an estimation for $\boldsymbol{f}^0_l$. To measure the errors between $\hat{\boldsymbol{f}}^0_l$ and $\boldsymbol{f}^0_l$, for $1\le l\le L$, we assume that

$$\sum_{k=0}^{K-1}\sum_{l=1}^{L}\left|f^0_{l,k}-\hat{f}^0_{l,k}\right|^2\le\mu_1^2\sigma_z^2 \tag{53}$$

where $\hat{f}^0_{l,k}$ is the entry in $\hat{\boldsymbol{f}}^0_l$, μ_1 is determined by both reconstruction algorithm and matrix $\boldsymbol{\Phi}$.

For the practical implementation, the signals are assumed to be compactly supported on both time and frequency. This assumption will absolutely cause energy loss during sampling. Support that $x^c_0(t)$ is the essential signal for $x_0(t)$ with the essential band limited into $[-\Omega/2,\Omega/2]$, to yield

$$\left\|x_0(t)-x^c_0(t)\right\|_2^2\le\epsilon_\Omega\left\|x_0(t)\right\|_2^2 \tag{54}$$

It is implied that the time-frequency coefficient $f^0_{l,k}$ is only obtained from $x^c_0(t)$, such that

$$\left\|x_0(t)-\sum_{k=0}^{K-1}\sum_{l=1}^{L}\left\langle x^c_0(t),M_{fl}T_{kT_0}g_l(t)\right\rangle M_{fl}T_{kT_0}\gamma_l(t)\right\|_2^2\le\epsilon_\Omega\left\|x_0(t)\right\|_2^2 \tag{55}$$

where $f^0_{l,k}=\left\langle x^c_0(t),M_{fl}T_{kT_0}g_l(t)\right\rangle$.

Actually, there is also a mismatch between the dual windows $g_l(t)$ and $\gamma_l(t)$, since optimization is used while calculating the expression of $\gamma_l(t)$. Given $\gamma_l(t)$, we assume that the accurate dual window for $\gamma_l(t)$ is $g'_l(t)$. Then the mismatch can be measured by **Theorem II.**

Theorem II: Assume that $x^c_0(t)$ is essential signal, $\gamma_l(t)$ and $g'_l(t)$ are dual windows, and $g_l(t)$ is the approximation for $g'_l(t)$ satisfying

$$\left(\sum_{k=0}^{K-1}\sum_{l=1}^{L}\left|\left\langle x^c_0(t),M_{fl}T_{kT_0}[g(t)-g'(t)]\right\rangle\right|^2\right)^{1/2}\le\mu_2\left\|x^c_0(t)\right\|_2 \tag{56}$$

Then we have

$$\left\|x^c_0(t)-\sum_{k=0}^{K-1}\sum_{l=1}^{L}\left\langle x^c_0(t),M_{fl}T_{kT_0}g_l(t)\right\rangle M_{fl}T_{kT_0}\gamma_l(t)\right\|_2\le\mu_2\sqrt{B_\gamma}\left\|x^c_0(t)\right\|_2 \tag{57}$$

where B_γ is the upper bound for the frame constructed by $\gamma_l(t)$.

Proof: See Appendix B.

With all the factors mentioned above, we now give a bounded total error as expressed in **Corollary I**.

Corollary I: Assume that $x_0(t)$ is the nonstationary signal without noises, $\hat{f}^0_{l,k}$ is the estimation for the time-frequency coefficients. Then the total deviation after reconstruction can be given by

$$\left\|x_0(t)-\sum_{k=0}^{K-1}\sum_{l=1}^{L}\hat{f}^0_{l,k}M_{fl}T_{kT_0}\gamma_l(t)\right\|_2\le\sqrt{\epsilon_\Omega}\left\|x_0(t)\right\|_2+\mu_2\sqrt{B_\gamma}\left\|x^c_0(t)\right\|_2+\mu_1\sigma_N\sqrt{B_\gamma} \tag{58}$$

VI. NUMERICAL SIMULATIONS

We now present some numerical experiments to illustrate the effectiveness of the proposed AIC system. The nonstationary signals used in the experiment are LFM and HF signals. One example of such signals is shown in Figure 1 and Figure 2. Here, the signals are compactly supported on the time interval [0, 0.512ms] and essentially bandlimited to [0.05MHz, 0.45MHz]. For LFM signals, the instantaneous frequency $f_i(t)$ is set to

$$f_i(t)=50000+781250000t \tag{59}$$

For HF signals, the maximum time duration between frequencies is 0.057ms, the hopping range is [0.05MHz, 0.45MHz]. According to the bandwidths of LFM and HF signals, we set $f_{Nyq}=1$MHz, such that $T_{Nyq}=1/f_{Nyq}=1us$.

Considering the properties of signals, we use a fixed window in different channels, created by Gaussian function with $\sigma=7\times10^{-6}$. Then the corresponding filter impulse response and magnitude response are shown in Figure 6. The support widths in time and frequency are 0.042ms and 0.14MHz.

A. LATTICES FOR COMPLETENESS

To make sure that lattices are complete to represent any nonstationary signals, we discuss the reasonable values of shifting parameters T_0 and f_l. Since the low-pass filters in different channels are the same, we set frequency-shifting parameter $f_l=lf_0$ to make a uniform division for the essential band. According to **Theorem I**, the frame bounds are

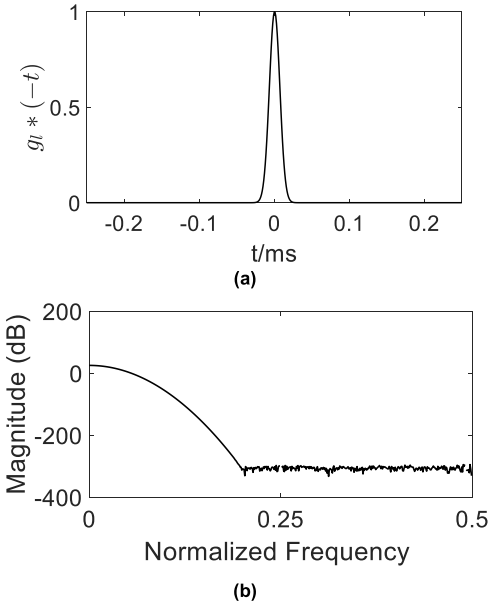

FIGURE 6. The impulse response and magnitude response for the low-pass filter: (a) the impulse response function; (b) the magnitude response function.

