CrossMark

ORIGINAL PAPER

Parameter estimation of multi-component chirp signals based on discrete chirp Fourier transform and population Monte Carlo

Peng Yang · Zheng Liu · Wen-Li Jiang

Abstract In this paper, a parameter estimation method for multi-component chirp signals with white Gaussian noise is proposed based on the modified discrete chirp Fourier transform (MDCFT) and the population Monte Carlo (PMC) methodology, in which the model order is unknown. By utilizing the integrability of linear parameters in the Bayesian model, this paper considers the posterior distribution of non-linear parameters. MDCFT was adopted to calculate the chirpogram of the observed data, and clear peaks can be detected in the discrete chirp Fourier transform domain. The importance function (IF) was constructed according to the peaks, and the PMC algorithm was employed to evaluate the posterior distribution. The proposed method cannot only use the selected IF to generate the sample fitting target function in the parameter region of interest, but can also utilize samples and importance weights to update the IF adaptively. The simulation results indicated that the proposed method can realize joint Bayesian model selection and parameter estimation of multi-component chirp signals. Compared with the two existing methods based on Monte Carlo methodology, the proposed method exhibits improved performance.

Keywords Multi-component chirp signals · Parameter estimation · Model selection · Bayesian inference · Modified discrete chirp Fourier transform (MDCFT) · Population Monte Carlo (PMC)

P. Yang (✉) · Z. Liu · W.-L. Jiang
College of Electronic Science and Engineering,
National University of Defense Technology,
Changsha 410073, Hunan, People's Republic of China
e-mail: yp2008zn@163.com

Z. Liu
e-mail: liuzheng@nudt.edu.cn

W.-L. Jiang
e-mail: jiangwl@nudt.edu.cn

1 Introduction

Chirp signals are widely used in radar and communication applications [1,2]. For example, the Doppler frequency of a moving object in the radar system is approximately proportional to the speed of the object, and the echo of the object under uniform accelerated motion is regarded as a chirp signal. The status and characteristics of a moving object can be reflected by the chirp signal parameters. Frequency and chirp rate are the prerequisites to obtain the speed and acceleration of the moving object. According to the amplitudes, the size and reflection characteristics of the object can be known. Therefore, the proper utilization of the limited observed data to estimate the chirp signal parameters accurately is significant.

Numerous researchers have explored the parameter estimation problem of chirp signals from different aspects. Traditional methods include algorithms based on the principal of maximum likelihood estimation (MLE) [3,4], discrete polynomial-phase transform (DPT) [5,6], time–frequency transform [7,8], and sparse signal reconstruction [9–11]. Most existing methods are based on the MLE principle, the estimation accuracy of which can be close to the Cramer–Rao low bound (CRLB). However, the computational load of such methods is large and cannot address multi-component chirp signals. The DPT-based methods utilize the instantaneous phase or instantaneous frequency to perform parameter estimation in a simple and efficient manner. However, such methods require a high signal-to-noise ratio (SNR). The idea behind time–frequency transform is to transform the detection and estimation problem of chirp signals into a line search problem of the time–frequency plane and then simplify the problem into a two-dimensional peak search problem through Radon or Hough transform. However, this method is limited by the calculation load, resolution of

🍇 Springer

time–frequency, and cross-term interference. Sparse reconstruction is based on the idea that signals can be discomposed into a linear expansion of chirplets while utilizing matching pursuit (MP) to identify the chirplet in the signals [11]. The obtained chirplet can be directly used to estimate the chirp signal parameters. However, this method also has several limitations. First, in the MP process, one chirp component can be decomposed into multiple chirplets. Second, a huge dictionary is adopted, the computation load of which is not applicable in practice.

Researchers have applied Bayesian inference to the parameter estimation problem of multi-component signals [12–15]. A numerical method is typically used to evaluate the joint posterior distribution of unknown parameters. Some researchers [14,15] have proposed one solution based on the Markov chain Monte Carlo (MCMC) method. This method uses the chirpogram generated from the discrete chirp Fourier transform (DCFT) to select initial parameters. The MCMC method combined with the random walk Metropolis–Hastings algorithm and the Gibbs sampler is utilized to calculate the posterior distribution of unknown parameters. However, this method needs a burn-in period, during which samples need to be abandoned. Moreover, the MCMC method is difficult to implement in a parallel manner, which significantly reduces its practicability. Researchers [16] have also combined importance sampling (IS) with MLE, which can deal with the parameter estimation problem of multi-component chirp signals. The estimation accuracy of this method can be close to that of the CRLB, but its computation load is huge, and its estimation precision is affected by the "picket fence" effect.

