RESEARCH **Open Access**

CrossMark

Compressive sampling-based CFO-estimation with exploited features

Chaojin Qing[1]* iD, Jiafan Wang[2], Chuan Huang[3] and Hongyuan Chen[4]

Abstract

Based on the compressed sensing (CS) technique, the carrier frequency offset (CFO) is estimated in compressive sampling scenarios. We firstly confirm the compressibility of estimation metric vector (EMV) of conventional maximum likelihood (ML)-based CFO estimation, and thus conduct the compressive sampling at receiver. By exploiting the EMV features, introducing a circle cluster, and proposing a novel coherence-pattern, we then form a feature-aided weight coherence (FAWC) optimization to optimize measurement-matrix. Besides the proposed FAWC optimization, by referencing compressive sampling matching pursuit (CoSaMP) algorithm and exploiting EMV features, a metric-feature based CoSaMP (MFB-CoSaMP) algorithm is proposed to improve the EMV-reconstruction accuracy, and to reduce computational complexity of classic CoSaMP. With reconstructed EMV, we finally develop a CFO estimation method to estimate the coarse CFO and fine CFO. Relative to weighted coherence minimization (WCM) and classic CoSaMP, the elaborate performance evaluations show that the FAWC and MFB-CoSaMP can independently or jointly improve accuracy of the CFO-estimation (including coarse CFO-estimation and fine CFO-estimation), and the improvement is robust to system parameters, e.g., sparsity level, number of measurements, etc. Furthermore, the mean squared error (MSE) of proposed CFO estimation method can almost reach to its Cramér-Rao lower Bound (CRLB) when a relative large number of measurements, a relative high carrier-to-noise ratio (CNR), and a reasonable length of observed signals can be obtained.

Keywords: Carrier frequency offset, Compressed sensing, Estimation metric, CoSaMP, Equivalent Likelihood function

1 Introduction

The carrier frequency offset (CFO), which is one of the well understood radio frequency (RF) impairment, may result in severe performance degradation at the receiver [1, 2]. To improve receiver performance, the CFO estimation has been studied comprehensively. In [3–5], the CFO estimations for additive white Gaussian noise (AWGN) channels, flat fading channels and frequency-selective fading channels are respectively addressed. Recently, the proposed compressive sensing (CS) approach [6, 7], which enables sub-Nyquist sampling of sparse or compressible signals in some domain, can be employed to reduce system complexity and to save power significantly. By exploiting the sparsity profile, the CS-based CFO estimation is presented in [8, 9] for multi-user uplink. Compared with the CFO estimation without utilizing CS, the estimation

accuracy is improved due to the priori information of sparse approximation. Although various methods of CFO estimation are proposed with and without utilizing CS, the sampling rate of these existing methods for CFO estimation, e.g., [3–5, 8, 9], needs to be at least the Nyquist rate, resulting in excessive power consumption and design difficulty for the analog-to-digital converter (ADC) when the high sampling rate is experienced [10, 11].

To reduce the sampling rate, the CS is introduced into synchronization issue in [12–14]. In [12], a fast and rough estimate of pseudo-noise (PN) code phase and Doppler frequency with a reduced number of parallel correlators (i.e., compressed correlators) is proposed, where the sparse expression is based on autocorrelation. For binary phase-shift keying (BPSK) signals and binary offset carrier (BOC) modulation signals, the 2-D compressed correlator (TDCC) technique for the rough estimate of PN code phase and Doppler frequency is introduced in [13, 14], respectively. Based on hypothesis testing for a code phase

*Correspondence: chjqing@mail.xhu.edu.cn
[1]School of Electrical Engineering and Electronic Information, Xihua University, 610039 Chengdu, China
Full list of author information is available at the end of the article

and Doppler frequency next to the true hypothesis can yield a non-negligible amount of signal energy, the compressed correlators technique in [12–14] tests a compressed hypothesis and coherently combines the signal energy in the neighboring hypotheses. Although the number of correlators is reduced, the compressed correlator technique can only roughly estimate Doppler frequency. Furthermore, the features of estimation metric vector (EMV) of CFO estimation are not exploited for compressive sampling and signal reconstruction. Thus, the CS-based CFO estimation, which includes coarse estimation and fine estimation, is not intensively investigated in [12–14].

By exploiting the features of EMV, a novel CS-based CFO estimation is proposed in this paper. The compressive sampling is introduced into Maximum Likelihood (ML)-Based CFO estimation to reduce sampling rate while holding estimation performance be not deteriorated significantly. We briefly describe some critical points of the proposed CS-based CFO estimation as the follows.

- **Feasibility analysis of compressive sampling:** Based on the compressibility of EMV in ML-based CFO estimation [15], we first verify the received signal can be obtained with compressive sampling, and the ADC requirement can be reduced.
- **Optimization of measurement matrix:** In compressive sampling, measurement matrix directly determines whether the reconstruction can be realized successfully [6, 7]. The designing of efficient measurement matrices becomes the core problem for higher probability of reconstruction. In [16], Baraniuk et al. have been proved that many random matrices are good measurement matrices, and some optimized methods can also be found in existing literatures, such as [17–22]. These existing methods, however, are not specially designed for CFO estimation, and thus cannot obtain an optimized performance (e.g., reconstruction-accuracy improvement for EMV recovery). To obtain a more suitable measurement matrix, we exploit the features of EMV. Firstly, the EMV is expressed as a *circle cluster* to reduce the block-sparsity to one (i.e., significant amplitudes are gathered in one sub-block when EMV is divided into multiple sub-blocks). With the special structure of block-sparsity, a novel coherence-pattern is proposed to fully utilize structure information of circle cluster. Then, a feature-aided weight coherence (FAWC) optimization, based on the algorithm of weighted coherence minimization (WCM) [22], is developed to optimize measurement-matrix without computational complexity increasing.
- **Reconstruction algorithm:** The reconstruction algorithm is another critical factor for successful

reconstruction. With the compressive sampling at receiver, many recovery algorithms are proposed. Among these reconstruction algorithms, we mainly reference the compressive sampling matching pursuit (CoSaMP) [23, 24], due to its high reconstruction accuracy and excellent robustness to noise. According to the CoSaMP algorithm and the EMV features, a metric feature-based CoSaMP (MFB-CoSaMP) algorithm is proposed to improve the EMV reconstruction accuracy, and to reduce the computational complexity of classic CoSaMP.

- **CFO estimation:** With the reconstructed EMV, we implement the CFO estimation by using a two-step procedure which includes coarse and fine CFO estimation. In the coarse of CFO estimation, the likelihood function is constructed according to EMV and to seek the local maximum. As for the fine CFO estimation, the received signal vector is recovered at Nyquist rate with the reconstructed EMV and then used to generate the likelihood function, with which an interpolation method is employed to seek the local maximum near to the value of coarse CFO estimation.

Performance evaluation shows that the proposed CS-based CFO estimation can be implemented with reduced sampling-rate, along with an acceptable estimation deterioration in terms of mean squared error (MSE). Compared with the optimization of weighted coherence minimization (WCM) [22] and the reconstruction algorithm CoSaMP, the elaborate performance evaluations present the proposed FAWC and MFB-CoSaMP can independently or jointly improve the accuracy of CFO-estimation (including coarse CFO-estimation and fine CFO-estimation), and the improvement is robust to system parameters, e.g., the sparsity level, the number of measurement, and the length of the observed signal. Furthermore, the MSE of proposed CFO-estimation can almost reach to its Cramér-Rao lower bound (CRLB) when the reasonable conditions can be obtained.

The main contributions of this paper are summarized as follows.

(a) We confirm the compressibility of CFO EMV in conventional maximum likelihood (ML)-based CFO estimation. Thus, the compressive sampling can be employed for CFO estimation.
(b) A novel FAWC optimization method is proposed by exploiting the features of EMV. Compared with WCM, the proposed FAWC can obtain a suitable measurement matrix for CFO estimation to improve the reconstruction accuracy with comparative computational complexity. Also, the proposed

method is robust to the designed parameters, and can easily to reach its convergence.

(c) An MFB-CoSaMP algorithm is proposed to reconstruct EMV by exploiting its features. Compared with the classic CoSaMP algorithm, the proposed method improves the recovery accuracy and reduces the computational complexity. Furthermore, the improvement of recovery-accuracy is robust against to the varying of parameters.

(d) We implement the CFO estimation (including coarse and fine estimation) with compressive sampling. Furthermore, the MSE performance can reach to its CRLB when reasonable system parameters are obtained.

The rest of this paper is organized as follows. In Section 2, we formulate the method of compressive sampling for CFO estimation, where the expression of sampling is derived from the ML-based approach in a conventional system model of Nyquist rate. Section 3 deals with the optimization of measurement matrix by exploiting EMV features. In Section 4, the CFO estimation method is proposed, where we present the MFB-CoSaMP recovery-method, coarse CFO estimation, and fine CFO estimation. Performance evaluations are shown in Section 5. Finally, Section 6 concludes this paper.

Notation: We use boldface letters to denote matrices and column vectors; **0** denotes the zero vector of arbitrary size; $(\cdot)^T$, $(\cdot)^H$, $(\cdot)^{-1}$, $(\cdot)^{\dagger}$, and $\lfloor \cdot \rfloor$, denote the transpose, conjugate transpose, matrix inversion, Moore-Penrose matrix inversion, and floor operation, respectively; $\mathbf{I}_P$ is $P \times P$ identity matrix; $G(i,j)$ is the (i,j)th element of the matrix of $\mathbf{G}$; we write $\|\cdot\|_p$ for the usual ℓ_p vector norm: $\|\mathbf{x}\|_p = \left(\sum x_i^p\right)^{1/p}$; $\mathrm{supp}\,(\mathbf{x}) = \{i : x_i \neq 0\}$ is the support set that denotes the index set of nonzero elements in $\mathbf{x}$; $\mathbf{\Phi}_T$ denotes the column sub-matrix comprising the T columns of $\mathbf{\Phi}$; $\mathbf{x}\,|_{\mathbf{T}}$ denotes the entries of the vector $\mathbf{x}$ in the set $\mathbf{T}$; the complementary set of set T is denoted by $\mathbf{T}^c$, Ø denotes empty set, and $E\{\cdot\}$ is the expectation operator.

2 Compressive sampling for CFO estimation

According to the conventional ML-based CFO estimation, we verify the feasibility of compressive sampling for CFO estimation in this section. In Subsection 2.1, We briefly describe the method of conventional ML-based CFO estimation. Then, in Subsection 2.2, we raise the compressible EMV, and summarize its features. Based on the compressibility of EMV, the feasibility of compressive sampling for CFO-estimation is verified in Subsection 2.3, according to the derivation result that the received signals (not EMV) can directly conduct compressive sampling.

2.1 Conventional ML-based CFO estimation

From [15], without compressive sampling, the observation of the sampled signal can be expressed as

$$r_k = e^{j(2\pi \Delta f k T_s + \theta)} + v_k, \ 1 \leq k \leq N, \tag{1}$$

where Δf is the frequency offset to be estimated, $T_s \leq 1/(2 \cdot \Delta f)$ is the sampling interval, θ is an unknown random phase with uniform probability density in $[0, 2\pi)$, and v_k is a sample of the complex AWGN with zero mean and variance σ^2. The carrier-to-noise ratio (CNR) ρ, which is the ratio of the signal to noise powers in (1), is defined as [15]

$$\rho \triangleq \frac{1}{2\sigma^2}. \tag{2}$$

In the conventional estimation method [15], the problem of ML estimation of the frequency Δf is to seek the maximum of the equivalent likelihood function

$$\begin{aligned}
\Lambda\left(\Delta\widetilde{f}\right) &\triangleq \left| \sum_{i=1}^{N} r_i e^{-j2\pi \Delta\widetilde{f} \cdot i T_s} \right|^2 \\
&= \sum_{k=1}^{N} \sum_{m=1}^{N} r_k r_m^* e^{-j2\pi \Delta\widetilde{f} T_s(k-m)},
\end{aligned} \tag{3}$$

where $\Delta\widetilde{f}$ is a tentative value for Δf.