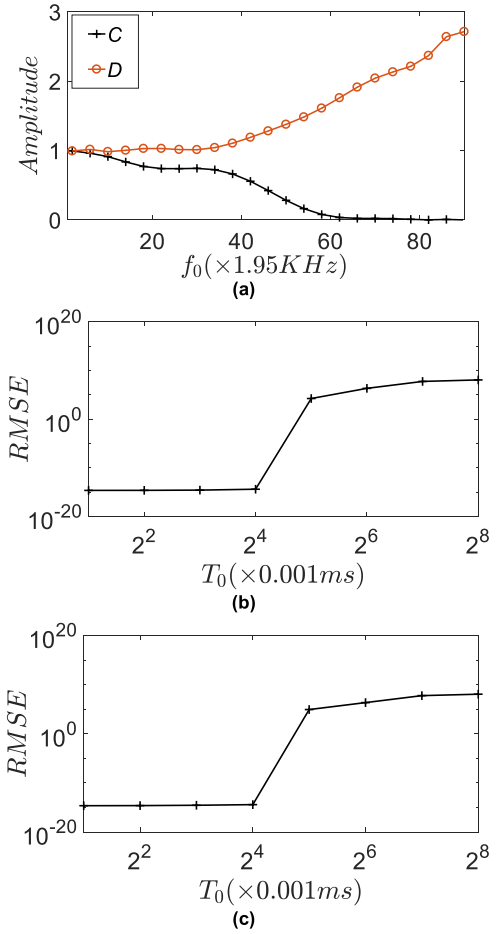

FIGURE 7. The analysis for lattice parameters T_0 and f_l: (a) the values of C, D with different f_0; (b) RMSE for LFM signals; (c) RMSE for HF signals.

determined by C, D and time-shifting parameter T_0. Firstly, we show the values of C, D with different f_0. The result is shown in Figure 7 (a). It is seen that small f_0 will contributes to the tightness of the frame, where C and D get close to 1 simultaneously. To guarantee the sampling density, in the following experiment, we set $f_0 = 62.5KHz$.

Then, we discuss the appropriate value for T_0. Actually, the constraint for T_0 in **Theorem I** is sufficient to construct a frame. Then with the filter shown in Figure 6, we have $T_0 \leq 7.14 \times 10^{-3} ms$. However, this constraint is not necessary. Given f_0, T_0 will directly determine the number of samples. So we want T_0 to be large enough to reduce the sample number as much as possible. We recover the original signals with the filter output to analyze the impact of T_0 in time-frequency representation. The result is shown in Figure 7 (b) and (c). In this figure, we introduce the Root Mean Square Error (RMSE) to measure the reconstruction error, such that

$$\text{RMSE} = \sqrt{\frac{\left\|\hat{x}(t) - x(t)\right\|_2^2}{\text{supp}\,|x(t)|}} \tag{60}$$

where $\hat{x}(t)$ is the reconstructed nonstationary signals. It is seen that low errors can also be achieved, even T_0 is larger than $7.14 \times 10^{-3} ms$. Therefore, we set $T_0 \leq 16 \times 10^{-3} ms$ for both LFM and HF signals, where time-frequency coefficients are able to perfectly represent the nonstationary signals.

B. RECONSTRUCTION FOR AIC SYSTEM AND COMPARISON

Now, we present the reconstruction effect using the samples from the proposed AIC system. We firstly use the system output to reconstruct the sparse time-frequency coefficients, and then recover the original nonstationary signals under the frame theory. The sparse optimization used in this paper is SBL algorithm [44]. As analysis above, we set $f_0 = 62.5KHz$. And T_0 is set to $16 \times 10^{-3} ms$ $8 \times 10^{-3} ms$, $4 \times 10^{-3} ms$, respectively. Corresponding, we have $L = 8$, $K = 34$, 67, and 134.

We set different values of M, such that different numbers of samples are obtained. We use the Monte Carlo method to conduct the experiment 500 times for each setting. For each time, we use RMSE to measure the reconstruction error. If RMSE < 0.01, the reconstruction is viewed as success; otherwise, it is false. The result is shown in Figure 8. It is seen that the original signals can be accurately reconstructed with high probability when M is large enough. That conforms to the basic CS theory. Besides, it is seen that the minimum M guaranteeing the high success probability increases with K. This is because that, for the same input signals, the ratio K/S is constant. Then according to **Lemma 1**, the minimum M is calculated by $M \geq \text{O}(S \times \log(K/S))$. The increasing in K will lead to a larger value of S. Then, more samples are required to guarantee accuracy reconstruction.

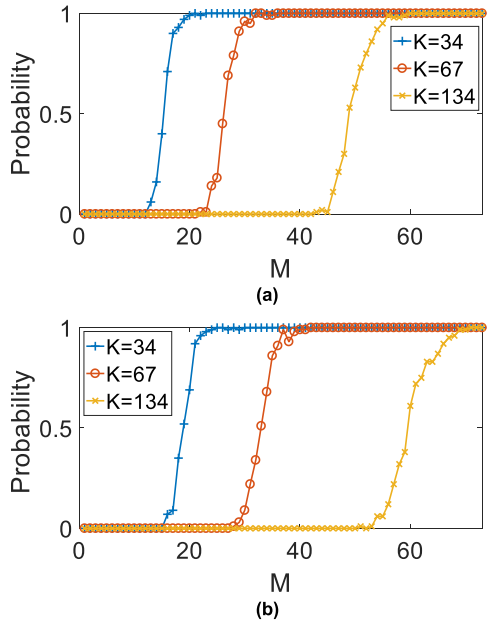

FIGURE 8. Reconstruction probability for different M and K: (a) for LFM signals; (b) for HF signals.

We further compare the effectiveness of the proposed AIC system with other sampling methods state-of-the-art. We introduce two widely-used sampling methods, including Nyquist sampling and Time-interleaving sampling, which are represented by the terms 'NS' and 'TIS'. And two other AIC systems, including RD and MWC, are also introduced to conduct the sub-Nyquist sampling for nonstationary signals. For TIS, the channel number is set to L. For RD, the sampling rate is LM/T. For MWC, the channel number is also L, the sampling rate in one channel is M/T. For the proposed AIC system, we set $K = 34$. Besides, we set $M = 20, M = 24$ for LFM and HF signals, respectively. Then the proposed AIC system achieves the accurate reconstruction with the probability up to 99%. The total sampling rate is the product of channel number and sampling rate (one channel), whereas the total sample number is calculated in the same way. The comparison is shown in Table 2.