Several methods are based on adaptive importance sampling (AIS). AIS, also called as the population Monte Carlo (PMC) [17,18], is an adaptive learning method, which can adaptively select and adjust the importance function (IF) according to the collected information. As a result, samples can be drawn around the true value with a greater possibility and the approximation to the object distribution can be obtained. In [17], a fixed number of preselected transition kernels were used and each of these kernels was assigned different weights at different iterations. In [19], the performance of the PMC method was improved through Rao–Blackwellization. A combination of the Gibbs sampler and PMC was proposed in [20]. This method first uses local maximum peaks in the chirpogram to construct the IF and then generates samples and uses resampling technology to prevent sample degeneration. Finally, the method iterates the above process and estimates the parameters and target function of multi-component chirp signals.

The performance of the AIS method is closely related to the selection of the IF. An inaccurate IF does not necessarily lead to the failure of the method but will significantly degrade performance. For chirp signals, DCFT has

good detection performance. However, the detected signals need to satisfy specific conditions [21]; otherwise, the detection performance of DCFT will degrade significantly. Thus, this paper adopts the modified discrete chirp Fourier transform (MDCFT) method [22–24] to detect peaks and construct IFs. The PMC is then used to estimate the parameters of multi-component chirp signals. In the iteration process of the PMC, on the one hand, samples are drawn from the IF constructed in the former iteration process; on the other hand, IF will be updated by the samples and their weights, and this IF will be used to draw samples in the next iteration. The whole iteration process can be stopped at any time. This method cannot only estimate the signal parameters but also the posterior distribution of unknown parameters. In addition, when model order is unknown, this method can realize the joint Bayesian model selection and parameter estimation of multi-component chirp signals.

The remainder of this paper is organized as follows: Section 2 mainly describes the mathematical model of multi-component chirp signals and the posterior distribution in the Bayesian framework. In Sect. 3, the idea and procedure of the proposed method is analyzed and the implementation steps and performance analysis are discussed. Section 4 verifies the performance of the proposed method through simulation experiments. Finally, Sect. 5 concludes the paper.

2 Problem description and theoretical background

2.1 Signal model

Let $\mathbf{y} = [y[0], y[1], \ldots, y[N-1]]^{\mathrm{T}}$ denote the observed vector of N data samples. $\mathbf{y}$ is generated from one of the following K models:

$$M_0 : y[n] = \varepsilon[n]$$
$$M_K : y[n] = \sum_{k=1}^{K} \left[A_k \exp(j2\pi W_k[n]) \right] + \varepsilon[n], \quad (1)$$

where $n = 0, \ldots, N-1$, K denotes the model order, and $W_k[n] = f_k n + \frac{s_k}{2} \left(n - \frac{N-1}{2} \right)^2$. f_k, s_k, and A_k, respectively, denote the center frequency, chirp rate, and complex amplitude of the kth chirp component. $\varepsilon[n]$ denotes the additive white Gaussian noise with zero mean and unknown variance σ_ε^2.

Equation (1) can be written into the following matrix form:

$$M_0 : \mathbf{y} = \boldsymbol{\varepsilon}$$
$$M_K : \mathbf{y} = \mathbf{D}\boldsymbol{\alpha} + \boldsymbol{\varepsilon}, \quad (2)$$

where $\boldsymbol{\varepsilon} = [\varepsilon[0], \varepsilon[1], \ldots, \varepsilon[N-1]]^{\mathrm{T}}$ is a noise vector; $\boldsymbol{\alpha} = [A_1, A_2, \ldots, A_K]^{\mathrm{T}}$ is a complex amplitude vector; and the $N \times K$ matrix $\mathbf{D}$ is defined asfollows:

$$[D]_{n+1,k} = \exp\left\{j2\pi\left[f_k n + \frac{s_k}{2}\left(n - \frac{N-1}{2}\right)^2\right]\right\},$$

where $n = 0, \ldots, N-1$, $k = 1, \ldots, K$.