2.2 Sparsity of CFO EMV

Assume $\Psi\left(\Delta\widetilde{f}\right) = \sum_{i=1}^{N} r_i e^{-j2\pi \Delta\widetilde{f} \cdot i T_s}$, then the vector form of $\Psi\left(\Delta\widetilde{f}\right)$ can be expressed as

$$\Psi\left(\Delta\widetilde{f}\right) = \mathbf{r}^T \cdot \mathbf{\Gamma}\left(\Delta\widetilde{f}\right), \tag{4}$$

where $\mathbf{r}$ and $\Psi\left(\Delta\widetilde{f}\right)$ are, respectively,

$$\mathbf{r} = [r_1, r_2, \cdots, r_N]^T, \tag{5}$$

and

$$\mathbf{\Gamma}\left(\Delta\widetilde{f}\right) = \left[e^{-j2\pi \Delta\widetilde{f} \cdot T_s}, \cdots, e^{-j2\pi \Delta\widetilde{f} \cdot N T_s} \right]^T. \tag{6}$$

In this paper, we name $\Psi\left(\Delta\widetilde{f}\right)$ as *CFO estimation metric* (EM). From (3), the equivalent likelihood function $\Lambda\left(\Delta\widetilde{f}\right)$ can be rewritten as

$$\Lambda\left(\Delta\widetilde{f}\right) \triangleq \left| \Psi\left(\Delta\widetilde{f}\right) \right|^2 = \left| \mathbf{r}^T \cdot \mathbf{\Gamma}\left(\Delta\widetilde{f}\right) \right|^2. \tag{7}$$

For grid search, P ($P \gtrsim N$) tentative values of Δf, denoted as $\Delta\widetilde{f}_1, \Delta\widetilde{f}_2, \cdots, \Delta\widetilde{f}_P$, are considered. For simplicity, we consider $P = N$ in this paper due to the same conclusions. According to the P tentative values, we form an EMV (denoted by $\widetilde{\mathbf{\Psi}}$) as

$$\widetilde{\mathbf{\Psi}} = \left[\Psi\left(\Delta\widetilde{f}_1\right), \Psi\left(\Delta\widetilde{f}_2\right), \cdots, \Psi\left(\Delta\widetilde{f}_P\right) \right]^T. \tag{8}$$

Substituting $\Psi\left(\Delta\widetilde{f}_p\right) = \mathbf{r}^T\boldsymbol{\Gamma}\left(\Delta\widetilde{f}_p\right), p = 1, 2, \cdots, P$ into (8), then we have

$$
\begin{aligned}
\widetilde{\boldsymbol{\Psi}} &= \left[\mathbf{r}^T\boldsymbol{\Gamma}\left(\Delta\widetilde{f}_1\right), \mathbf{r}^T\boldsymbol{\Gamma}\left(\Delta\widetilde{f}_2\right), \cdots, \mathbf{r}^T\boldsymbol{\Gamma}\left(\Delta\widetilde{f}_P\right)\right]^T \\
&= \left[\boldsymbol{\Gamma}\left(\Delta\widetilde{f}_1\right), \boldsymbol{\Gamma}\left(\Delta\widetilde{f}_2\right), \cdots, \boldsymbol{\Gamma}\left(\Delta\widetilde{f}_P\right)\right]^T\mathbf{r} \\
&= \widetilde{\boldsymbol{\Gamma}}r,
\end{aligned}
\tag{9}
$$

where

$$
\begin{aligned}
\widetilde{\boldsymbol{\Gamma}} &= \left[\boldsymbol{\Gamma}\left(\Delta\widetilde{f}_1\right), \boldsymbol{\Gamma}\left(\Delta\widetilde{f}_2\right), \cdots, \boldsymbol{\Gamma}\left(\Delta\widetilde{f}_P\right)\right]^T \\
&= \begin{pmatrix}
e^{-j2\pi\Delta\widetilde{f}_1\cdot T_s} & e^{-j2\pi\Delta\widetilde{f}_1\cdot 2T_s} & \cdots & e^{-j2\pi\Delta\widetilde{f}_1\cdot NT_s} \\
e^{-j2\pi\Delta\widetilde{f}_2\cdot T_s} & e^{-j2\pi\Delta\widetilde{f}_2\cdot 2T_s} & \cdots & e^{-j2\pi\Delta\widetilde{f}_2\cdot NT_s} \\
\vdots & \vdots & \ddots & \vdots \\
e^{-j2\pi\Delta\widetilde{f}_P\cdot T_s} & e^{-j2\pi\Delta\widetilde{f}_P\cdot 2T_s} & \cdots & e^{-j2\pi\Delta\widetilde{f}_P\cdot NT_s}
\end{pmatrix}.
\end{aligned}
\tag{10}
$$

In (8), the EMV is approximately sparse. That is, for the element-amplitudes of EMV (i.e., $\left|\Psi\left(\Delta\widetilde{f}_1\right)\right|, \cdots, \left|\Psi\left(\Delta\widetilde{f}_P\right)\right|$), only a few amplitudes are significant and the rest are nearly zero or negligible.

The examples are given in Fig. 1 to illustrate the compressibility of EMV, where $N = 64$, $P = 64$, $T_s = 10^{-9}$s,

$\Delta f.T_s \in (-0.5, 0.5)$ is the normalized CFO. Note that we just consider the noise-free case in order to reveal the CFO features obviously. Four cases of normalized CFO, i.e., $\Delta f.T_s = -0.4976$ (near to -0.5), $\Delta f.T_s = -0.2441$ (between -0.5 and 0), $\Delta f.T_s = 0.3241$ (between 0 and 0.5), and $\Delta f.T_s = 0.4757$ (near to 0.5), are given in (a)$-$(d), respectively. From (a)$-$(d) in Fig. 1, and a large number of other experiments, the *intrinsic features* of EMV could be summarized as follows.

(a) Only a few element-amplitudes in EMV are significant.
(b) The significant amplitudes only gather in one cluster when the normalized CFOs connect to a circle from -0.5 to 0.5. In this paper, this cluster in a circle is denominated as *circle cluster* (i.e., significant amplitudes form a cluster in a circle).

Note that the CFO estimation metrics are not in a cluster in a strict sense for the special case that the normalized CFOs are located near to -0.5 (or 0.5). When the normalized CFOs are located near to -0.5 (or 0.5), some significant amplitudes appear near to 0.5 (or -0.5). Thus, we still describe this feature as *cluster* due to its cycle periodicity when the normalized CFOs are connected from -0.5 to 0.5 to a circle. For convenience, we call this
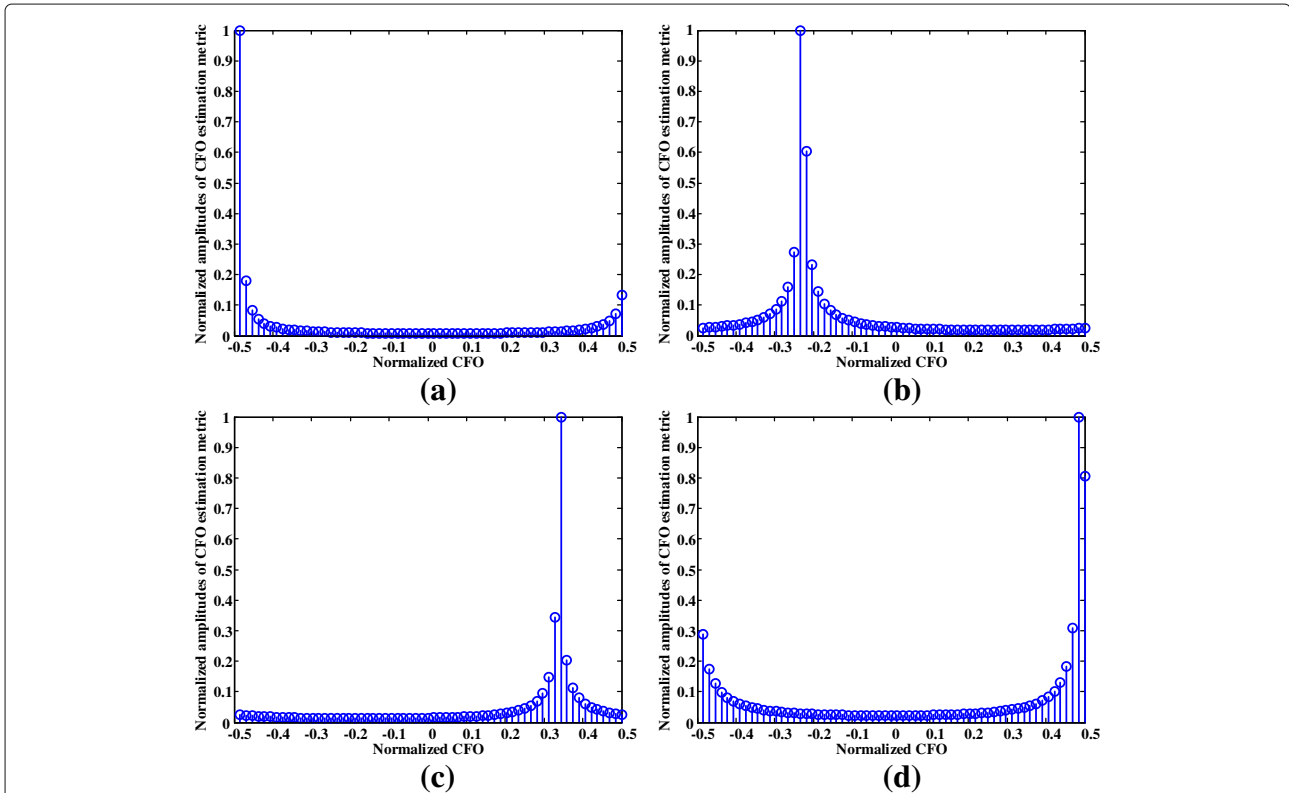
Fig. 1 Examples of the compressibility of CFO estimation metrics for noise-free cases. Where $N = 64$, $P = 64$, $T_s = 10^{-9}$s. **a** $\Delta f.T_s = -0.4976$ (normalized CFO is near to -0.5); **b** $\Delta f.T_s = -0.2441$ (normalized CFO is between -0.5 and 0); **c** $\Delta f.T_s = 0.3241$ (normalized CFO is between 0 and 0.5); and **d** $\Delta f.T_s = 0.4757$ (normalized CFO is near to 0.5)

cluster in this paper as *circle cluster*, i.e., a cluster in a circle.

According to the *intrinsic features*, EMV can be compressed according to the compressed sensing theory [6, 7]. For expression convenience, we also call the *intrinsic features* of EMV as *EMV features* in this paper.

2.3 Feasibility of compressive sampling for CFO estimation

As verified in Subsection 2.1, the CFO EMV, i.e., $\widetilde{\Psi}$, can be compressed. However, the compressibility of CFO EMV $\widetilde{\Psi}$ dose not mean that the received signal $\mathbf{r}$ can conduct compressive sampling, for the reason that the sparsity lies in CFO EMV $\widetilde{\Psi}$ rather than the received signal $\mathbf{r}$. Thus, we need to further analyze the compressibility of CFO EMV $\widetilde{\Psi}$ and whether it could be mapped to the compressive sampling of the receive-signal $\mathbf{r}$.

Based on the compressed sensing theory [6, 7], an $M \times P$ ($M \ll N \le P$) measurement matrix $\mathbf{\Phi}$ can be employed to compress the EMV $\widetilde{\Psi}$ due to its sparsity. Then, an $M \times 1$ measurements, denoted as $\mathbf{y}$, is given by

$$\mathbf{y} = \mathbf{\Phi}\widetilde{\Psi}. \tag{11}$$

Substituting $\widetilde{\Psi} = \widetilde{\Gamma}\mathbf{r}$ (see (9)) into(11), we can derive

$$\mathbf{y} = \mathbf{\Phi}\widetilde{\Gamma}\mathbf{r} = \mathbf{\Theta}\mathbf{r}, \tag{12}$$

where the $M \times N$ matrix $\mathbf{\Theta} = \mathbf{\Phi}\widetilde{\Gamma}$ is defined as *sensing matrix*, and can be expressed as

$$\mathbf{\Theta} = [\mathbf{\Theta}_1, \mathbf{\Theta}_2, \cdots, \mathbf{\Theta}_M]^T, \tag{13}$$

where $\mathbf{\Theta}_m = [\theta_{m1}, \theta_{m2}, \cdots, \theta_{mN}]^T$, $m = 1, 2, \cdots, M$.

Fortunately, the derived expression in (12) can be directly employed to perform the compressive sampling of received signal $\mathbf{r}$ due to its form $\mathbf{y} = \mathbf{r}$. Note that, since M is significantly smaller than N, $\mathbf{y} = \mathbf{\Theta}\mathbf{r}$ infers that $\mathbf{r}$ (not EMV) can be compressed by the $M \times N$ matrix $\mathbf{\Theta}$, i.e., the compressive sampling of received signal $\mathbf{r}$ can be directly conducted. With the sensing matrix $\mathbf{\Theta}$, we can adopt the generic circuit architecture of analog-to-information converter (AIC) [25] or modulated wideband converter (MWC) model [26] to implement compressive sampling. Due to $M \ll N$, the sampling rate can be naturally reduced, i.e., the ADCs at sub-Nyquist rate can be employed for CFO estimation.

After conducting the compressive sampling according to (12), we will use reconstruction-approach to reconstruct the EMV and then perform the CFO estimation based on the reconstructed EMV. Especially, the reconstruction accuracy is mainly decided by the measurement matrix and reconstruction algorithms [6, 7]. Thus, we optimize the measurement-matrix in Section 3 and improve the reconstruction-algorithm in Section 4 for a better reconstruction-accuracy of EMV.

3 Optimization of measurement matrix

In CS theory, the measurement matrix plays an important role to determine the performance of reconstruction [6, 7], because a more efficient measurement matrix for the compressive sampling leads to the higher probability of reconstruction. In [16], Baraniuk et al. have been proved that many random matrices are good measurement matrices. The optimized methods can be found in [17–22]. However, these existing methods are not specially designed for CFO estimation. Thus, the *EMV features* (see Subsection 2.2) are not exploited for the optimizing of measurement matrices. Usually, the CFO EMV appears as *intrinsic features* that only a few element-amplitudes in EMV are significant, and the significant amplitudes gather together to form a *circle cluster*, as depicted in Fig. 1. To optimize measurement matrix $\mathbf{\Phi}$, we exploit the EMV features and propose a FAWC optimization method in this paper.

In [22], the optimization-method, i.e., the WCM, is proposed for a block sparse case. According to the *EMV features*, we can also see that the sparsity of EMV is typically block-sparse case, i.e., nonzero entries in EMV gather in some clusters. Furthermore, when *circle cluster* is introduced, EMV is its special case that the block-sparsity is one. That is, the nonzero entries in EMV occur only in one cluster. Therefore, WCM optimization-method in [22] is mainly referenced herein by the proposed FAWC. The main differences between the proposed FAWC and WCM are given as following:

(a) The *circle cluster* is introduced to reduce the block-sparsity to one with the sub-block-length K (i.e., the sparsity level), while WCM has its sub-block uncertainty. See Fig. 1a, an example that the normalized CFOs are located near to -0.5 is considered. Assuming $K = 7$ (i.e., the amplitudes which are less than 0.05 are treated as ignorable), our FAWC with *circle cluster* has only one subblock with non-ignorable amplitudes and exact number of non-ignorable amplitudes in that subblock (i.e., subblock-length is $K = 7$), while the method in [22] has to consider two subblocks (i.e., the block-sparsity is 2) with non-ignorable amplitudes and the numbers of non-ignorable amplitudes in that two sub-blocks are uncertain. In fact, the actual numbers of non-ignorable amplitudes in Fig. 1a are, respectively, 3 and 4 in the two sub-blocks of non-ignorable amplitudes according to the method in [22]. However, the two sub-blocks have to be considered as owning seven non-ignorable amplitudes to cover all possibility (i.e., the actual numbers of non-ignorable amplitudes maybe 1,2,...,7 in the two sub-blocks).