It is seen that the proposed AIC system reduces the sampling rate and sample number significantly. Meanwhile, the total sampling rate and sample number are also below 'Nyquist' method. For TIS, it is seen that although sampling rate (one channel) and sample number (one channel) are reduced, the total sampling rate and sample number are the same with 'Nyquist' method. That is the essential difference between AIC system and TIS, since the AIC system achieves the sub-Nyquist sampling for signals. Compared with RD and MWC, we know that, with the same total sample number, the proposed AIC system achieves the lowest sampling rate. Moreover, the simplified AIC further reduces the total sampling rate by increase the sample number in one channel. That means the high utilization of system channels also contributes to the decrease in the total sampling rate.

TABLE 2. Comparison for effectiveness.

		Channe-ls	Sampli-ng rate /KHz (one channel)	Total sampli-ng rate /KHz	Sample number (one channel)	Total samp-le numbe r
LFM	NS	1	1000	1000	512	512
	TLS	8	125	1000	64	512
	RD	1	312.5	312.5	160	160
	MWC	8	39.06	312.5	20	160
	Proposed	160	1.95	312.5	1	160
	Simplified	8	1.95	15.6	20	160
HF	NS	1	1000	1000	512	512
	TLS	8	125	1000	64	512
	RD	1	375	375	192	192
	MWC	8	46.9	375	24	192
	Proposed	192	1.95	375	1	192
	Simplified	8	1.95	15.6	24	192

For further comparison, we use the samples from the AIC systems, including RD, MWC and proposed AIC system, to conduct the reconstruction for nonstationary signals. For LFM and HF signals, M is set to 20, 40, and 60, respectively. The reconstructed signals are shown in Figure 9 and Figure 10. To measure the reconstruction effect, RMSEs are also calculated in Table 3.

TABLE 3. Comparison of reconstruction error.

		RD	MWC	Proposed
LFM	M=20	9.88×10^{-2}	6.23×10^{-2}	1.62×10^{-3}
	M=40	6.56×10^{-2}	5.78×10^{-2}	1.91×10^{-5}
	M=60	2.23×10^{-2}	7.40×10^{-3}	1.52×10^{-5}
HF	M=20	1.02×10^{-1}	5.99×10^{-2}	3.57×10^{-3}
	M=40	6.31×10^{-2}	3.37×10^{-2}	7.41×10^{-5}
	M=60	9.50×10^{-3}	7.62×10^{-3}	4.91×10^{-5}

It is seen from Figure 9 and Figure 10 that by increasing the value of M, the reconstruction effect is improved. And the proposed AIC system achieves a better reconstruction than RD and MWC. For $M = 20$ or $M = 40$, RD and MWC fail to make an effect reconstruction for LFM and HF signals, whereas the proposed AIC system reconstructs the original signals preferably. And as shown in Table 3, for $M = 20$ or $M = 40$, RMSEs of RD and MWC are larger than the proposed AIC system. When $M = 60$, the reconstruction effects for RD and MWC are obviously improved. However, the reconstruction error is still larger than the proposed AIC system. Therefore, the effectiveness of the proposed AIC system is verified.

C. THE IMPACT OF NOISES

To explore the impact of noises, we recalculate the RMSE with noises in different Signal-to-Noise Ratio (SNR). With $K = 34$, we set $M \geq 20$ and $M \geq 24$ for LFM and HF signals, respectively, where the accurate reconstruction probability without noises is up to 99%. For each setting, the experiment is conducted 500 times. Similarly,

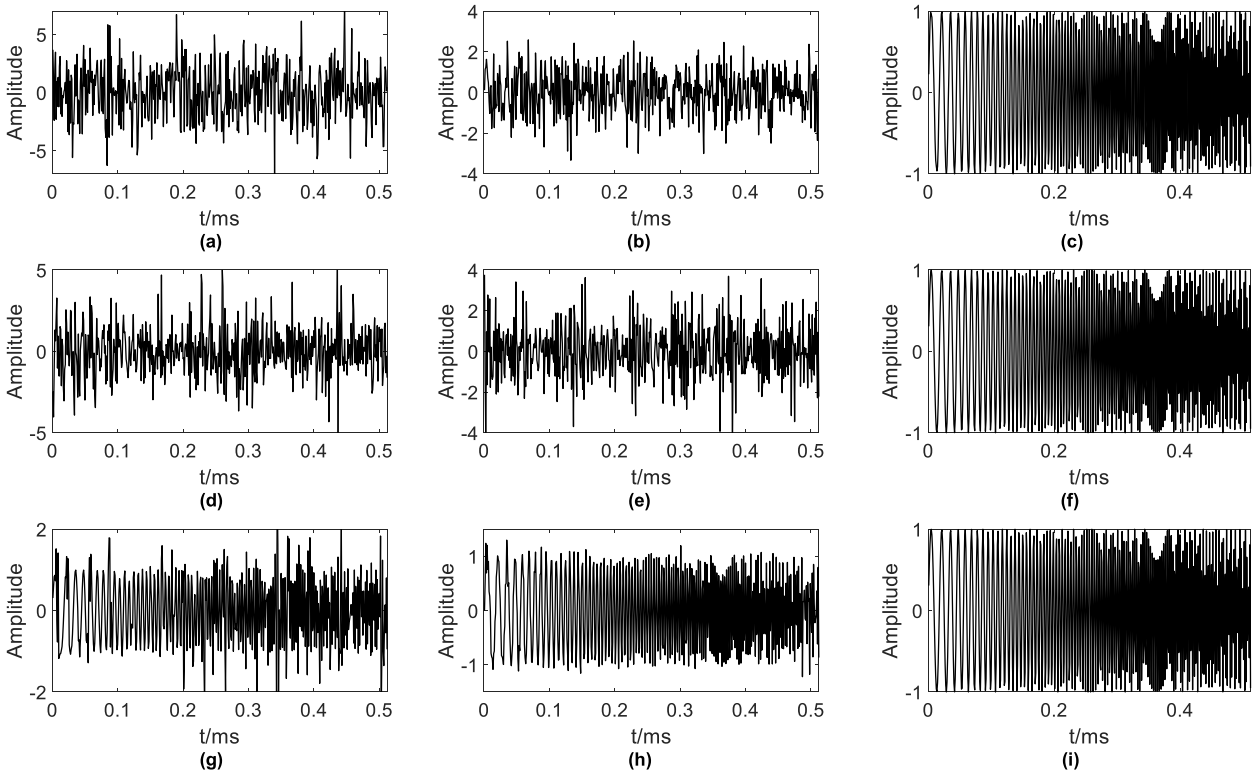

FIGURE 9. Reconstructed LFM signals: (a) RD with $M = 20$ (b) MWC with $M = 20$ (c) Proposed AIC system with $M = 20$ (d) RD with $M = 40$ (e) MWC with $M = 40$ (f) Proposed AIC system with $M = 40$ (g) RD with $M = 60$ (h) MWC with $M = 60$ (i) Proposed AIC system with $M = 60$.