The parameter estimation problem of multi-component chirp signals can be described as follows: Assuming parameter set $\boldsymbol{\Theta} \triangleq \left[\boldsymbol{\alpha}^{\mathrm{T}}, \boldsymbol{f}^{\mathrm{T}}, \boldsymbol{s}^{\mathrm{T}}, \sigma_\varepsilon^2, K\right]^{\mathrm{T}}$ is unknown, the objective is to estimate model order K, center frequency $\boldsymbol{f} = [f_1, f_2, \ldots, f_K]^{\mathrm{T}}$ and chirp rate $\boldsymbol{s} = [s_1, s_2, \ldots, s_K]^{\mathrm{T}}$ from the data y.

2.2 Bayesian model and aims

In the Bayesian model, data and signal parameters can be regarded as random variables. According to Eq. (2), the likelihood of $\boldsymbol{y}$ is assumed as:

$$p\left(\boldsymbol{y}|\boldsymbol{\Theta}\right) = \frac{1}{\left(\pi\sigma_\varepsilon^2\right)^N}\exp\left\{-\frac{1}{\sigma_\varepsilon^2}\left(\boldsymbol{y} - \boldsymbol{D}\boldsymbol{\alpha}\right)^{\mathrm{H}}\left(\boldsymbol{y} - \boldsymbol{D}\boldsymbol{\alpha}\right)\right\}. \quad (3)$$

Using Bayesian methodology $p\left(\boldsymbol{\Theta}|\boldsymbol{y}\right) = p\left(\boldsymbol{y}|\boldsymbol{\Theta}\right)p\left(\boldsymbol{\Theta}\right)/p\left(\boldsymbol{y}\right)$, the posterior distribution of the unknown parameter set $\boldsymbol{\Theta}$ can be expressed as follows:

$$\begin{aligned}p\left(\boldsymbol{\Theta}|\boldsymbol{y}\right) &\propto p\left(\boldsymbol{y}|\boldsymbol{\Theta}\right)p\left(\boldsymbol{\Theta}\right)\\&\propto p\left(\boldsymbol{y}|\boldsymbol{\alpha}, \boldsymbol{f}, \boldsymbol{s}, \sigma_\varepsilon^2, K\right)p\left(\boldsymbol{\alpha}, \boldsymbol{f}, \boldsymbol{s}, \sigma_\varepsilon^2, K\right).\end{aligned} \quad (4)$$

The Bayesian inference has to assume the prior of the unknown parameters reasonably and construct the following form [12,13,18]:

$$\begin{aligned}&p\left(\boldsymbol{\alpha}, \boldsymbol{f}, \boldsymbol{s}, \sigma_\varepsilon^2, K\right)\\&= p\left(\boldsymbol{\alpha}|\boldsymbol{f}, \boldsymbol{s}, \sigma_\varepsilon^2, K\right)p\left(\boldsymbol{f}, \boldsymbol{s}|K\right)p\left(\sigma_\varepsilon^2\right)p\left(K\right).\end{aligned} \quad (5)$$

The amplitude vector $\boldsymbol{\alpha}$, center frequency vector $\boldsymbol{f}$, and the chirp rate vector $\boldsymbol{s}$ determine the components of each chirp. The amplitude vector $\boldsymbol{\alpha}$ and the additive noise level σ_ε^2 are closely related since they both determine the amplitude of the observed signal as well as the SNR. Thus, $\boldsymbol{\alpha}$ is closely associated with $\boldsymbol{f}, \boldsymbol{s}, \sigma_\varepsilon^2$, and K. The prior of $\boldsymbol{\alpha}$ can be expressed as $p\left(\boldsymbol{\alpha}|\boldsymbol{f}, \boldsymbol{s}, \sigma_\varepsilon^2, K\right)$. This paper regards the center frequency and chirp rate of each chirp component as a set of parameters to draw samples. Considering their close correlation, the prior of $\boldsymbol{f}$ and $\boldsymbol{s}$ is expressed as $p\left(\boldsymbol{f}, \boldsymbol{s}|K\right)$.