(b) The concerned patterns of Gram matrix $\mathbf{G}$ (defined in Eq. (15)) are different. For example, the concerned

patterns of WCM and FAWC are given in Fig. 2a and Fig. 2b, respectively. In Fig. 2a, WCM considers three blocks of size 7, and its concerned patterns are based on sub-block coherence. Unlike WCM, the concerned patterns in the proposed FAWC are mainly based on the significant amplitudes. Furthermore, minimizing the sub-block coherence is the main task of WCM in [22], while we minimize the coherence close to the maximum of the significant amplitudes in CFO-estimation metric. The more appropriate minimization of coherence is exploited according to EMV features, which will be verified in the later section.

(c) The measurement matrix is optimized on the basis of complex matrix, rather than optimizing real measurement matrix.

A summary of the proposed FAWC is exhibited in Table 1. Some details of the proposed FAWC are explained as follows.

1). Objective of optimization
According to Eq. (9), the sparse vector $\widetilde{\boldsymbol{\Psi}} = \widetilde{\boldsymbol{\Gamma}}\mathbf{r}$. Then, we have

$$\mathbf{r} = \widetilde{\boldsymbol{\Gamma}}^{\dagger}\widetilde{\boldsymbol{\Psi}} = \mathbf{D}\widetilde{\boldsymbol{\Psi}}, \tag{14}$$

where $\mathbf{D} = \widetilde{\boldsymbol{\Gamma}}^{\dagger}$ is just for expression convenience. Equation (14) indicates that $\mathbf{D}$ can be viewed as a dictionary under the CS framework. Then the Gram matrix of $\mathbf{E} = \boldsymbol{\Phi}\mathbf{D}$ with normalized columns can be expressed as

$$\mathbf{G} = \mathbf{E}^{H}\mathbf{E} = \mathbf{D}^{H}\boldsymbol{\Phi}^{H}\boldsymbol{\Phi}\mathbf{D}. \tag{15}$$

Similar to [22], the optimization objective in this paper, which minimizes the total coherence of the concerned pattern (the red entries in Fig. 2b, denoted

by μ_{C}^{t}), non-concerned pattern (the green entries in Fig. 2b, denoted by μ_{NC}^{t}) and the normalization penalty (denoted by η) of Gram matrix $\mathbf{G}$, is given by

$$\boldsymbol{\Phi} = \arg\min_{\boldsymbol{\Phi}} \left\{ \frac{1}{2}\eta + (1-\alpha)\,\mu_{NC}^{t} + \alpha\mu_{C}^{t} \right\}, \tag{16}$$

where $0 < \alpha < 1$ is a weighting parameter between the total coherence of the concerned pattern and the total coherence of non-concerned pattern. The normalization penalty η, the total coherence of non-concerned pattern μ_{NC} and the total coherence of concerned pattern μ_{C} are defined as

$$\begin{cases} \eta = \sum\limits_{(i,j)\in\boldsymbol{\Omega}_{I}} \left|G\left(i,j\right) - 1\right|^{2} = \sum\limits_{j=1}^{P} \left|G\left(j,j\right) - 1\right|^{2}, \\ \mu_{NC}^{t} = \sum\limits_{(i,j)\in\boldsymbol{\Omega}_{NC}} \left|G\left(i,j\right)\right|^{2}, \\ \mu_{C}^{t} = \sum\limits_{(i,j)\in\boldsymbol{\Omega}_{N}} \left|G\left(i,j\right)\right|^{2}. \end{cases} \tag{17}$$

Where $i = 1, 2, \cdots, P$ and $j = 1, 2, \cdots, P$; $\boldsymbol{\Omega}_{I}$, $\boldsymbol{\Omega}_{NC}$ and $\boldsymbol{\Omega}_{C}$ are the index sets of diagonal entries, non-concerned pattern and concerned pattern of Gram matrix, respectively, (i.e., the index set of the yellow entries, the green entries and the red entries in Fig. 2b). By defining a complete set $\boldsymbol{\Omega} = \{(i,j)|1 \le i \le P, 1 \le j \le P\}$, then we have

$$\begin{cases} \boldsymbol{\Omega}_{I} = \left\{(i,j)\,\middle|\,i = j\right\}, \\ \boldsymbol{\Omega}_{C} = \left\{(i,j)\,\middle|\,|i-j| \le \left\lfloor\frac{K}{2}\right\rfloor,\, i \ne j\right\} \\ \qquad \cup \left\{(i,j)\,\middle|\,|i-j| \ge P - \left\lfloor\frac{K}{2}\right\rfloor,\, i \ne j\right\}, \\ \boldsymbol{\Omega}_{NC} = \boldsymbol{\Omega} - \boldsymbol{\Omega}_{I} - \boldsymbol{\Omega}_{C}. \end{cases} \tag{18}$$

In Eq. (18), $\boldsymbol{\Omega}_{NC}$ is expressed by the difference set of $\boldsymbol{\Omega}$, $\boldsymbol{\Omega}_{I}$ and $\boldsymbol{\Omega}_{C}$.

(a)   (b)
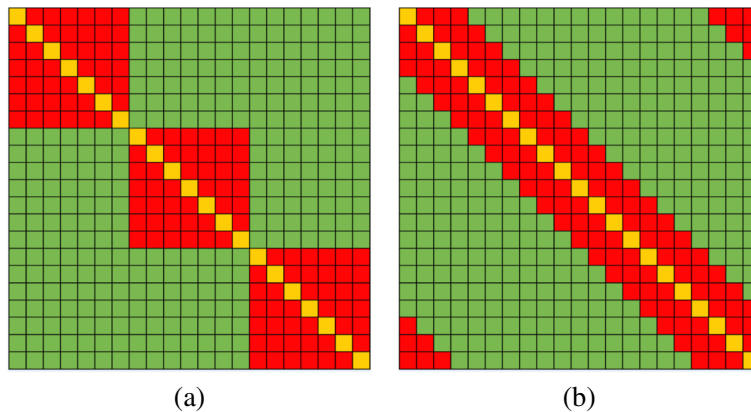
Fig. 2 The difference of the patterns in Gram matrix, where **a** is the patterns in [22] with three blocks of size 7, **b** is the patterns proposed in this paper. The the diagonal entries are in *yellow*, the off-diagonal entries belonging to the non-concerned patterns are in *green*, and the off-diagonal entries belonging to the concerned patterns are in *red*

Table 1 FAWC optimization

Objective: Measurement matrix optimization with

given dictionary **D**, i.e.,

$$\mathbf{\Phi} = \arg\min_{\mathbf{\Phi}} \left\{ \tfrac{1}{2}\eta + (1-\alpha)\,\mu_{NC}^t + \alpha\mu_C^t \right\}.$$

Initialization: Set $n = 0$, and calculate the

eigenvalue decomposition of $\mathbf{DD}^H$, i.e.,

$$\mathbf{DD}^H = \mathbf{U\Lambda U}^H.$$

Then, we calculate the initial value of **Φ**

according to

$$\mathbf{\Phi}^{(0)} = \begin{bmatrix} \mathbf{I}_M & 0 \end{bmatrix} \mathbf{\Lambda}^{-\frac{1}{2}} \mathbf{U}^H.$$

Repeat:

a). Update $\mathbf{G}^{(n)}$ according to $\mathbf{\Phi}^{(n)}$:

$$\mathbf{G}^{(n)} = \left(\mathbf{\Phi}^{(n)}\mathbf{D}\right)^H \mathbf{\Phi}^{(n)}\mathbf{D}.$$

b). Calculate $h_t\left(\mathbf{G}^{(n)}\right)$ according to the

forthcoming Eq. (24), and form the matrix

$$\mathbf{\Upsilon} = \mathbf{\Lambda}^{-\frac{1}{2}}\mathbf{U}^H \mathbf{D} h_t\left(\mathbf{G}^{(n)}\right) \mathbf{U}\mathbf{D}^H \mathbf{\Lambda}^{-\frac{1}{2}}.$$

c). Calculate the eigenvalue decomposition of **Υ**,

and find its M top eigenvalues $\mathbf{\Delta}_M$ and the

corresponding eigenvectors $\mathbf{V}_M$ of **Υ**.

d). Update measurement matrix according to

$$\mathbf{\Phi}^{(n+1)} = \mathbf{\Delta}_M^{\frac{1}{2}}\mathbf{V}_M^H \mathbf{\Lambda}^{-\frac{1}{2}}\mathbf{U}^H.$$

e). $n = n + 1$.

Until: Convergence criterion is satisfied.

---

2). *Initialization of optimization*

Duarte-Carvajalino and Sapiro [27] proposed designing **Φ** by minimizing $\left\| \mathbf{D}^T\mathbf{\Phi}^T\mathbf{\Phi D} - \mathbf{I}_P \right\|_F^2$, which is used to initialize **Φ** in [22] for the algorithm of WCM.

Different from the real dictionary in [22] and [27], the dictionary $\mathbf{D} = \widetilde{\mathbf{\Gamma}}^\dagger$ is a complex matrix due to the complex value of CFO EM. Thus, we initialize **Φ** by minimizing $\left\| \mathbf{D}^H\mathbf{\Phi}^H\mathbf{\Phi D} - \mathbf{I}_P \right\|_F^2$, i.e.,

$$\mathbf{\Phi}^{(0)} = \min_{\mathbf{\Phi}} \left\| \mathbf{D}^H\mathbf{\Phi}^H\mathbf{\Phi D} - \mathbf{I}_P \right\|_F^2. \tag{19}$$

The objective (19) can be solved by using the eigenvalue decomposition (EVD) of $\mathbf{DD}^H$, i.e.,

$$\mathbf{DD}^H = \mathbf{U\Lambda U}^H, \tag{20}$$

where **U** is a unitary matrix, and **Λ** is a real diagonal matrix in which the diagonal entries are the eigenvalues of $\mathbf{DD}^H$. Then, the initial value of **Φ**, denoted by $\mathbf{\Phi}^{(0)}$, can be determined by

$$\mathbf{\Phi}^{(0)} = \begin{bmatrix} \mathbf{I}_M & 0 \end{bmatrix} \mathbf{\Lambda}^{-\frac{1}{2}}\mathbf{U}^H. \tag{21}$$

3). *The nth Iteration of optimization*

According to [22], the value of **Φ** in the nth iteration, i.e., $\mathbf{\Phi}^{(n+1)}$ is given by

$$\mathbf{\Phi}^{(n+1)} = \mathbf{\Delta}_M^{\frac{1}{2}}\mathbf{V}_M^H \mathbf{\Lambda}^{-\frac{1}{2}}\mathbf{U}^H, \tag{22}$$

where **U** and **Λ** can be obtained from eigenvalue decomposition of $\mathbf{DD}^H$ (see (20)); $\mathbf{\Delta}_M$ are the top eigenvalues and $\mathbf{V}_M$ are the corresponding eigenvectors of

$$\mathbf{\Upsilon} = \mathbf{\Lambda}^{-\frac{1}{2}}\mathbf{U}^H \mathbf{D} h_t\left(\mathbf{G}^{(n)}\right) \mathbf{D}^H \mathbf{U}\mathbf{\Lambda}^{-\frac{1}{2}}. \tag{23}$$

In (23), $h_t\left(\mathbf{G}^{(n)}\right)$ is defined as

$$\begin{aligned} h_t\left(\mathbf{G}^{(n)}\right) &\triangleq \tfrac{1}{3}h_\eta\left(\mathbf{G}^{(n)}\right) + \tfrac{2}{3}\alpha h_{\mu_C}\left(\mathbf{G}^{(n)}\right) \\ &+ \tfrac{2}{3}(1-\alpha)\,h_{\mu_{NC}}\left(\mathbf{G}^{(n)}\right). \end{aligned} \tag{24}$$

Where the entries of $h_\eta\left(\mathbf{G}^{(n)}\right)$, $h_{\mu_C}\left(\mathbf{G}^{(n)}\right)$ and $h_{\mu_{NC}}\left(\mathbf{G}^{(n)}\right)$ are defined as

$$\begin{cases} h_\eta\left(\mathbf{G}^{(n)}\right)(i,j) = \begin{cases} 1, & (i,j) \in \mathbf{\Omega}_I \\ \mathbf{G}^{(n)}(i,j), & else \end{cases} \\ h_{\mu_C}\left(\mathbf{G}^{(n)}\right)(i,j) = \begin{cases} \mathbf{G}^{(n)}(i,j), & (i,j) \in \mathbf{\Omega}_C \\ 0, & else \end{cases} \\ h_{\mu_{NC}}\left(\mathbf{G}^{(n)}\right)(i,j) = \begin{cases} \mathbf{G}^{(n)}(i,j), & (i,j) \in \Omega_{NC} \\ 0, & else \end{cases} \end{cases}. \tag{25}$$

For measurement-matrix optimizing, the proposed FAWC satisfies the conditions of surrogate objective of the bound-optimization method. Moreover, its iterative minimization can guarantee the convergence to a local solution. The proofs are abbreviated here for the reasons that the similar proofs can be obtained from Appendix B and Appendix A in [22] for the conditions of bound-optimization method and the convergence, respectively. Similar to [22], the computation complexity of the proposed optimization algorithm is also $O(N^3)$ (same as WCM), due to the application of EVD (the complexity of EVD is N^3). Therefore, the proposed FAWC maintains a comparative computation-complexity with WCM.