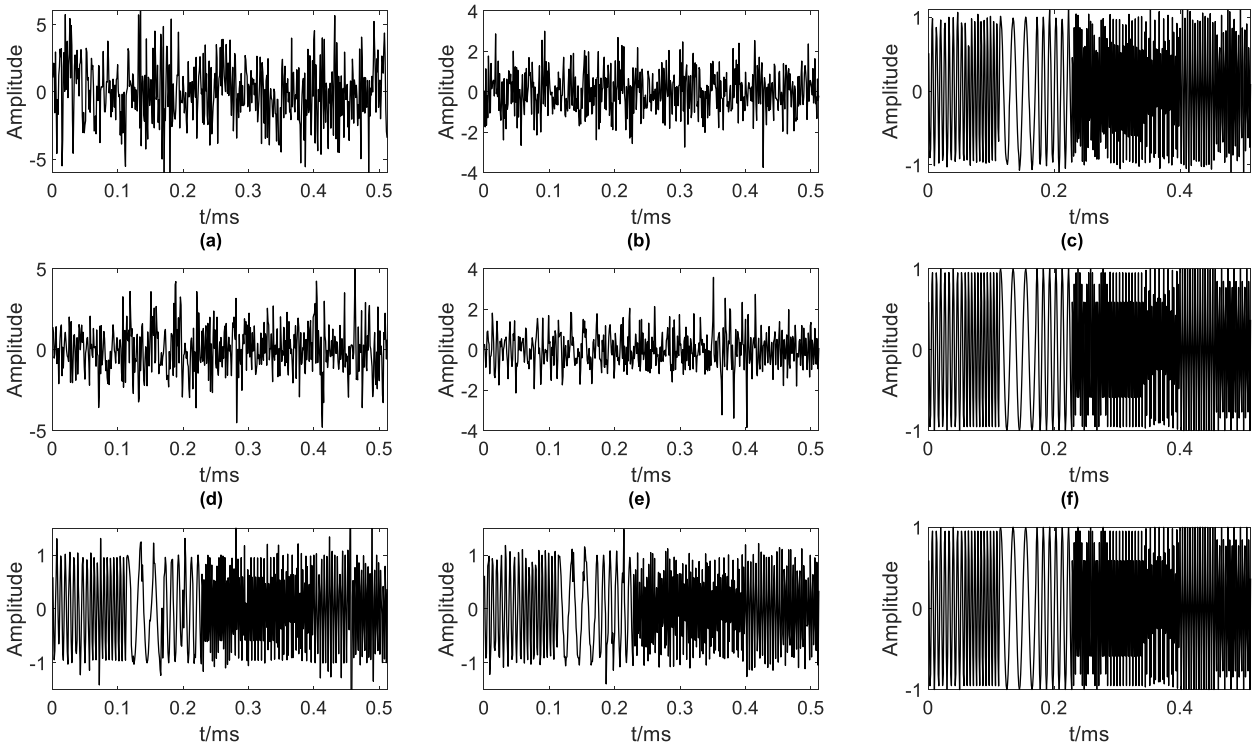

FIGURE 10. Reconstructed HF signals: (a) RD with $M = 20$ (b) MWC with $M = 20$ (c) Proposed AIC system with $M = 20$ (d) RD with $M = 40$ (e) MWC with $M = 40$ (f) Proposed AIC system with $M = 40$ (g) RD with $M = 60$ (h) MWC with $M = 60$ (i) Proposed AIC system with $M = 60$.

If RMSE < 0.01, the reconstruction is viewed as accurate; otherwise, it is false.

The result is shown in Figure 11. It is seen that the high probability can also be guaranteed if SNR is large.

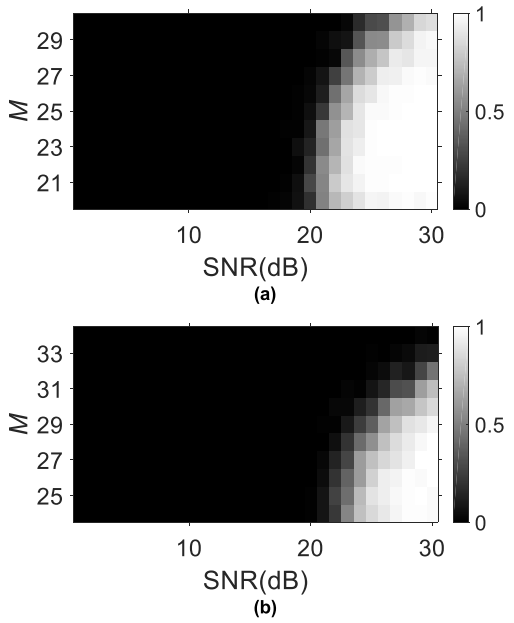

FIGURE 11. The reconstruction probability with noises: (a) for LFM signals; (b) for HF signals.

Meanwhile, the probability is also impacted by M. When we increase M slightly, the reconstruct effect is improved, as the high probability can be obtained with a smaller SNR. However, when we continue to increase the value of M, the reconstruction effect is worse. This is mainly because that the increasing in M also introduces more measurement noises. Therefore, in practical implementation, we can increase the M appropriately to enhance the robustness of AIC system. For example, in this experiment, we can choose $M = 21$ and $M = 25$ for LFM and HF signals, respectively. Then the best reconstruct effect can be obtained.