Referring to the selected parameter priors in [12,13,18], the prior of each parameter is assumed as

$$p\left(\boldsymbol{\alpha}|\boldsymbol{f}, \boldsymbol{s}, \sigma_\varepsilon^2, K\right) = N\left(\boldsymbol{\alpha}; 0, \delta^2\sigma_\varepsilon^2\left(\boldsymbol{D}^{\mathrm{H}}\boldsymbol{D}\right)^{-1}\right), \quad (6)$$

$$p\left(\sigma_\varepsilon^2\right) = \mathrm{IG}\left(\sigma_\varepsilon^2; \upsilon_0, \gamma_0\right), \quad (7)$$

where δ^2 denotes an expected SNR of the observed signals; $\boldsymbol{\alpha}$ is subject to maximum entropy Gaussian distribution of zero mean; $N\left(\boldsymbol{x}; \boldsymbol{m_x}, \boldsymbol{\Sigma_x}\right)$ indicate that $\boldsymbol{x}$ obeys the multivariate

Gaussian distribution, the mean of which is $\boldsymbol{m_x}$ and covariance matrix is $\boldsymbol{\Sigma_x}$; and IG $(x; \alpha, \beta)$ denotes that $\boldsymbol{x}$ obeys the inverse Gamma distribution, the hyper-parameters of which are (α, β). Assuming that the prior probability of each model is equal, $p\left(K\right) = 1/K$. The normalized center frequencies satisfy $0 < f_1 < f_2 < \cdots < f_K < 1$, and the normalized chirp rates satisfy $0 \le s_k \le 2$. Therefore, we can obtain $p\left(\boldsymbol{f}, \boldsymbol{s}|K\right) = (1/2)^K$.

Based on this assumption, we can derive the posterior distribution of K, $\boldsymbol{f}$, and $\boldsymbol{s}$ on $\boldsymbol{y}$ as follows (please refer to "Appendix 1" for the detailed deviation):

$$p\left(\boldsymbol{f}, \boldsymbol{s}, K|\boldsymbol{y}\right) \propto \left(\gamma_0 + \boldsymbol{y}^{\mathrm{H}}\boldsymbol{P}\boldsymbol{y}\right)^{-(N+\upsilon_0)} \times \frac{\left(1 + \delta^2\right)}{K \cdot 2^K}. \quad (8)$$

The posterior distribution of Eq. (8) is highly nonlinear and therefore cannot be expressed in a closed form. Section 3 will describe the feasibility of approximation to the posterior using the proposed method.

2.3 DCFT and MDCFT

This subsection describes the DCFT and MDCFT and explains the advantage of the latter over the former.

The DCFT for discrete chirp signals with length N is defined as

$$X_c\left(k, l\right) = \frac{1}{\sqrt{N}}\sum_{n=0}^{N-1}x\left(n\right)W_N^{kn+ln^2}, \quad 0 \le k, l \le N-1, \quad (9)$$

where k and l, respectively, represent the frequency and chirp rate; $W_N = \exp\left(-j2\pi/N\right)$; $x\left(n\right)$ can represent $x\left(n\right) = \exp\left[\frac{j2\pi}{N}\left(k_0 n + l_0 n^2\right)\right] = W_N^{-(k_0 n + l_0 n^2)}$; k_0 and l_0 are both integers; and $0 \le k_0, l_0 \le N-1$.

The literature [21] introduces the basic properties of $X_c\left(k, l\right)$. Given that DCFT is a linear transform, it will not give rise to the cross-term interference of Wigner–Ville distribution. This method can deal with multi-component chirp signals. The multiple peaks in the DCFT domain, respectively, correspond to each target signal. However, this method requires signals to satisfy specific conditions [21,25]. For example, the value N should be a prime number, and the chirp rate should be an integer (that is, the value should locate on the divided lattice point). Otherwise, the peak in the chirpogram calculated by the DCFT will be unclear, which may result in a sharp decline in detection performance.