4 CFO estimation method

Based on the compressive sampling (see Section 2) and the optimized measurement-matrix (see Section 3), the proposed CFO estimation method first reconstructs EMV. Then, we estimate the coarse CFO by seeking the maximum of the equivalent likelihood function according to the reconstructed EMV. Finally, for the fine CFO estimation, the received signal of Nyquist rate is recovered from the reconstructed EMV, and the likelihood function interpolation locates the local maximum from the result of coarse CFO estimation.

4.1 Sparse reconstruction of EMV

In this subsection, we will present the proposed MFB-CoSaMP reconstruction-method for EMV (i.e., $\widetilde{\Psi}$ recovery. The proposed reconstruction-method mainly exploits *EMV features* as priori information, and thus improves its reconstruction-accuracy. We denote the reconstructed EMV as $\overset{\smile}{\Psi}$ and implement the CFO estimations (including coarse and fine CFO estimation) on the basis of the reconstructed EMV.

Among these currently available CS signal recovery algorithms, our proposed MFB-CoSaMP mainly references the CoSaMP algorithm due to its high reconstruction accuracy and excellent robustness to noise [23, 24]. By further referencing the methodology of model-based CoSaMP [28], an excellent method of support-set identification is developed. The objective of MFB-CoSaMP is to recovery the EMV, i.e., the algorithm output is $\overset{\smile}{\Psi}$. We describe some critical points of MFB-CoSaMP in detail as follows.

A.1 Initialization of MFB-CoSaMP

The input parameters and initialization of MFB-CoSaMP are similar to the CoSaMP algorithm. As for the input parameters, we also need the measurement matrix $\mathbf{\Phi}$, the noisy measurements $\mathbf{y}$ and the sparsity level K. In the initialization step, the initial target vector $\overset{\smile}{\Psi}^{(0)}$ and initial residual $\boldsymbol{v}$ are, respectively, set as a zero vector and $\mathbf{y}$, for the reason that no priori can be obtained.

A.2 Identification based on EMV proxy

Similar to classical CoSaMP algorithm in [23, 24], we form an EMV proxy $\mathbf{u}$ for CFO estimation, i.e,

$$\mathbf{u} = \mathbf{\Phi}^{H}\mathbf{v}, \tag{26}$$

where $\mathbf{\Phi}$ is the measurement matrix optimized in Section 3, and $\mathbf{v}$ is the residual of each iteration. For description convenience, the $P \times 1$ vector $\mathbf{u}$ is expressed as $\mathbf{u} = [u_1, u_2, ..., u_P]^T$. Unlike CoSaMP algorithm in which $2K$ largest components of the proxy $\mathbf{u}$ are located, the MFB-CoSaMP firstly locates the maximal amplitude in $\mathbf{u}$, i.e.,

$$W_1 = \{i : |u_i| = \max\{|u_1|, |u_2|, \cdots, |u_P|\}\}. \tag{27}$$

According to the W_1, we then locate the other $2K-1$ indexes to form support-set $\mathbf{W}_1$. In CoSaMP, the signal components that carry a lot of energy locate in the identification process whereas the EMV features indicate that the significant amplitudes only gather in one circle cluster. Thus, the maximal amplitude is of the special importance to determine the location of circle cluster due to its usual high reliability. Starting from W_1, we then search $2K-1$ indexes nearest W_1 to form a circle cluster. The

identification-result, i.e., the support set $\mathbf{W}_1$, is given by

$$\mathbf{W}_1 = \widetilde{\mathbf{W}}_{11} \bigcup \widetilde{\mathbf{W}}_{12}, \tag{28}$$

where $\widetilde{\mathbf{W}}_{11}$ is defined as

$$\widetilde{\mathbf{W}}_{11} = \begin{cases} \{f_I(W_1 - K)\}, \\ \quad \text{if } |u_{f_I(W_1-K)}| \geq |u_{f_I(W_1+K)}|; \\ \{f_I(W_1 + K)\}, \\ \quad \text{if } |u_{f_I(W_1-K)}| < |u_{f_I(W_1+K)}|. \end{cases} \tag{29}$$

Where $f_I(X)$ is an index-indication function defined as

$$f_I(X) = \begin{cases} P + X, & \text{if } X \leq 0 \\ X, & \text{if } 0 < X \leq P \\ X - P, & \text{other} \end{cases}. \tag{30}$$

In (28), $\widetilde{\mathbf{W}}_{12}$ is determined by the different values of W_1:

$$\widetilde{\mathbf{W}}_{12} = \begin{cases} \{1, \cdots, f_I(W_1 + K - 1)\} \\ \quad \bigcup \{W_1 - K + 1, \cdots, P\}, W_1 > P - K + 1; \\ \{1, \cdots, W_1 + K - 1\} \\ \quad \bigcup \{f_I(W_1 - K + 1), \cdots, P\}, W_1 < K; \\ \{W_1 - K + 1, \cdots, W_1 + K - 1\}, \text{other}. \end{cases} \tag{31}$$

A.3 Support-set merger and metric-vector estimation

After obtaining the identified support-set $\mathbf{W}_1$, we unite the support-set of current approximation $\overset{\smile}{\Psi}^{(k-1)}$ to construct a merger support-set T in the kth iteration, i.e.,

$$\text{T} \leftarrow \text{supp}\left(\overset{\smile}{\Psi}^{(k-1)}\right) \bigcup \mathbf{W}_1. \tag{32}$$

Based on the merged support-set T, a least-square estimation is employed. Denoting $\mathbf{b}$ as $\mathbf{b} = [b_1, b_2, ..., b_P]^T$ and the estimated metric-vector as $\mathbf{b}|_T$, we have

$$\mathbf{b}|_T \leftarrow (\mathbf{\Phi}_T)^{\dagger}\mathbf{y}. \tag{33}$$

Besides the estimated components $\mathbf{b}|_T$, the other components of $\mathbf{b}$ are set as zeros, i.e.,

$$\mathbf{b}|_{T^c} \leftarrow \mathbf{0}. \tag{34}$$

Compared with CoSaMP algorithm, the procedures of support-set merger and metric-vector estimation in the proposed MFB-CoSaMP are similar, just with different support-set T.

A.4 Identification based on EMV

In CoSaMP algorithm, K largest components of the estimated $\mathbf{b}$ are located. By contrast, the MFB-CoSaMP locates the maximal amplitude in $\mathbf{b}$, i.e.,

$$W_2 = \{i : |b_i| = \max\{|b_1|, |b_2|, \cdots, |b_P|\}\}. \tag{35}$$

On the basis of W_2 and the *EMV features*, $K-1$ indexes nearest to W_2 in the circle cluster are searched. The

identification-result, i.e., the support set $\mathbf{W}_2$, is given by

$$\mathbf{W}_2 = \begin{cases} \mathbf{W}_2^{(o)}, & \text{if } K \text{ is odd} \\ \mathbf{W}_{21}^{(e)} \bigcup \mathbf{W}_{22}^{(e)}, & \text{if } K \text{ is even} \end{cases}, \tag{36}$$

where $\mathbf{W}_2^{(o)}$, $\mathbf{W}_{21}^{(e)}$, and $\mathbf{W}_{22}^{(e)}$ are, respectively, given by

$$\mathbf{W}_2^{(o)} = \begin{cases} \left\{ f_I \left(W_2 - \left\lfloor \frac{K}{2} \right\rfloor \right), \cdots, P \right\} \\ \quad \bigcup \left\{ 1, \cdots, W_2 + \left\lfloor \frac{K}{2} \right\rfloor \right\}, W_2 \le \left\lfloor \frac{K}{2} \right\rfloor; \\ \left\{ 1, \cdots, f_I \left(W_2 + \left\lfloor \frac{K}{2} \right\rfloor \right) \right\} \\ \quad \bigcup \left\{ W_2 - \left\lfloor \frac{K}{2} \right\rfloor, \cdots, P \right\}, W_2 > P - \left\lfloor \frac{K}{2} \right\rfloor; \\ \left\{ W_2 - \left\lfloor \frac{K}{2} \right\rfloor, \cdots, W_2 + \left\lfloor \frac{K}{2} \right\rfloor \right\}, \text{other.} \end{cases} \tag{37}$$

$$\mathbf{W}_{21}^{(e)} = \begin{cases} f_I(W_2 - \lfloor K/2 \rfloor), \\ \quad \text{if } \left| u_{f_I(W_2 - \lfloor K/2 \rfloor)} \right| \ge \left| u_{f_I(W_2 + \lfloor K/2 \rfloor)} \right|; \\ f_I(W_2 + \lfloor K/2 \rfloor), \\ \quad \text{if } \left| u_{f_I(W_2 - \lfloor K/2 \rfloor)} \right| < \left| u_{f_I(W_2 + \lfloor K/2 \rfloor)} \right|. \end{cases} \tag{38}$$

and

$$\mathbf{W}_{22}^{(e)} = \begin{cases} \left\{ f_I \left(W_2 - \left\lfloor \frac{K}{2} \right\rfloor + 1 \right), \cdots, P \right\} \\ \quad \cup \left\{ 1, \cdots, W_2 + \left\lfloor \frac{K}{2} \right\rfloor - 1 \right\}, W_2 < \left\lfloor \frac{K}{2} \right\rfloor; \\ \left\{ 1, \cdots, f_I \left(W_2 + \left\lfloor \frac{K}{2} \right\rfloor - 1 \right) \right\} \\ \quad \cup \left\{ W_2 - \left\lfloor \frac{K}{2} \right\rfloor + 1, \cdots, P \right\}, \\ \quad\quad\quad\quad\quad\quad W_2 > P - \left\lfloor \frac{K}{2} \right\rfloor + 1; \\ \left\{ W_2 - \left\lfloor \frac{K}{2} \right\rfloor + 1, \cdots, W_2 + \left\lfloor \frac{K}{2} \right\rfloor - 1 \right\}, \text{other.} \end{cases} \tag{39}$$

In $(37)-(39)$, the index-indication function $f_I(X)$ is defined in Eq. (30).

With the novel support-set $\mathbf{W}_2$, the components of $\mathbf{b}$ which indexes lie in $\mathbf{W}_2$ are reserved, while the others are set as zeros, i.e.,

$$\mathbf{b} \left|_{\mathbf{W}_2^c} \leftarrow \mathbf{0}. \right. \tag{40}$$

A.5) Update of EMV

In the kth iteration, the metric vector $\overset{\smile(k)}{\mathbf{\Psi}}$ should be updated according to $\mathbf{b}$ in (33), (34), and (40). Then, we have

$$\overset{\smile(k)}{\mathbf{\Psi}} \leftarrow \mathbf{b}. \tag{41}$$

With current samples $\mathbf{y}$ and updated metric-vector $\overset{\smile(k)}{\mathbf{\Psi}}$, the residual $\mathbf{v}$ (i.e., the part of the metric-vector that has not been approximated) is replaced by

$$\mathbf{v} \leftarrow \mathbf{y} - \mathbf{\Phi} \overset{\smile(k)}{\mathbf{\Psi}}. \tag{42}$$

After K iterations from *A.2* to *A.5*, the halting criterion of MFB-CoSaMP is satisfied. Therefore, the reconstructed $\overset{\smile}{\mathbf{\Psi}}$ is $\overset{\smile(K)}{\mathbf{\Psi}}$, i.e.,

$$\overset{\smile}{\mathbf{\Psi}} = \overset{\smile(K)}{\mathbf{\Psi}}. \tag{43}$$

A summary of MFB-CoSaMP algorithm is exhibited in Table 2. Compared with the classic CoSaMP, the proposed

Table 2 MFB-CoSaMP algorithm

Input: Measurement matrix $\mathbf{\Phi}$, noisy measurements $\mathbf{y}$, and sparsity level K.

Output: CFO EMV $\overset{\smile}{\mathbf{\Psi}}$

Initial: $\overset{\smile(0)}{\mathbf{\Psi}} \leftarrow \mathbf{0}$, $\mathbf{v} \leftarrow \mathbf{y}$, $k \leftarrow 0$.

Repeat:

a). $k = k + 1$.

b). Form the metric-vector proxy: $\mathbf{u} = \mathbf{\Phi}^H \mathbf{v}$.

c). Identify the circle-cluster location according to $\mathbf{u}$

$W_1 = \{i : |u_i| = \max \{|u_1|, |u_2|, \cdots, |u_P|\}\}$;

$\mathbf{W}_1 \leftarrow$ the 2K indexes nearest to W_1 in index set $\{1, 2, \cdots, P\}$ including W_1.

d). Merge the support set:

$\mathsf{T} \leftarrow \text{supp} \left(\overset{\smile(k-1)}{\mathbf{\Psi}} \right) \bigcup \mathbf{W}_1$.

e). Least square estimation: $\mathbf{b} |_\mathsf{T} \leftarrow (\mathbf{\Phi}_\mathsf{T})^\dagger \mathbf{y}$.

f). $\mathbf{b} |_{\mathsf{T}^c} \leftarrow \mathbf{0}$.

g). Identify circle-cluster location according to $\mathbf{b}$

$W_2 = \{i : |b_i| = \max \{|b_1|, |b_2|, \cdots, |b_P|\}\}$;

$\mathbf{W}_2 \leftarrow$ the K indexes nearest to W_2 in index set $\{1, 2, \cdots, P\}$ including W_2.

h). $\mathbf{b} \left|_{\mathbf{W}_2^c} \leftarrow \mathbf{0} \right.$.

i). Prune to obtain the next approximation:

$\overset{\smile(k)}{\mathbf{\Psi}} \leftarrow \mathbf{b}$.

j). Update current samples $\mathbf{v} \leftarrow \mathbf{y} - \mathbf{\Phi} \overset{\smile(k)}{\mathbf{\Psi}}$.

Until: $k = K$

$\overset{\smile}{\mathbf{\Psi}} \leftarrow \overset{\smile(K)}{\mathbf{\Psi}}$.