D. IMPLEMENTATION ASPECTS

Considering the implementation in practical circuit, the proposed AIC system mainly contains the following components: modulator, filter, mixer, (pseudo-)random sequence generator, integrator and ADC. Typically, the mixer, (pseudo-)random sequence generator, integrator and ADC components are commonly shared by some current CS-based AIC systems [28], [34]. Therefore, in this sub-section, we mainly analyze the implementation of modulator and filter.

Under the proposed scheme, the modulating function is the conventional frequency modulating signal, such that it can be realized exploiting the widely used frequency synthesis technique. Nowadays, the frequency synthesizer with high stability and precision is available [51], [52]. Stable output signals at the frequency up to 60 GHz will provide a high working band for the proposed AIC system.

Analog filter is the key missing piece for the proposed AIC system. For the conventional low-pass filter, the filter designing mainly focuses on the performance in magnitude response. However, in this paper, we may pay more attention to the filter's impulse response. The latest technique in continuous-time analog filter designing provides a feasible approach to design an integrated analog Gaussian filter [53], [54]. The rational approximation for the Gaussian function in time will guarantee the accurate time-frequency representation for nonstationary signals.

Considering all the aspects above, we make a comparison about implementation complexity between the proposed AIC system and some other sampling methods state-of-the-art, including:

1) Nyquist sampling(NS);
2) Time-interleaving sampling (TIS);
3) RD;
4) MWC.

It is seen from Table 4 that compared with Nyquist and T-Nyquist sampling methods, CS-based AIC systems (RD, MWC and Proposed) use more components in implementation. This is because to achieve the lower sampling rate, a more complex system construction is needed. Similarly, we can see that the proposed AIC system has the most components, denoting that it even has a higher complexity than RD and MWC. However, it is important to note that the proposed AIC system still makes sense in the application, because we have a further lower sampling rate that RD and MWC. Besides, we reduce the working rate for (pseudo-)random sequence generator, such that it allows a more flexible choice of (pseudo-)random sequence generator in the application.

TABLE 4. Comparison for implementation.

	Modulator	Filter	Mixer	Sequence generator	Integrator	ADC
NS	/	/	/	/	/	√
TIS	/	/	/	/	/	√
RD	/	/	√	√	√	√
MWC	/	/	√	√	√	√
Proposed	√	√	√	√	√	√

Another problem in hardware implementation is the non-ideal factors for the analog realization of the proposed AIC system. Since most components of the proposed AIC system works in analog form, the non-ideal analog devices will absolutely bring errors during the sampling and reconstruction. Therefore, further studies are required to search a satisfactory solution.

VII. DISCUSSION

In this section, we make a discussion about the connections between the proposed AIC system and some other related works, and then make a conclusion about this paper.

A. RELATED WORKS

The most direct precedent for this paper is the CS theory, such works as [5], [39], and [41]. Besides, as shown in work [8], the random measurement matrix has been proven to be effective for some applications. Considering more about the CS theory, these works mainly concentrate on the feasibility of compressive sampling and sparse reconstruction. Inspired by

these works, we try to transport the theoretical achievements into practice. Therefore, we design an analog front-end to deal with the analog signals directly, and propose the feasible architecture to achieve the compressive sampling. This work can be viewed as an application of CS theory for nonstationary signals. Indeed, our numerical experiments also echo some results in CS theory.

Actually, the idea of CS-based AIC system has been studied previously in some works, and the item 'Xampling' has been proposed to convey a balance between CS techniques and traditional concepts in sampling theory [46], [47]. Some AIC systems have been proven to be achievable, such as MWC and RD. Indeed, our proposed system is another case for 'Xampling' system. The main difference of our work lies in the input signal. For these AIC systems, sparsity for input signals in frequency domain is necessary to guarantee the perfect reconstruction. However, in this paper, we study the nonstationary signals, which are usually not sparse in frequency. To deal with this problem, in this paper, we propose a novel sampling scheme to realize the compressive sampling in time-frequency domain. We show that it is feasible to recover the nonstationary signals accurately with a high probability.

This paper is highly inspired by fruitful results in time-frequency analysis, such works as [42] and [43]. Besides, the conception of CS in time-frequency domain has also been studied in some works such as [48] and [49]. These works explore the feasibility of sparse representation in time-frequency and provide the theoretical analysis for the reconstruction process. Compared with these works, we consider more about practical implementation. So the time-frequency representation is more practical. Actually, the proposed AIC system is also inspired by the work [50], where a Gabor-based sampling system is proposed for short pulses. Despite the similarity, in this paper, we focus on the sub-Nyquist sampling for nonstationary signals. Therefore, we extend the Gabor frame to a more general situation with irregular lattices and variable window function. Moreover, we present a novel architecture with advantages in practical implementation.

B. CONCLUSION

In this paper, we propose a novel AIC system to achieve the sub-Nyquist sampling for nonstationary signals. We present a simple architecture such that it can be implemented with existing analog devices and ADCs. The sub-Nyquist sampling is achieved by exploring the sparsity in time-frequency domain. With the proposed system, we can sample the nonstationary signals at a low rate and recover the original signals without using a priori information.

The sampling process is conducted in a multi-channel AIC system consisting of low-pass filters, integrators, and ADCs. We provide a description for the system construction schematically, and then analyze the sampling process to reveal the essential relationship between time-frequency coefficients and system samples. For practical implementation, we further provide the simplified system construction to reduce inner-channels and (pseudo-)random sequence generators. In reconstruction, we establish the reconstruction model for time-frequency coefficients, and then construct a frame to make sure that these coefficients are complete to recover the original signals. We consider the reconstruction error with the existence of noises and mismatch. It is shown that the proposed AIC system has bounded total error. We verify the effectiveness by numerical experiments. It is seen that the proposed AIC system outperforms the other sampling methods state-of-the-art.

For practical application, there are some other aspects we should pay attention to. Firstly, the accuracy reconstruction for nonstationary signals depends on the good sparsity in time-frequency domain. If the signals are not sparse sufficiently, we may need more samples to guarantee accurate reconstructed. In words, the proposed system is only suitable for the nonstationary signals which have good sparsity in time-frequency domain. Secondly, compared with some other CS-based AIC systems state-of-the-art, the proposed AIC system contains more components in circuit implementation. Therefore, further study is still required to simplify hardware implementation. Finally, the efficient sampling for nonstationary signals is determined by the bandwidth of input signals. That means the fixed system parameters may not suitable for all input signals. The feasible solution is to conduct a coarse signal analysis before the sampling process. However, the specific implementation scheme requires further study.