Specific to this shortcoming, this paper employs MDCFT to perform peak detection. MDCFT is defined as

$$X_M\left(k, l\right) = \frac{1}{\sqrt{N}}\sum_{n=0}^{N-1}x'\left(n\right)W_N^{kn+(l/N)n^2}, \quad 0 \le k, l \le N-1, \quad (10)$$

where $x'\left(n\right) = W_N^{-(k_0 n + (l_0/N)n^2)}$.

Compared with DCFT, MDCFT reduces N times of the search step length of the chirp rate and its estimation range turns into $k' \in (0, 1, \ldots, N-1)$, and $l' \in (0, 1/N, \ldots, N-1/N)$, that is, $k' = k$ and $l' = l/N$. Therefore, MDCFT increases the estimation accuracy of the chirp rate, detects the adjacent peaks with smaller interval, and weakens the "picket fence" effect. In addition, the MDCFT has no limitation of the signal length.

In the parameter estimation problem of multi-component chirp signals, frequency and chirp rate do not necessarily lie in the divided lattice points, and the signal length is not necessarily equal to prime. Thus, MDCFT is more suitable for transforming the signal into the DCFT domain. Considering that the mathematical model adopted in this paper includes the center frequency and chirp rate, Eq. (10) is redefined as

$$
\begin{aligned}
& X'_M (k, l) \\
& = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} x'(n) \, W_N^{kn + \frac{l}{N}\left(n - \frac{N-1}{2}\right)^2}, \quad 0 \leq k, l \leq N-1.
\end{aligned}
\tag{11}
$$

2.4 Generic PMC methodology

PMC is an adaptive learning methodology that concurrently implements multiple IFs and adaptively selects and adjusts IFs in the iteration process according to each performance. As a result, samples have a greater probability to occur around the true value. In the initial stage of the PMC method, the choice of IFs in the full-exploration parameter space contributes to the rapid identification of parameter region of interest. In the follow-up stage, the choice of IFs with better local characteristics can contribute to a finer search of the region adjacent to the true parameters.

The general procedure of the PMC is as follows [17,18, 26].

We assume that $p(\boldsymbol{\theta})$ is the objective function, whereas $q^{(t)}(\boldsymbol{\theta})$ is the IF in the tth iteration. When samples cannot be directly drawn from $p(\boldsymbol{\theta})$, the alternative method is to draw sample $\boldsymbol{\theta}^{(i,t)}$ from $q^{(t)}(\boldsymbol{\theta})$. $\boldsymbol{\theta}^{(i,t)}$ denotes the ith sample in the tth iteration. The importance weight of the sample is $\tilde{\omega}^{(i,t)} = p\left(\boldsymbol{\theta}^{(i,t)}\right) / q\left(\boldsymbol{\theta}^{(i,t)}\right)$. Therefore, the random observation $\chi = \left\{\boldsymbol{\theta}^{(i,t)}, \omega^{(i,t)}\right\}_{i=1}^{R}$ derived from the samples and importance weights can approximate the target function, where $\omega^{(i,t)}$ represents the normalized weight corresponding to sample $\boldsymbol{\theta}^{(i,t)}$ for $i = 1, \ldots, R$, and R denotes the total number of samples.

In the initial stage of the method, D IFs $g(\boldsymbol{\theta}; \boldsymbol{\xi}_1), \ldots, g(\boldsymbol{\theta}; \boldsymbol{\xi}_D)$ are selected. The overall IF constructed in the tth iteration can be expressed as follows:

$$
q^{(t)}(\boldsymbol{\theta}) = \sum_{d=1}^{D} \alpha_d^{(t)} g\left(\boldsymbol{\theta}; \boldsymbol{\xi}_d^{(t)}\right)
\tag{12}
$$

where $\sum_{d=1}^{D} \alpha_d^{(t)} = 1$. $d = 1, \ldots, D$; $q^{(t)}(\boldsymbol{\theta})$ is a mixture of D IFs; $g\left(\boldsymbol{\theta}; \boldsymbol{\xi}_d^{(t)}\right)$ and $\alpha_d^{(t)}$, respectively, represent the dth IF and its candidate probability in the tth iteration, that is, $\alpha_d^{(t)}$ could be interpreted as frequency of using the dth IF for sampling; $\boldsymbol{\xi}_d^{(t)}$ represents the construction parameter of function $g\left(\boldsymbol{\theta}; \boldsymbol{\xi}_d^{(t)}\right)$; and $d^{(i,t)} \in \{1, \ldots, D\}$ denotes the index of the IF used for sampling $\boldsymbol{\theta}^{(i,t)}$ and obeys multinomial distribution $M\left(\alpha_1^{(t)}, \ldots, \alpha_D^{(t)}\right)$.