MFB-CoSaMP can improve the reconstruction accuracy due to its priori information exploited from the EMV features. In CoSaMP algorithm, the support set $\mathbf{W}_1$ locates 2K largest components of $\{|u_1|, |u_2|, ..., |u_P|\}$. Based on the EMV features that the significant amplitudes gather in a circle cluster, the MFB-CoSaMP first locates the index of maximal amplitude in $\mathbf{u}$ in the $\mathbf{W}_1$ identification and then search the $2K - 1$ indexes nearest to it. The method of $\mathbf{W}_1$ identification is also adopted for $\mathbf{W}_2$ identification, except for the K largest components of $\{|b_1|, |b_2|, ..., |b_P|\}$ with the support set $\mathbf{W}_2$. In MFB-CoSaMP, the maximum of amplitudes is of special importance to determine the location of circle cluster due to its usual highest reliability.

In addition to the accuracy improvement, MFB-CoSaMP can also reduce the computational complexity. The comparison of computational complexity between CoSaMP and MFB-CoSaMP is assessed as follows. Due to the same processing for initialization, support set merger, metric-vector estimating and updating, CoSaMP, and MFB-CoSaMP have the same computational complexity in these procedures. The main differences lie in the procedure of support-set identification, which is presented in *A.2* and *A.4*. In *A.2* (or *A.4*), the CoSaMP locates 2K (or K) largest components of the proxy $\mathbf{u}$ (or $\mathbf{b}$) from

the total $P \times 1$ space. Thus, the classic CoSaMP requires $\sum_{i=1}^{2K}(P-i) + \sum_{i=1}^{K}(P-i)$ real additions in each iteration. In comparison to CoSaMP, MFB-CoSaMP only locates the maximum of $\mathbf{u}$ (or $\mathbf{b}$) in the total $P \times 1$ space in $A.2$ (or $A.4$), and directly chooses the locations of the other $2K - 1$ (or $K - 1$) components whose locations are located nearest to the location of the maximum. Then, MFB-CoSaMP requires $2P$ real additions in each iteration. Obviously, $2P < \sum_{i=1}^{2K}(P-i) + \sum_{i=1}^{K}(P-i)$ for reasonable $K \geq 1$. Therefore, the proposed MFB-CoSaMP reduces the computational complexity compared to the classic CoSaMP.

4.2 Coarse CFO estimation
Represented by the coarse CFO estimation, denoted as $\Delta \widehat{f}_{\text{coarse}}$, the reconstructed EMV, i.e., $\breve{\Psi}$ in (43), could be obtained as
$$\breve{\Psi} = \left[\breve{\Psi}\left(\Delta \widetilde{f}_1\right), \breve{\Psi}\left(\Delta \widetilde{f}_2\right), \cdots, \breve{\Psi}\left(\Delta \widetilde{f}_P\right)\right]^T,$$
then, $\Delta \widehat{f}_{\text{coarse}}$ can be derived as
$$\Delta \widehat{f}_{\text{coarse}} = \arg\max_{\Delta \widetilde{f}_p}\left\{\left|\breve{\Psi}\left(\Delta \widetilde{f}_p\right)\right|^2\right\}, \tag{44}$$

where $p = 1, 2, ..., P$.

4.3 Fine CFO estimation
To implement the fine CFO estimation, we first utilize the reconstructed EMV to recover the received signal $\mathbf{r}$ with Nyquist rate. Then, an interpolation method is employed to construct the equivalent likelihood-function. Finally, we seek the local maximum by using the constructed likelihood-function to estimate the fine CFO.

From (9), we have $\mathbf{r} = \widetilde{\Gamma}^{\dagger}\widetilde{\Psi}$. With the reconstructed EMV (i.e., $\breve{\Psi}$), the received signal $\mathbf{r}$ (sampled with Nyquist rate) can be expressed as
$$\mathbf{r} = \widetilde{\Gamma}^{\dagger}\left(\breve{\Psi} + \mathbf{n}\right) = \widetilde{\Gamma}^{\dagger}\breve{\Psi} + \widetilde{\Gamma}^{\dagger}\mathbf{n}, \tag{45}$$

where $\mathbf{n}$ is the $N \times 1$ noise vector, which is caused by the inaccurate reconstruction and approximate sparsity of $\widetilde{\Psi}$. Then, an approximation of $\mathbf{r}$, denoted as $\breve{\mathbf{r}}$, can be given by
$$\breve{\mathbf{r}} = \widetilde{\Gamma}^{\dagger}\breve{\Psi} = \mathbf{r} - \widetilde{\Gamma}^{\dagger}\mathbf{n}. \tag{46}$$

In (46), if the dominant element-amplitudes in EMV (i.e., $\widetilde{\Psi}$) can be reconstructed accurately and its well sparse-representation can be obtained, the effect of the noise vector $\mathbf{n}$ will be insignificant. Fortunately, with a good recovery-algorithm (e.g., MFB-CoSaMP) and sufficient observations (i.e., relatively large N) at a relatively high CNR, it is feasible to ignore the effect of noise vector $\mathbf{n}$.

With the recovered $\breve{\mathbf{r}}$ and the coarse CFO estimation $\Delta \widehat{f}_{\text{coarse}}$, we estimate the fine CFO (denoted as $\Delta \widehat{f}_{\text{fine}}$) near to $\Delta \widehat{f}_{\text{coarse}}$, where the frequency range for searching $\Delta \widehat{f}_{\text{fine}}$ is assumed to be $\left[\Delta \widehat{f}_{\text{coarse}} - \zeta, \Delta \widehat{f}_{\text{coarse}} + \zeta\right]$ with $\zeta > 0$. Without loss of generality, ζ is chosen as the half search-step of the coarse CFO estimation.

According to Eq. (6), we use the tentative frequency $\breve{\Delta f}$ in $\left[\Delta \widehat{f}_{\text{coarse}} - \zeta, \Delta \widehat{f}_{\text{coarse}} + \zeta\right]$ to construct $N \times 1$ vector $\breve{\Gamma}\left(\breve{\Delta f}\right)$ as

$$\breve{\Gamma}\left(\breve{\Delta f}\right) = \left[e^{-j2\pi\breve{\Delta f}\cdot T_s}, e^{-j2\pi\breve{\Delta f}\cdot 2T_s}, \cdots, e^{-j2\pi\breve{\Delta f}\cdot NT_s}\right]^T. \tag{47}$$

After replacing $\mathbf{r}$ and $\Gamma\left(\widetilde{\Delta f}\right)$ with $\breve{\mathbf{r}}$ (in (46)) and $\breve{\Gamma}\left(\breve{\Delta f}\right)$ (in (47)), respectively, we express the equivalent likelihood function as
$$\Lambda\left(\breve{\Delta f}\right) = \left|\left(\widetilde{\Gamma}^{\dagger}\breve{\Psi}\right)^T \breve{\Gamma}\left(\breve{\Delta f}\right)\right|^2. \tag{48}$$

Thus, the fine CFO estimation $\Delta \widehat{f}_{\text{fine}}$ can be obtained by seeking the maximum of the equivalent likelihood function $\Lambda\left(\breve{\Delta f}\right)$, i.e.,
$$\Delta \widehat{f}_{\text{fine}} = \arg\max_{\breve{\Delta f}}\left\{\left|\left(\widetilde{\Gamma}^{\dagger}\breve{\Psi}\right)^T \breve{\Gamma}\left(\breve{\Delta f}\right)\right|^2\right\}. \tag{49}$$

In (49), the pseudo-inverse $\widetilde{\Gamma}^{\dagger}$ can be computed and stored in advance to save the processing resources during the fine CFO estimation.

5 Performance evaluation
In this section, we will evaluate the performance of proposed methods. For the proposed FAWC, we will evaluate its cost function, recovery performance, and robustness. To evaluate the proposed MFB-CoSaMP, we will consider the reconstruction accuracy and robustness. For their combines, the coarse and fine CFO-estimations are evaluated, respectively.

5.1 Performance of optimized measurement-matrix
To verify the effectiveness of the proposed optimization-method FAWC in Section 3, comparisons against the WCM method in [22] are given in this subsection.

Firstly, we give the evolution of the cost function in (16) (i.e., $\frac{1}{2}\eta + (1-\alpha)\mu_{NC}^t + \alpha\mu_C^t$) in Fig. 3 to observe its convergence behavior, where, $N = 128$, $P = N = 128$, $K = 5$, and $M = N/2 = 64$. Three cases, i.e., $\alpha = 0.1$,
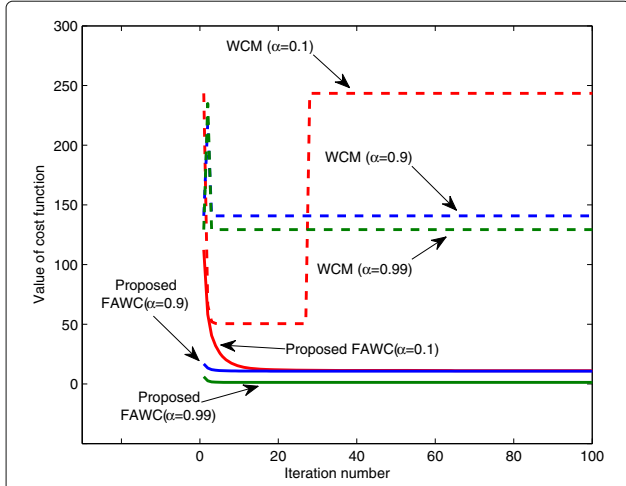
Fig. 3 The cost functions of different construction methods, i.e., the proposed FAWC method and WCM method, where $N = 128$, $P = 128$, $K = 5$, and $M = 64$. Three cases, i.e., $\alpha = 0.1$, $\alpha = 0.9$, and $\alpha = 0.99$ are considered
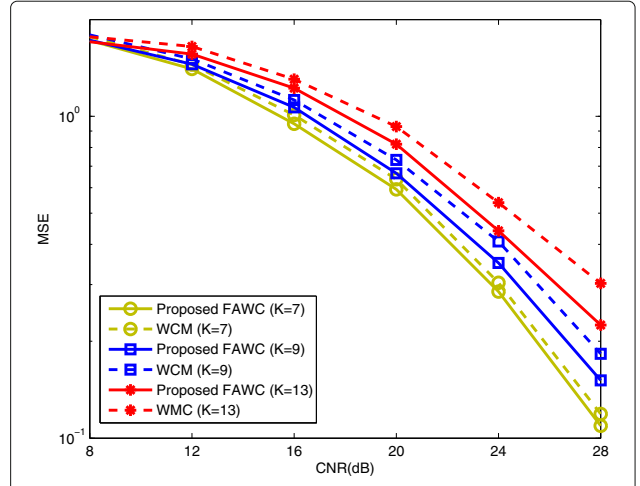

Fig. 4 MSE vs. CNR with different measurement-matrices (optimized by WCM and proposed FAWC, respectively) and sparsity K, where $N = 128$, $P = 128$, $\alpha = 0.9$, and $M = 64$ are considered

$\alpha = 0.9$, and $\alpha = 0.99$ are, respectively, considered. For both FAWC and WCM with a relatively large number of iterations, the increasing α decreases the value of cost function of FAWC. After around 20 iterations, the proposed FAWC has a stable value of cost function, while jumping occurs in WCM. Furthermore, for the given dictionary $\mathbf{D} = \widetilde{\mathbf{\Gamma}}^{\dagger}$ according to the tentative CFOs, FAWC has a smaller value of cost function to capture small coherence of concerned patterns and non-concerned patterns.

Similar to the WCM, $\alpha \approx 1$ is also a good value for the proposed FAWC method. To avoid completely ignoring the coherence of non-concerned patterns, setting $\alpha = 1$ is not considered in the following simulations of this subsection. At the same time, for the sake of fairness, both WCM and proposed FAWC method employ the classic CoSaMP method to reconstruct EMV. Note that we do not adopt the proposed MFB-CoSaMP for the EMV recovery in this subsection. We just expect to actually reveal the improvement from the optimization of measurement matrix, rather than our reconstruction method. We will evaluate the MSE performance, and the MSE in this paper is defined as

$$MSE = E\left\{ \frac{\|\mathbf{X} - \widehat{\mathbf{X}}\|_2^2}{\|\mathbf{X}\|_2^2} \right\}, \qquad (50)$$

where $E\{\cdot\}$ denotes the expectation operator, $\widehat{\mathbf{X}}$ is the estimation of $\mathbf{X}$.