APPENDIX A
PROOF OF THEOREM I

We present some basic properties for the translation and modulation operators as

$$\langle x, T_{KT_0} g\rangle = \langle T_{-KT_0} x, g\rangle, \quad \langle x, M_{f_l} g\rangle = \langle M_{-f_l} x, g\rangle \tag{61}$$

$$\left(T_{KT_0} x\right)^{\wedge} = M_{-KT_0} X, \quad \left(M_{f_l} x\right)^{\wedge} = T_{KT_0} X \tag{62}$$

where $(\cdot)^{\wedge}$ is to conduct the FT. Then for any $x(t) \in L_2(\mathbb{R})$, we have

$$\sum_{k,l\in\mathbb{Z}} \left|\langle x, M_{f_l} T_{kT_0} g_l\rangle\right|^2 = \sum_{k,l\in\mathbb{Z}} \left|\langle X, T_{f_l} M_{-kT_0} G_l\rangle\right|^2 \tag{63}$$

For convenience, we ignore the phase factors in $T_{-f_l} M_{kT_0} G_l$, since they vanish in the inner product. Then

$$\begin{aligned}\sum_{k,l\in\mathbb{Z}} \left|\langle X, T_{f_l} M_{-kT_0} G_l\rangle\right|^2 &= \sum_{k,l\in\mathbb{Z}} \left|\langle T_{-f_l} X, M_{-kT_0} G_l\rangle\right|^2 \\ &= \sum_{k,l\in\mathbb{Z}} \left|\langle T_{-f_l} X G_l{}^*, M_{-kT_0}\rangle\right|^2 \end{aligned} \tag{64}$$

Assume that the support interval for $G_l(f)$ is I_l. In words, given $l \in \mathbb{Z}$, $T_{-f_l} X G_l{}^*$ is also compactly supported on interval I_l. Since $\max_l |\mathrm{supp}(G_l(f))| \le \frac{1}{T_0}$, we have $I_l \le \frac{1}{T_0}$.

Then $\left\{\frac{1}{T_0} M_{-kT_0}\right\}_{k\in\mathbb{Z}}$ is an orthonormal basis for $L_2(I_l)$. So we have

$$\sum_{k,l\in\mathbb{Z}} \left|\langle T_{-f_l} X G_l{}^*, M_{-kT_0}\rangle\right|^2$$

$$= \frac{1}{T_0}\sum_{l\in\mathbb{Z}}\int_{-\infty}^{+\infty}\left|T_{-f_l}X(f)\cdot {G_l}^*(f)\right|^2$$
$$= \frac{1}{T_0}\sum_{l\in\mathbb{Z}}\int_{-\infty}^{+\infty}\left|X(f)\cdot T_{f_l}{G_l}^*(f)\right|^2$$
$$= \frac{1}{T_0}\sum_{l\in\mathbb{Z}}\int_{-\infty}^{+\infty}|X(f)|^2\left|T_{f_l}{G_l}^*(f)\right|^2 df \quad (65)$$

Such that

$$\frac{C}{T_0}\|X(f)\|^2 \le \frac{1}{T_0}\sum_{l\in\mathbb{Z}}\int_{-\infty}^{+\infty}|X(f)|^2\left|T_{f_l}{G_l}^*(f)\right|^2 df \le \frac{D}{T_0}\|X(f)\|^2 \quad (66)$$

Then **Theorem I** is proven.

APPENDIX B
PROOF OF THEOREM II

As the approximate dual window for $g(t)$, $\gamma(t)$ is also supposed to construct a frame with the bounds A_γ and B_γ, such that

$$A_\gamma\|x\|^2 \le \sum_{k,l\in\mathbb{Z}}\left|\langle x, M_{f_l}T_{kT_0}\gamma_l\rangle\right|^2 \le B_\gamma\|x\|^2 \quad (67)$$

$$\left\|\sum_{k,l\in\mathbb{Z}}c_{k,l}M_{f_l}T_{kT_0}\gamma_l\right\|_2^2 \le B_\gamma\sum_{k,l\in\mathbb{Z}}^{|c_{k,l}|^2} \quad (68)$$

Then we have

$$\left\|x_0^c(t)-\sum_{k=0}^{K-1}\sum_{l=1}^{L}\langle x_0^c(t), M_{f_l}T_{kT_0}g_l(t)\rangle M_{f_l}T_{kT_0}\gamma_l(t)\right\|_2^2$$
$$= \left\|\sum_{k=0}^{K-1}\sum_{l=1}^{L}\langle x_0^c(t), M_{f_l}T_{kT_0}[g_l'(t)-g_l(t)]\rangle M_{f_l}T_{kT_0}\gamma_l(t)\right\|_2^2$$
$$\le B_\gamma\sum_{k=0}^{K-1}\sum_{l=1}^{L}\left|\langle x_0^c(t), M_{f_l}T_{kT_0}[g_l'(t)-g_l(t)]\rangle\right|^2 \quad (69)$$

Such that

$$\left\|x_0^c(t)-\sum_{k=0}^{K-1}\sum_{l=1}^{L}\langle x_0^c(t), M_{f_l}T_{kT_0}g_l(t)\rangle M_{f_l}T_{kT_0}\gamma_l(t)\right\|_2 \le \mu_2\sqrt{B_\gamma}\left\|x_0^c(t)\right\|_2 \quad (70)$$

Then **Theorem II** is proven.