The MSE performance of EMV recovery is given in Figs. 4 and 5, where $N = 128$, $P = N = 128$, $\alpha = 0.9$, and the reconstruction method is classic CoSaMP. Note that the main purpose of introducing the circle-cluster in Section 3 is to solve the uncertainty of sub-block when the

normalized CFO is near to $+0.5$ or -0.5. For other cases with the same CoSaMP recovery algorithm, similar MSE performance can be obtained from the two optimization methods. Thus, in this simulation, the unknown normalized CFO randomly generated is near to $+0.5$ or -0.5. We employ the interval $[\,0.45, 0.5)\bigcup(-0.5, -0.45\,]$ to represent the space near to $+0.5$ or -0.5. The K-sparsity EMV is formed in each simulation by the following procedure: (a) generate a noise-free EMV with a normalized-CFO near to $+0.5$ or -0.5 randomly; (b) find the maximum among EMV element-amplitudes; (c) set the elements in EMV to zeros except the maximum-amplitude element and the other $K - 1$ elements which indexes are the
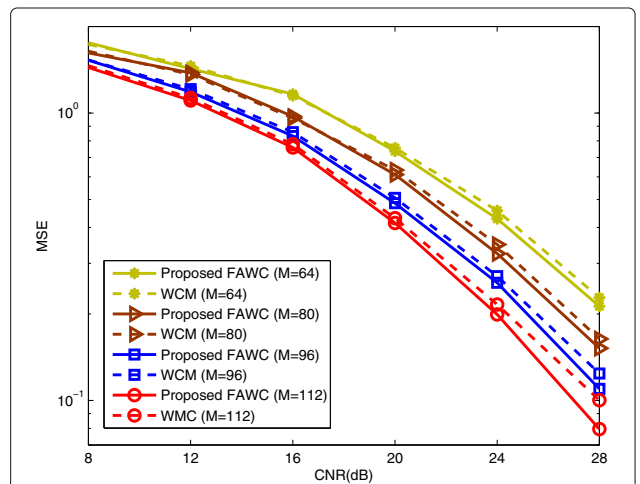
Fig. 5 MSE vs. CNR with different measurement-matrices (optimized by WCM and FAWC, respectively) and M, where $N = 128$, $P = 128$, $\alpha = 0.9$, and $K = 13$ are considered

nearest to the maximum-amplitude element (similar to form the $\mathbf{W}_2$ in *A.4* of Section 4). The formed EMV passes through the noise channel and then generates measurements according to the different measurement matrices optimize by WCM and proposed optimization-method, respectively. Different Ks are adopted while M is kept as $N/2 = 64$ unchanged in Fig. 4. From Fig. 4, the proposed FAWC optimization-method slightly reduces the MSE, compared with WCM. With the increasing K, a much easier distinguishment of MSE improvement can be observed, due to the increasing significance of suitable concerned-patterns with a larger K. The similar conclusions can also be seen in Fig. 5, where different M are adopted while K is kept as 13. From Fig. 5, the proposed FAWC optimization-method slightly improves the MSE performance.

In Fig. 6, M, P, and K vary according to the varying N is simulated, where $M = N/2, P = N, K = \lceil N/10 \rceil, \alpha = 0.9$, and three cases of CNR (i.e., $\rho = 30$ dB, $\rho = 40$ dB, and $\rho = 50$ dB) are considered. From Fig. 6, with the increasing CNR, the improvement of MSE is observed apparently.

Besides coping with the uncertainty of sub-block, the proposed optimization-method can also help to improve the proposed reconstruction method, which can be seen in the later simulations.

5.2 Effectiveness of MFB-CoSaMP

In this subsection, we compare the reconstruction performance of EMV when CoSaMP and MFB-CoSaMP are, respectively, adopted. To really present the merits of MFB-CoSaMP, a Gaussian random matrix [16], which is generated with each entry independently drawn from a Gaussian distribution with zero mean and unit variance, is employed as the measurement matrix for both

algorithms. We do not use the optimized measurement matrices (e.g., the matrix optimized by WCM method or proposed FAWC in this paper) to avoid any improvement brought from the optimization of measurement matrix.

Similar to the MSE evaluation in Subsection 5.1 of this section, the same procedure is adopted to generate K-sparsity EMV and to pass through the noise channel. Then, we employ Gaussian random-matrix to compress EMV and obtain the measurements. To compare the reconstruction performance of CoSaMP and MFB-CoSaMP, Figs. 7, 8, and 9 give the MSE performance. In Fig. 7, the MSE performance with different sparsity Ks are considered (i.e., $K = 7$, $K = 9$, and $K = 13$), and $N = 128$, $P = N = 128$, $M = N/2 = 64$. It can be seen that the proposed MFB-CoSaMP effectively reduces MSE relative to the classic CoSaMP. The similar conclusion can also be derived from the cases of Figs. 8 and 9. In Fig. 8, the MSE comparison with different numbers of measurement, where $N = 128$, $P = N = 128$, $K = 13$, and four cases of measurements (i.e., $M = 64$, $M = 80$, $M = 96$, and $M = 112$) are considered. Particularly, in Fig. 9, the M, P, and K are approximate linear-varying with the change of N, i.e., N varies from 128 to 256, while $M = N/2$, $P = N$, and $K = \lceil N/20 \rceil$. Actually, K does not vary linearly with N, describing its as approximate linear-varying for description convenience. Besides these basic parameters, three cases of CNR, i.e., $\rho = 10$ dB, $\rho = 20$ dB, and $\rho = 30$ dB are considered. Compared with CoSaMP, the MFB-CoSaMP obviously improve the MSE performance in Figs. 8 and 9. Besides the MSE improvement, Fig. 9 also illuminates that more significant improvement can be obtained with the increase of CNR, due to more remarkable EMV features in the higher CNR. The MSE
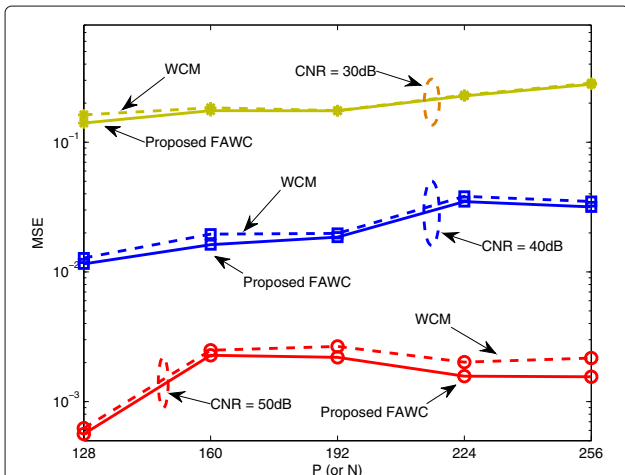
Fig. 6 MSE vs. P (or N) with different measurement-matrices (optimized by WCM and FAWC respectively), where $M = N/2, P = N$, $\alpha = 0.9$, and $K = \lceil N/10 \rceil$ are, respectively, considered
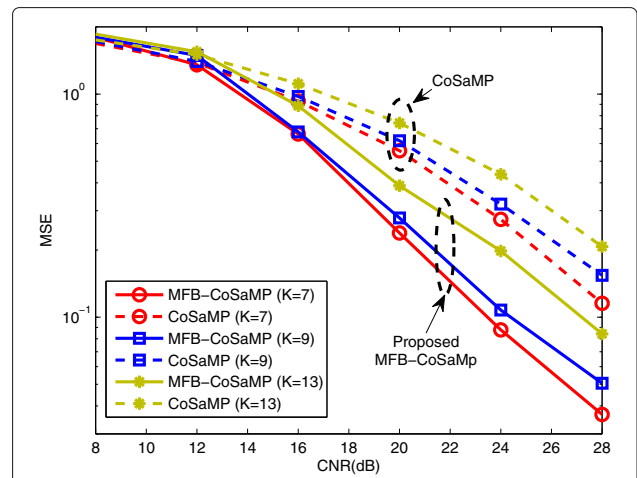
Fig. 7 MSE of EMV reconstruction with different reconstruction algorithms (i.e., CoSaMP and the proposed MFB-CoSaMP) and different sparsity K (i.e., $K = 7$, $K = 9$, and $K = 13$ are, respectively, considered). Where $N = 128$, $P = N = 128$, and $M = N/2 = 64$ are considered
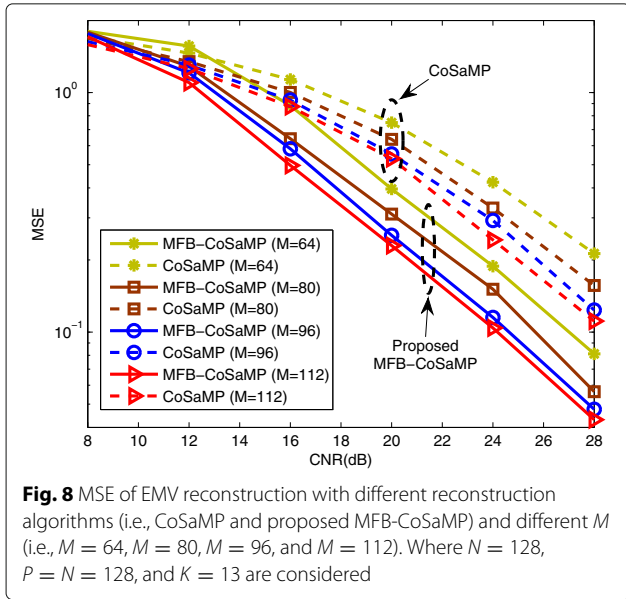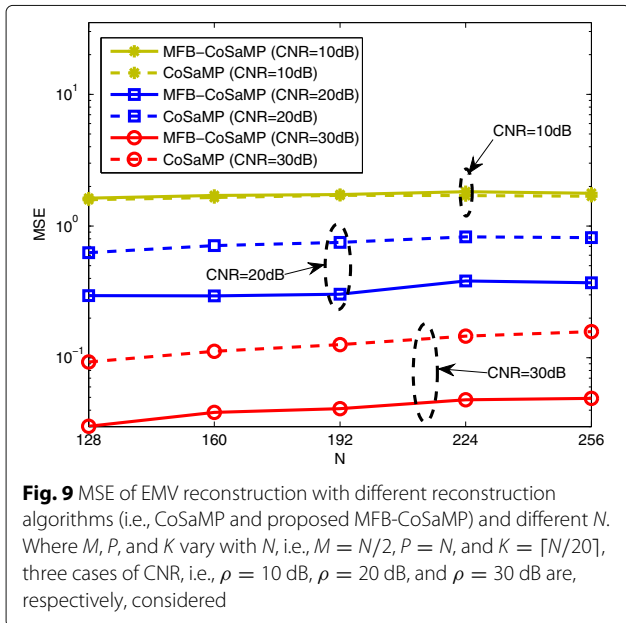
Fig. 8 MSE of EMV reconstruction with different reconstruction algorithms (i.e., CoSaMP and proposed MFB-CoSaMP) and different M (i.e., $M = 64$, $M = 80$, $M = 96$, and $M = 112$). Where $N = 128$, $P = N = 128$, and $K = 13$ are considered

performance improvements in Figs. 7, 8, and 9 are mainly due to the prior-information developed from the EMV features (see step c and g) in Table 2, or see *A.2* and *A.4* in Subsection 4.1 for details).

5.3 Performance of CFO estimation

In this subsection, we discuss the influence of proposed methods in the coarse and fine CFO estimation, respectively. For the convenience of expression, some abbreviations are given as follows.

- "WCM + CoSaMP" denotes that the measurement matrix is optimized by WCM method, and the reconstruction algorithm is CoSaMP.

Fig. 9 MSE of EMV reconstruction with different reconstruction algorithms (i.e., CoSaMP and proposed MFB-CoSaMP) and different N. Where M, P, and K vary with N, i.e., $M = N/2$, $P = N$, and $K = \lceil N/20 \rceil$, three cases of CNR, i.e., $\rho = 10$ dB, $\rho = 20$ dB, and $\rho = 30$ dB are, respectively, considered

- "FAWC + CoSaMP" represents that the measurement matrix is optimized by proposed FAWC optimization method, and the reconstruction algorithm is CoSaMP.
- "WCM + MFB-CoSaMP" denotes that the measurement matrix is optimized by WCM method, and the reconstruction algorithm is MFB-CoSaMP.
- "FAWC + MFB-CoSaMP" represents that the measurement matrix is optimized by FAWC, and the reconstruction algorithm is MFB-CoSaMP.
- "ML (Nyquist Rate)" denotes ML-based coarse CFO estimation with the Nyquist-rate sampling.

Unlike the known sparsity in aforementioned simulations, the sparsity K during the CFO estimation is usually unknown in practical systems. Thus, a sparsity level (i.e., inexact sparsity) is employed in this Section. To obtain a reasonable sparsity level, we use the maximum-amplitude of EMV to set the threshold. The maximum-amplitude can be expressed as $\gamma = \max \left\{ \left| \Psi \left(\Delta \tilde{f}_1 \right) \right|, \left| \Psi \left(\Delta \tilde{f}_2 \right) \right|, \cdots, \left| \Psi \left(\Delta \tilde{f}_P \right) \right| \right\}$, where $\Psi \left(\Delta \tilde{f}_p \right)$, $p = 1, 2, \cdots, P$ is defined in (4). Three thresholds, i.e., $Th_1 = 0.1 \times \gamma$, $Th_2 = 0.05 \times \gamma$, and $Th_3 = 0.01 \times \gamma$, are considered. The amplitude larger than Th_i, $i = 1, 2, 3$, is viewed as the significant amplitude under the threshold Th_i, and the number of significant amplitudes is counted as sparsity-K in each experiment. The 10^5 statistical experiment results are given in Tables 3 and 4, where the ceiling operator is employed to make the mean value and variance be an integer.

From Tables 3 and 4, choosing the moderate threshold Th_2 can commendably cover significant amplitudes and holds a relatively small K for a good reconstruction accuracy, while Th_3 keeps a better approximation at the cost of a bigger K. In this paper, we always consider the case $P = N$. Increasing P or N can make the significant amplitudes of EMV more concentrated and easier to cover with a smaller K. The sparsity levels are given as follows.

- For $N = 128$, we choose $K = 10$ as the sparsity-level according to its mean with Th_2. Because Th_1 is high and only one element of EMV is reserved (according to the mean), Th_3 results in a too big K to ensure the measurements $M \ll N$.
- For $N = 256$, we would like to choose the sparsity level as $K = 23$ according to the mean with Th_3.

Table 3 Mean value of spasity-K with different thresholds

$P \ (= N)$	128	256	512	1024
Mean-value of K (Th_1)	1	1	1	1
Mean-value of K (Th_2)	10	6	4	3
Mean-value of K (Th_3)	61	23	12	7

Table 4 Variance of spasity-*K* with different thresholds

$P(=N)$	128	256	512	1024
Variance of K (Th_1)	0	0	0	0
Variance of K (Th_2)	35	36	23	13
Variance of K (Th_3)	1792	828	503	278

- For $N = 512$ and $N = 1024$, the threshold Th_3 is usually employed, while the mean and variance with Th_3 are simultaneously considered, since N is large enough to cover more significant-amplitudes of EMV. Then, the sparsity levels are, respectively, chosen as $35\,(12 + \lceil \sqrt{503} \rceil)$ and $37\,(7 + \lceil \sqrt{3 \times 287} \rceil)$.
- For simulation convenience, we also choose sparsity-level as $K = \lceil N/10 \rceil$ in some simulations.

C.1 Performance evaluation of coarse CFO estimation

Compared with the WCM optimization, we firstly verify the correct probability of coarse CFO-estimation can be improved by using the proposed FAWC. The correct coarse CFO estimation is defined as

$$\left| \Delta f - \widehat{\Delta f}_{\text{coarse}} \right| T_s \leq \frac{\triangle}{2}, \tag{51}$$

i.e., the offset between the estimated CFO and the real CFO is no more than a half search-step $\triangle$. In this paper, the search-step for coarse CFO estimation is set as $\triangle = 1/P$.

The correct probability of coarse CFO estimation is given in Fig. 10, where $N = 128$, $P = N = 128$, $\alpha = 0.9$, $M = N/2 = 64$, and $K = 10$ (according to Th_2 in Table 3). The measurement matrix is optimized by FAWC and WCM, respectively. As for the unknown normalized CFO, which is randomly generated in $[4.5, +0.5)$ or $(-0.5, -4.5]$ to illuminate the proposed optimization method, can solve the uncertainty of sub-block. Compared with "WCM + CoSaMP", "FAWC + CoSaMP" proves that the proposed measurement matrix can improve the correct probability due to its same recovery algorithm (i.e., CoSaMP). Similar to "WCM + CoSaMP", "WCM + MFB-CoSaMP" shows recovery effectiveness of MFB-CoSaMP since the WCM is utilized in both methods. Although MFB-CoSaMP presents a significant improvement relative to "WCM + CoSaMP", the proposed measurement matrix (i.e., the measurement matrix optimized by the FAWC) can arouse further improvement. Thus, the correct probability of "FAWC + MFB-CoSaMP", which is nearest to that of ML, can obtain the best improvement of coarse CFO estimation. Figure 10 manifests the effectiveness of the FAWC and MFB-CoSaMP, i.e., FAWC and MFB-CoSaMP can independently or jointly improve the correct probability of coarse CFO estimation.
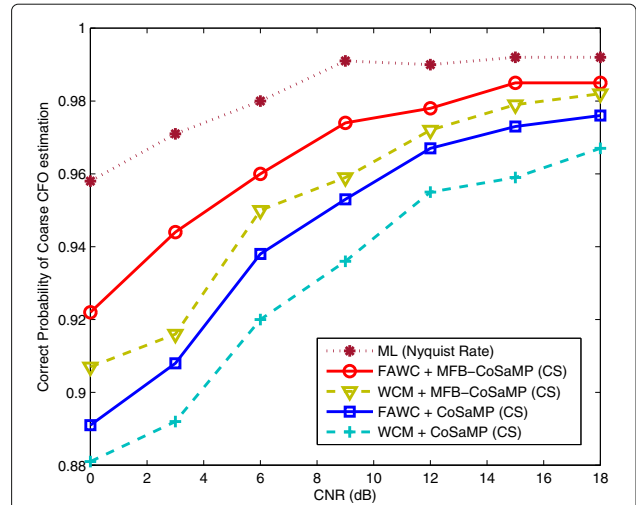

Fig. 10 Correct probability of coarse CFO estimation. Where different measurement-matrices (i.e., constructed by the proposed FAWC method and WCM method), $N = 128$, $P = N = 128$, $\alpha = 0.9$, $K = 10$, and $M = N/2 = 64$ are considered

To elaborate the parameter influence, we, respectively, investigate the influences of K, M, and N in Figs. 11, 12, and 13, where normalized CFO is randomly generated in $[-0.5, +0.5)$. With different K (i.e., $K = 7$, $K = 9$, $K = 13$) and the same other parameters of the simulation in Fig. 10, the correct probability of coarse CFO estimation is given in Fig. 11. Again, we can see that the proposed FAWC and MFB-CoSaMP can jointly improve the correct-probability of coarse CFO estimation, compared with "WCM + CoSaMP". Besides the improvement
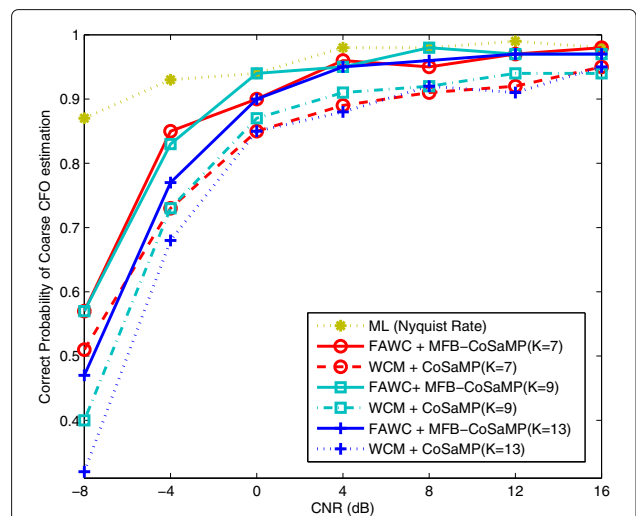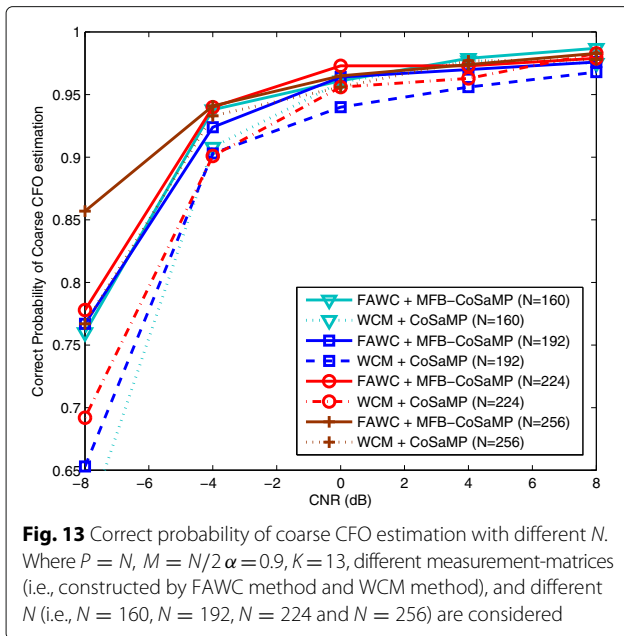

Fig. 11 Correct probability of coarse CFO estimation with different K. Where $N = 128$, $P = N = 128$, $\alpha = 0.9$, $M = N/2 = 64$, different measurement-matrices (i.e., constructed by the FAWC method and WCM method), and different K (i.e., $K = 7$, $K = 9$, $K = 13$) are considered

Fig. 12 Correct probability of coarse CFO estimation with different *M*. Where $N = 128$, $P = N = 128$, $\alpha = 0.9$, $K = 13$, different measurement-matrices (i.e., constructed by FAWC method and WCM method), and different *M* (i.e., $M = 80$, $M = 96$, and $M = 112$) are considered
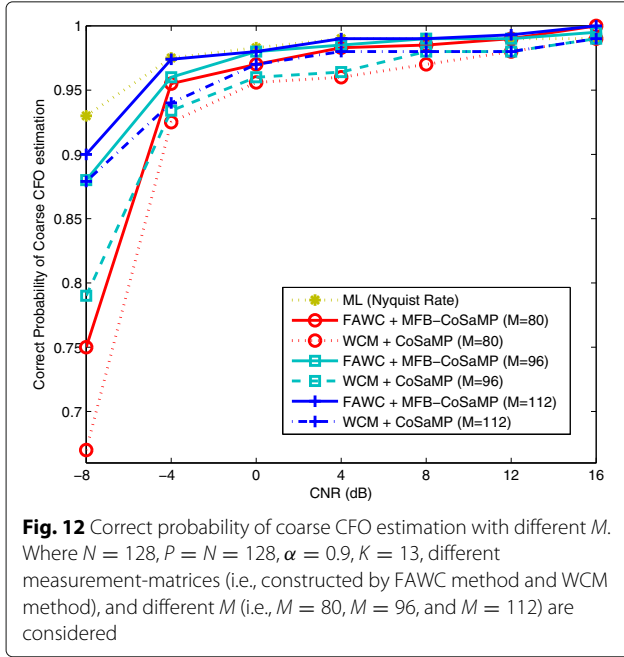
of correct probability, it looks like the smaller *K* (near the sparsity-level $K = 10$) obtains better correct probability when CNR is relatively low (e.g., $\rho \leq -4$ dB). When $\rho \geq -4$ dB, the influence of *K* (near the sparsity-level) is not clear.

During the simulation in Fig. 11, we fix sparsity $K = 13$ (near to 10, and $\lceil N/10 \rceil = \lceil 128/10 \rceil = 13$), change *M*, and keep other parameters the same. The curves of correct probability with different *M* are plotted in Fig. 12, where

Fig. 13 Correct probability of coarse CFO estimation with different *N*. Where $P = N$, $M = N/2$ $\alpha = 0.9$, $K = 13$, different measurement-matrices (i.e., constructed by FAWC method and WCM method), and different *N* (i.e., $N = 160$, $N = 192$, $N = 224$ and $N = 256$) are considered

$N = 128$, $P = N = 128$, $\alpha = 0.9$, $K = 13$, different measurement-matrices (i.e., optimized by the FAWC method and WCM method), and different *M*s (i.e., $M = 80$, $M = 96$, and $M = 112$) are considered. With the increase of *M*, higher correct probability can be obtained, and the improvement is much easier to observe in lower CNR. In Fig. 13, *M*, *P*, and *K* vary with the change of *N*, where $M = N/2$, $P = N$, $\alpha = 0.9$, $K = \lceil N/10 \rceil$. Compared with "WCM + CoSaMP", the "FAWC + MFB-CoSaMP" improve the correct probability of coarse CFO estimation. When CNR is relatively low, e.g., $\rho \leq -4$ dB, the bigger *N* obtains a higher correct probability for both "WCM + CoSaMP" and "FAWC + MFB-CoSaMP", while this rule is not certain for higher CNR. Even so, it is clear that the improvement from "FAWC + MFB-CoSaMP" obviously exists.

According to the aforementioned coarse CFO estimation, our "FAWC + MFB-CoSaMP" can effectively improve its correct probability compared with the conventional "WCM + CoSaMP".

C.2 Performance of fine CFO estimation

Under the compressive sampling scenario, our objective is to obtain a better MSE of fine CFO-estimation, compared with the conventional CS-based method. Furthermore, the MSE performance, which can reach the Cramér-Rao lower Bound (CRLB) of Nyquist rate, is also expected. From [15], the CRLB with Nyquist rate is given by

$$CRLB = \frac{3}{2\pi^2 T_s^2} \cdot \frac{1}{\rho N \left(N^2 - 1\right)}. \tag{52}$$
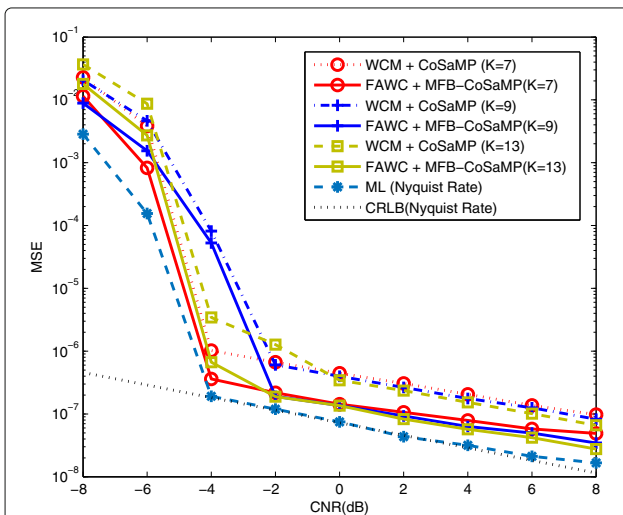
where the CNR ρ is defined in (2).

Firstly, we investigate the MSE performance of fine CFO estimation with different numbers of measurement. The performance evaluation is given in Fig. 14, where $N = 128$ $P = N$, $\alpha = 0.9$, $K = 10$, different measurement matrices (i.e., constructed by the proposed optimization method and WCM method), and different *M* (i.e., $M = 64$, $M = 96$, and $M = 112$) are considered. From Fig. 14, the proposed method, i.e., "FAWC + MFB-CoSaMP", improves the MSE performance, compared with the conventional "WCM + CoSaMP". It is obvious that the increasing *M* makes a better MSE for both "FAWC + MFB-CoSaMP" and "WCM + CoSaMP". Regrettably, sustained increasing of *M* cannot obtain significant MSE-improvement for the relative large *M* (e.g., $M \geq 96$) and relative high CNR (CNR $\rho \geq -2$ dB).