REFERENCES

[1] C. E. Shannon, "A mathematical theory of communication," *Bell Syst. Tech. J.*, vol. 27, no. 3, pp. 623–656, Jul./Oct. 1948.
[2] H. Nyquist, "Certain topics in telegraph transmission theory," *Proc. IEEE*, vol. 90, no. 2, pp. 280–305, Apr. 2002.
[3] H. J. Landau, "Sampling, data transmission, and the Nyquist rate," *Proc. IEEE*, vol. 55, no. 10, pp. 1701–1706, Oct. 1967.
[4] D. L. Donoho and M. Elad, "Optimally sparse representation in general (nonorthogonal) dictionaries via ℓ_1 minimization," *Proc. Nat. Acad. Sci. USA*, vol. 100, no. 5, pp. 2197–2202, 2003.
[5] D. L. Donoho, "Compressed sensing," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1289–1306, Apr. 2006.
[6] R. Baraniuk, M. Davenport, R. Devore, and M. Wakin, "A simple proof of the restricted isometry property for random matrices," *Constructive Approximation*, vol. 28, no. 3, pp. 253–263, 2008.
[7] S. Liu, Y. D. Zhang, and T. Shan, "Sparsity-based frequency-hopping spectrum estimation with missing samples," in *Proc. IEEE Radar Conf. (RadarConf)*, May 2016, pp. 1043–1047.
[8] E. J. Candes and T. Tao, "Near-optimal signal recovery from random projections: Universal encoding strategies?" *IEEE Trans. Inf. Theory*, vol. 52, no. 12, pp. 5406–5425, Dec. 2006.
[9] S. K. Sharma, E. Lagunas, S. Chatzinotas, and B. Ottersten, "Application of compressive sensing in cognitive radio communications: A survey," *IEEE Commun. Surveys Tuts.*, vol. 18, no. 3, pp. 1838–1860, 3rd Quart., 2016.
[10] H. Sun, A. Nallanathan, C.-X. Wang, and Y. Chen, "Wideband spectrum sensing for cognitive radio networks: A survey," *IEEE Wireless Commun.*, vol. 20, no. 2, pp. 74–81, Apr. 2013.
[11] S. Liu, Y. Ma, and T. Shan, "Segmented discrete polynomial-phase transform with coprime sampling," *J. Eng.*, vol. 2019, no. 19, pp. 5619–5621, Oct. 2019.
[12] M. Porat and Y. Zeevi, "The generalized Gabor scheme of image representation in biological and machine vision," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 10, no. 4, pp. 452–467, Jul. 1987.
[13] F. J. Owens and M. S. Murphy, "A short-time Fourier transform," *Signal Process.*, vol. 14, no. 1, pp. 3–10, 1988.
[14] D. Griffin and J. Lim, "Signal estimation from modified short-time Fourier transform," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 32, no. 2, pp. 236–243, Apr. 1984.
[15] I. Daubechies, "The wavelet transform, time-frequency localization and signal analysis," *IEEE Trans. Inf. Theory*, vol. 36, no. 5, pp. 961–1005, Sep. 1990.
[16] S. V. Bozhokin and I. M. Suslova, "Double wavelet transform of frequency-modulated nonstationary signal," *Tech. Phys.*, vol. 58, no. 12, pp. 1730–1736, Dec. 2013.
[17] E. Wigner, "On the quantum correction for thermodynamic equilibrium," *Phys. Rev.*, vol. 40, no. 5, pp. 749–759, Jun. 1932.
[18] W. Martin and P. Flandrin, "Wigner-ville spectral analysis of nonstationary processes," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 33, no. 6, pp. 1461–1470, Dec. 1985.
[19] B. Le, T. W. Rondeau, J. H. Reed, and C. W. Bostian, "Analog-to-digital converters: A review of the past, present, and future," *IEEE Signal Process. Mag.*, vol. 22, no. 6, pp. 69–77, Nov. 2005.
[20] C. R. Parkey and W. B. Mikhael, "Cumulant statistical adaptation of nonlinear post conversion correction for TI-ADCs," in *Proc. IEEE Autotestcon*, Nov. 2015, pp. 121–126.
[21] C. R. Parkey and W. B. Mikhael, "Linearized adaptation of non-linear post conversion correction for TIADCs: A behavioral model study," *IEEE Instrum. Meas. Mag.*, vol. 18, no. 4, pp. 46–49, Aug. 2015.
[22] A. Bonnetat, J.-M. Hode, G. Ferre, and D. Dallet, "An adaptive all-digital blind compensation of dual-TIADC frequency-response mismatch based on complex signal correlations," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 62, no. 9, pp. 821–825, Sep. 2015.
[23] F. Guo, L. Yang, and Z. Zhang, "TDOA/FDOA estimation method based on dechirp," *IET Signal Process.*, vol. 10, no. 5, pp. 486–492, Jul. 2016.
[24] D. Wei and Y.-M. Li, "Generalized sampling expansions with multiple sampling rates for lowpass and bandpass signals in the fractional Fourier transform domain," *IEEE Trans. Signal Process.*, vol. 64, no. 18, pp. 4861–4874, Sep. 2016.
[25] A. Bhandari and A. I. Zayed, "Shift-invariant and sampling spaces associated with the fractional Fourier transform domain," *IEEE Trans. Signal Process.*, vol. 60, no. 4, pp. 1627–1636, Apr. 2012.
[26] O. Taheri and S. A. Vorobyov, "Segmented compressed sampling for analog-to-information conversion: Method and performance analysis," *IEEE Trans. Signal Process.*, vol. 59, no. 2, pp. 554–572, Feb. 2011.
[27] J. A. Tropp, J. N. Laska, M. F. Duarte, J. K. Romberg, and R. G. Baraniuk, "Beyond Nyquist: Efficient sampling of sparse bandlimited signals," *IEEE Trans. Inf. Theory*, vol. 56, no. 1, pp. 520–543, Jan. 2010.
[28] T. Ragheb, J. N. Laska, H. Nejati, S. Kirolos, R. G. Baraniuk, and Y. Massoud, "A prototype hardware for random demodulation based compressive analog-to-digital conversion," in *Proc. 51st Midwest Symp. Circuits Syst.*, Aug. 2008, pp. 37–40.
[29] P. J. Pankiewicz, T. Aridsen, and T. Larsen, "Sensitivity of the random demodulation framework to filter tolerances," in *Proc. IEEE Eur. Signal Process. Conf.*, Aug./Sep. 2015, pp. 534–538.