Based on the statistics of sparsity-*K* (listed in Tables 3 and 4), the MSE performance with different *K* is shown in Fig. 15, where $N = 128$ $P = N$, $M = 96$, $\alpha = 0.9$. In addition to the sparsity level $K = 10$, three cases of *K*, i.e., $K = 7$, $K = 9$, and $K = 13$ are, respectively, considered. From Fig. 15, the proposed method,
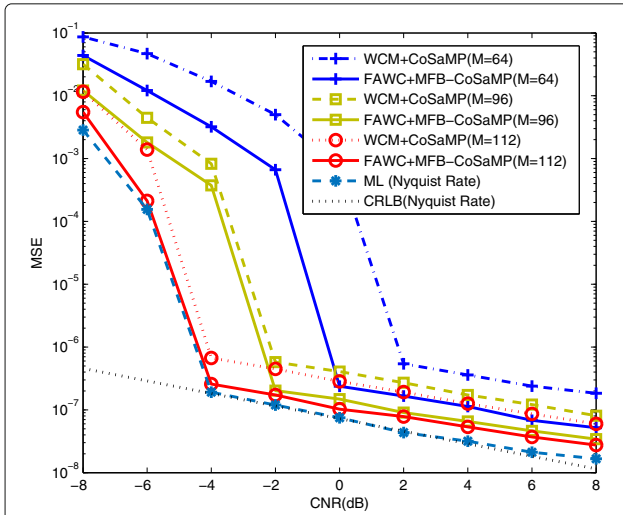
Fig. 14 MSE of fine CFO-estimation with different *M*. Where $N = 128$ $P = N, \alpha = 0.9, K = 10$ (according to Table 3), different measurement-matrices (i.e., constructed by the proposed FAWC method and WCM method), and different *M* (i.e., $M = 64$, $M = 96$, and $M = 112$) are considered

i.e., "FAWC + MFB-CoSaMP", has a better MSE performance than conventional "WCM + CoSaMP". For both "FAWC + MFB-CoSaMP" and "WCM + CoSaMP", the smallest MSE is reached at $K = 7$ under the low CNR (e.g., $\rho \leq -2$ dB). When CNR is high (e.g., $\rho \geq 2$ dB) the smallest MSE is obtained by $K = 13$. That is to say, the higher CNR or the lower noise makes a larger K which is chosen to cover enough significant amplitudes and get a

better EMV approximation. Actually, the influence of the given sparsity-K in Fig. 15 is not significant.

Unlike coarse CFO-estimation that only EMV requires to recover, the fine CFO-estimation usually need the recovered signal **r** at Nyquist rate to construct the equivalent likelihood function (see Subsection 4.1). An approximation of received signal **r** is given in (46). In (46), we require a good enough approximation of EMV, and thus the enough "energy" of EMV could be covered. It looks like a larger K is more effective for a relative high CNR (e.g., in Fig. 15). However, larger K usually results in a worse reconstruction accuracy. From Tables 3 and 4, a larger N can make the "energy" more concentrated in EMV. To balance K and covered "energy", a larger N is a good choice.

To verify the feasibility that the larger N can get the better MSE performance, the simulation is given in Fig. 16, where $P = N, M = \lceil 0.85 \times N \rceil, K = \lceil 0.1 \times N \rceil, \alpha = 0.9$, $N = 128$, $N = 192$, and $N = 256$ are, respectively, considered. As we expected, increasing N can reduce the MSE for both "FAWC + MFB-CoSaMP" and "WCM + CoSaMP", and the proposed "FAWC + MFB-CoSaMP" can obtain a smaller MSE than "WCM + CoSaMP" for each N.

Another phenomenon observed in Fig. 16 is that the larger N, the closer MSE to its CRLB could be obtained. To verify this, an extended simulation is given in Fig. 17, where $P = N, M = \lceil 0.85 \times N \rceil, K = \lceil 0.1 \times N \rceil$, $\alpha = 0.9$, and three cases of N (i.e., $N = 256$, $N = 512$, and $N = 1024$) are considered. It is obvious that the large N can make MSE easily reach the CRLB in spite of the insignificant discrepancy.

Fig. 15 MSE of fine CFO-estimation with different *K*. Where $N = 128$ $P = N, M = 96, \alpha = 0.9$, different measurement-matrices (i.e., constructed by the FAWC method and WCM method), and different *K* (i.e., $K = 7$, $K = 9$, and $K = 13$) are considered
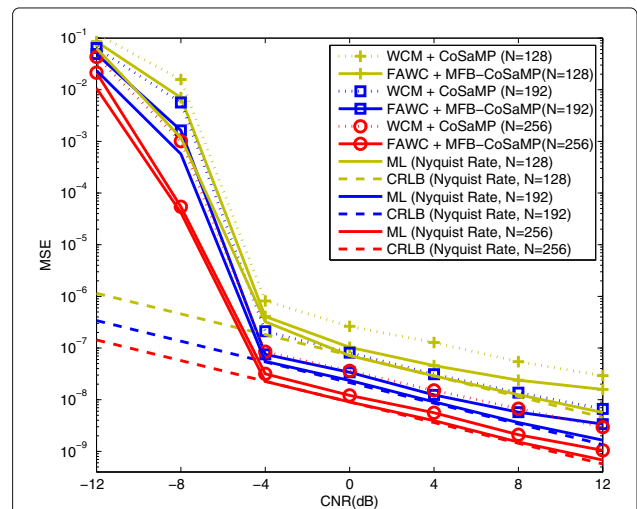
Fig. 16 MSE of fine CFO-estimation with different *N*. Where $P = N$, $M = \lceil 0.85 \times N \rceil, K = \lceil 0.1 \times N \rceil, \alpha = 0.9$, different measurement-matrices (i.e., constructed by the proposed optimization method and WCM method), and different *N* i.e., $N = 128$, $N = 192$, and $N = 256$, are considered

Fig. 17 MSE of fine CFO-estimation with different N. Where $P = N$, $M = \lceil 0.85 \times N \rceil$, $K = \lceil 0.1 \times N \rceil$, $\alpha = 0.9$, different measurement-matrices (i.e., constructed by the proposed FAWC method and WCM method), and different N i.e., $N = 256$, $N = 512$, and $N = 1024$, are considered
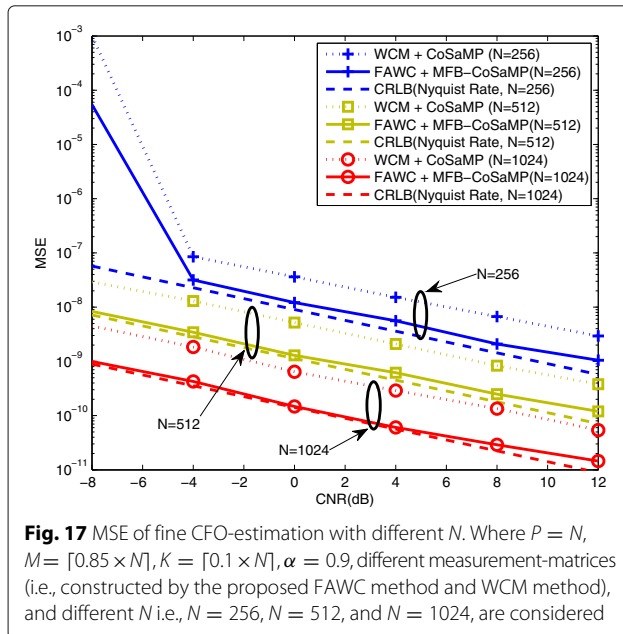
6 Conclusions

In this paper, a preliminary study for CFO estimation based on compressed sensing has been exhibited. We first confirmed that compressive sampling is feasible for ML-based CFO estimation. To solve the number uncertainty of sub-block in block-sparsity CS scenarios, we then introduce the circle cluster, propose a new coherence-pattern, and form an FAWC optimization-method by exploiting the features of EMV. Compared with WMC, the proposed FAWC shows improvements in full-performance evaluations. FAWC can obtain smaller value of cost function to capture small coherence, can obtain a better convergence. Beside the properties of small coherence and good convergence, FAWC effectively solves the uncertainty of sub-block and thus improve the reconstruction accuracy and robustness to sparsity level, receive signal length, and the measurements. Furthermore, based on the EMV features, the MFB-CoSaMP is proposed to boost the support set mergence, improve the reconstruction accuracy, reduce the computational complexity, and hold the improvement robustness against the simulation parameters. Finally, the jointed "FAWC + MFB-CoSaMP" has been verified by the elaborate performance evaluations. For example, the MSE performance which is close to the CRLB is better than that of "WCM + CoSaMP", "WCM + MFB-CoSaMP", or "FAWC + MFB-CoSaMP", while the improvement is robust to the simulation parameters (e.g., sparsity level, number of measurement, and receive signal length).

Authors' contributions
All authors read and approved the final manuscript.

Competing interests
The authors declare that they have no competing interests.

Author details
[1]School of Electrical Engineering and Electronic Information, Xihua University, 610039 Chengdu, China. [2]Department of Electrical and Computer Engineering, Texas A&M University, College Station, 77840 Texas, USA. [3]National Key Laboratory of Science and Technology on Communications, University of Electronic Science and Technology of China, 611731 Chengdu, China. [4]Alps Electric Co., Ltd., 141–8501 Tokyo, Japan.

References
1. L Haring, A Czylwik, M Speth, in *Proc. 2004 International OFDM-Workshop*. Analysis of synchronization impairments in multiuser OFDM systems, (Dresden, 2004), pp. 91–95
2. X Wang, B Hu, A low-complexity ML estimator for carrier and sampling frequency offsets in OFDM systems. IEEE Commun. Lett. **18**(3), 503–506 (2014)
3. M Morelli, U Mengali, Feedforward frequency estimation for PSK: A tutorial review. Eur. Trans. Telecommun. **9**(2), 103–116 (1998)
4. W Kuo, M Fitz, Frequency offset compensation of pilot symbol assisted modulation in frequency flat fading. IEEE Trans. Commun. **45**(11), 1412–1416 (1997)
5. M Morelli, U Mengali, Carrier-frequency estimation for transmissions over selective channels. IEEE Trans. Commun. **48**(9), 1580–1589 (2000)
6. D Donoho, Compressed sensing. IEEE Trans. Inf. Theory. **52**(4), 1289–1306 (2006)
7. E Cands, J Romberg, T Tao, Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information. IEEE Trans. Inf. Theory. **52**(2), 489–509 (2006)
8. P Cheng, Z Chen, Y Guo, L Gui, in *Proc IEEE International Symposium on Personal Indoor and Mobile Radio Communications (PIMRC)*. Distributed Bayesian compressive sensing based blind carrier-frequency offset estimation for interleaved OFDMA uplink, (London, 2013), pp. 801–806
9. J Zhang, K Niu, Z He, in *Proc IEEE International Conference on Communications (ICC)*. Multi-layer distributed Bayesian compressive sensing based blind carrier-frequency offset estimation in uplink OFDMA systems, (Kuala Lumpur, 2016), pp. 1–5
10. J Zhou, M Ramirez, S Palermo, S Hoyos, Digital-assisted asynchronous compressive sensing front-end. IEEE J. Emer. Sel. Topics Circuits Syst. **2**(3), 482–492 (2012)
11. X Chen, E Sobhy, Z Yu, S Hoyos, J Silva-Martinez, S Palermo, B Sadler, A sub-Nyquist rate compressive sensing data acquisition front-end. IEEE J. Emerg. Sel. Topics Circuits Syst. **2**(3), 542–551 (2012)
12. S Kong, A deterministic compressed GNSS acquisition technique. IEEE Trans. Veh. Technol. **62**(2), 511–521 (2013)
13. S Kong, B Kim, Two-dimensional compressed correlator for fast PN code acquisition. IEEE Trans. Wireless Commun. **12**(11), 5859–5867 (2013)
14. B Kim, S Kong, Two-dimensional compressed correlator for fast acquisition of BOC(m, n) signals. IEEE Trans. Veh. Technol. **63**(6), 2662–2672 (2014)
15. M Luise, R Reggiannini, Carrier frequency recovery in all-digital modems for burst-mode transmissions. IEEE Trans. Commun. **43**(234), 1169–1178 (1995)
16. R Baraniuk, M Davenport, R DeVore, M Wakin, A simple proof of the restricted isometry property for random matrices.Constr. Approx. **28**(3), 253–263 (2008)

17. V Abolghasemi, S Ferdowsi, S Sanei, A gradient-based alternating minimization approach for optimization of the measurement matrix in compressive sensing. Signal Proc. **92**(4), 999–1009 (2012)
18. G Li, Z Zhu, D Yang, L Chang, H Bai, On projection matrix optimization for compressive sensing systems. IEEE Trans. Signal Process. **61**(11), 2887–2898 (2013)
19. W Chen, M Rodrigues, I Wassell, Projection design for statistical compressive sensing: a tight frame based approach. IEEE Trans. Signal Process. **61**(8), 2016–2029 (2013)
20. L Zelnik-Manor, K Rosenblum, Y Eldar, Dictionary optimization for block-sparse representations. IEEE Trans. Signal Process. **60**(5), 2386–2395 (2012)
21. N Cleju, Optimized projections for compressed sensing via rank-constrained nearest correlation matrix. Appl. Comput. Harmon. Anal. **36**(3), 495–507 (2014)
22. L Zelnik-Manor, K Rosenblum, Y Eldar, Sensing matrix optimization for block-sparse decoding. IEEE Trans. Signal Process. **59**(9), 4300–4312 (2011)
23. D Needell, J Tropp, Cosamp: iterative signal recovery from incomplete and inaccurate samples. Commun. ACM. **5**(12), 93–100 (2010)
24. D Needell, J Tropp, CoSaMP: iterative signal recovery from noisy samples. Appl. Comput. Harmon. Anal. **26**(3), 301–321 (2008)
25. J Tropp, J Laska, M Duarte, J Romberg, R Baraniuk, Beyond Nyquist: efficient sampling of sparse bandlimited signals. IEEE Trans. Inf. Theory. **56**(1), 520–544 (2010)
26. M Mishali, Y Eldar, Blind multiband signal reconstruction: compressive sensing for analog signals. IEEE Trans. Signal Process. **57**(3), 993–1009 (2009)
27. J Duarte-Carvajalino, G Sapiro, Learning to sense sparse signals: simultaneous sensing matrix and sparsifying dictionary optimization. IEEE Trans. Image Process. **18**(7), 1395–1408 (2009)
28. R Baraniuk, V Cevher, M Duarte, C Hegde, Model-based compressive sensing. IEEE Trans. Inf. Theory. **56**(4), 1982–2001 (2010)