[30] A. Harms, W. U. Bajwa, and R. Calderbank, "A constrained random demodulator for sub-Nyquist sampling," *IEEE Trans. Signal Process.*, vol. 61, no. 3, pp. 707–723, Feb. 2013.
[31] M. Wakin, S. Becker, E. Nakamura, M. Grant, E. Sovero, D. Ching, J. Yoo, J. Romberg, A. Emami-Neyestanak, and E. Candes, "A nonuniform sampler for wideband spectrally-sparse environments," *IEEE J. Emerg. Sel. Topics Circuits Syst.*, vol. 2, no. 3, pp. 516–529, Sep. 2012.
[32] D. E. Bellasi, L. Bettini, C. Benkeser, T. Burger, Q. Huang, and C. Studer, "VLSI design of a monolithic compressive-sensing wideband analog-to-information converter," *IEEE J. Emerg. Sel. Topics Circuits Syst.*, vol. 3, no. 4, pp. 552–565, Dec. 2013.
[33] M. Pelissier and C. Studer, "Non-uniform wavelet sampling for RF analog-to-information conversion," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 65, no. 2, pp. 471–484, Feb. 2018.
[34] M. Mishali and Y. C. Eldar, "From theory to practice: Sub-Nyquist sampling of sparse wideband analog signals," *IEEE J. Sel. Topics Signal Process.*, vol. 4, no. 2, pp. 375–391, Apr. 2010.
[35] E. Israeli, S. Tsiper, D. Cohen, E. Shoshan, R. Hilgendorf, A. Reysenson, and Y. C. Eldar, "Hardware calibration of the modulated wideband converter," in *Proc. IEEE Global Commun. Conf.*, Dec. 2014, pp. 948–953.
[36] Z. Xu, Z. Li, and J. Li, "Broadband cooperative spectrum sensing based on distributed modulated wideband converter," *Sensors*, vol. 16, no. 10, pp. 1602–1613, 2016.
[37] D. Adams, Y. C. Eldar, and B. Murmann, "A mixer front end for a four-channel modulated wideband converter with 62-dB blocker rejection," *IEEE J. Solid-State Circuits*, vol. 52, no. 5, pp. 1286–1294, May 2017.
[38] J. Ville, "Theorie et applications de la notion de signal analytique," *Cables Transmiss.*, vol. 2A, no. 1, pp. 61–74, Jan. 1948.
[39] E. J. Candès, "The restricted isometry property and its implications for compressed sensing," *Comp. Rendus Mathematique*, vol. 346, nos. 9–10, pp. 589–592, May 2008.
[40] E. J. Candès and M. B. Wakin, "An introduction to compressive sampling," *IEEE Signal Process. Mag.*, vol. 25, no. 2, pp. 21–30, Mar. 2008.
[41] S. Foucart and H. Rauhut, *A Mathematical Introduction to Compressive Sensing*. Basel, Switzerland: Birkhäuser, 2013.
[42] S. Yu, X. You, W. Ou, X. Jiang, K. Zhao, Z. Zhu, Y. Mou, and X. Zhao, "STFT-like time frequency representations of nonstationary signal with arbitrary sampling schemes," *Neurocomputing*, vol. 204, pp. 211–221, Sep. 2016.
[43] I. Djurović, E. Sejdić, and J. Jiang, "Frequency-based window width optimization for-transform," *AEU-Int. J. Electron. Commun.*, vol. 62, no. 4, pp. 245–250, Apr. 2008.
[44] Y. Huang, J. L. Beck, S. Wu, and H. Li, "Bayesian compressive sensing for approximately sparse signals and application to structural health monitoring signals for data loss recovery," *Probabilistic Eng. Mech.*, vol. 46, pp. 62–79, Oct. 2016.
[45] C. Lu, Z. Lin, and S. Yan, "Smoothed low rank and sparse matrix recovery by iteratively reweighted least squares minimization," *IEEE Trans. Image Process.*, vol. 24, no. 2, pp. 646–654, Feb. 2015.
[46] D. Cohen, K. V. Mishra, and Y. C. Eldar, "Spectral coexistence in radar using Xampling," in *Proc. IEEE Radar Conf. (RadarConf)*, May 2017, pp. 1185–1190.
[47] M. Mishali, Y. C. Eldar, and A. J. Elron, "Xampling: Signal acquisition and processing in union of subspaces," *IEEE Trans. Signal Process.*, vol. 59, no. 10, pp. 4719–4734, Oct. 2011.
[48] P. Flandrin and P. Borgnat, "Time-frequency energy distributions meet compressed sensing," *IEEE Trans. Signal Process.*, vol. 58, no. 6, pp. 2974–2982, Jun. 2010.
[49] E. Sejdić, I. Orović, and S. Stanković, "Compressive sensing meets time–frequency: An overview of recent advances in time–frequency processing of sparse signals," *Digit. Signal Process.*, vol. 77, pp. 22–35, Jun. 2018.
[50] E. Matusiak and Y. C. Eldar, "Sub-Nyquist sampling of short pulses," *IEEE Trans. Signal Process.*, vol. 60, no. 3, pp. 1134–1148, Mar. 2012.
[51] X. Song, Y. Cui, X. Zhang, and K. Liu, "Design of low phase noise frequency synthesizer based on DDS+PLL," *Chin. J. Electron Devices*, vol. 42, no. 4, pp. 947–952, 2019.
[52] S. Rong, J. Yin, and H. C. Luong, "A 0.05-to 10-GHz, 19-to 22-GHz, and 38-to 44-GHz frequency synthesizer for software-defined radios in 0.13-μm CMOS process," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 63, no. 1, pp. 109–113, 2016.
[53] J. M. H. Karel, S. A. P. Haddad, S. Hiseni, R. L. Westra, W. A. Serdijn, and R. L. M. Peeters, "Implementing wavelets in continuous-time analog circuits with dynamic range optimization," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 59, no. 2, pp. 229–242, Feb. 2012.
[54] Y. Tong, Y. He, H. Li, W. Yu, and Y. Long, "Analog implementation of wavelet transform in switched-current circuits with high approximation precision and minimum circuit coefficients," *Circuits, Syst., Signal Process.*, vol. 33, no. 8, pp. 2333–2361, Aug. 2014.

QIANG WANG (Graduate Student Member, IEEE) received the M.S. degree from the Army Engineering University of PLA–Shijiazhuang, Shijiazhuang, China, in 2017, where he is currently pursuing the Ph.D. degree. His current research interests include compressive sensing and analog to information conversion.

CHEN MENG received the Ph.D. degree from the Nanjing University of Science & Technology, Nanjing, China, in 2006. He is currently a Professor with the Missile Engineering Department, Army Engineering University of PLA–Shijiazhuang, Shijiazhuang, China. His current research interests include automatic test systems (ATSs), network for equipment support, and analog to information conversion.

CHENG WANG received the Ph.D. degree from the Ordnance Engineering College, Shijiazhuang, China, in 2006. He is currently a Lecturer with the Missile Engineering Department, Army Engineering University of PLA–Shijiazhuang. His research interests include automatic test systems (ATSs), analog to information conversion, and intelligent signal processing.