To evaluate the performance of the dth IF $g\left(\boldsymbol{\theta}; \boldsymbol{\xi}_d^{(t)}\right)$, the sum of weights corresponding to all the samples generated from the IF is utilized. $q^{(t)}(\boldsymbol{\theta})$ is determined through all IFs and candidate probability. Therefore, the sum of the weights of samples generated in the true parameter region is large. The formula is as follows:

$$
\alpha_d^{(t)} = \frac{\sum_{i=1}^{R} \omega^{(i,t-1)} \Pi_{g\left(\boldsymbol{\theta}; \boldsymbol{\xi}_d^{(t)}\right)} \boldsymbol{\theta}^{(i,t-1)}}{\sum_{i=1}^{R} \omega^{(i,t-1)}}
\tag{13}
$$

where $\Pi_{g\left(\boldsymbol{\theta}; \boldsymbol{\xi}_d^{(t)}\right)} \boldsymbol{\theta}^{(i,t-1)}$ is an indicator function taking the value of 1 when sample $\boldsymbol{\theta}^{(i,t-1)}$ is generated from IF $g\left(\boldsymbol{\theta}; \boldsymbol{\xi}_d^{(t)}\right)$ and is 0 otherwise.

R samples are drawn from the overall IF in each iteration. Such samples and weights are used to approximate the distribution of random variables. When the times of iterations are sufficiently large, the proposed method can accurately estimate the target distribution.

3 Parameter estimation method based on the MDCFT and the PMC

3.1 Idea of the proposed method

Researchers [16,27] have applied the PMC method to the parameter estimation of multi-sinusoids. Thus, this paper further solves the problem of joint model selection and parameter estimation of multi-component chirp signals. To enable the samples generated from the IFs to occur around the true value with a greater possibility, IFs are elaborately chosen through the MDCFT and are adaptively updated through the PMC.

According to the characteristics of multi-component chirp signals, this paper adopts the MDCFT for peak detection in the DCFT domain. Compared with the conventional DCFT, the MDCFT has the following advantages [24,25]: (1) reduces the "picket fence" effect, (2) has no limit on the signal length, and (3) reduces the number of mirror peaks [14]

or local peaks and decreases the difficulty in peak detection. Furthermore, the MDCFT utilizes the detected peaks to construct IFs and then employs PMC to realize joint model selection and parameter estimation. In the PMC iteration process, on one hand, IFs are constructed according to the samples generated in the former iteration; on the other hand, samples and weights are used to update IFs that will be used in the next iteration. The whole iteration process can be stopped at any time. This method can estimate not only the parameters but also the posterior distribution of the observed data on unknown parameters.

3.2 Proposed method

3.2.1 Construction of IFs

According to Eq. (11), MDCFT is used to transform data into the chirpogram for peak detection. The detected peaks are ranked from large to small based on their amplitudes. The top L spectral peaks with larger amplitudes are retained, and the parameter space of l th peak is assumed to have a bivariate normal distribution, noted as $\varphi(\boldsymbol{\theta}_l)$, where $\boldsymbol{\theta}_l \triangleq (f_l, s_l)$. f_l and s_l, respectively, represent the central frequency and chirp rate of the lth peak. The constructed bivariate normal distribution is expressed as

$$\varphi(\boldsymbol{\theta}_l) \sim N(\boldsymbol{\mu}_l, \boldsymbol{\Sigma}_l), \tag{14}$$

where $\boldsymbol{\mu}_l$ and $\boldsymbol{\Sigma}_l$, respectively, denote the corresponding mean and covariance of parameter $\boldsymbol{\theta}_l$. $\boldsymbol{\Sigma}_l = \begin{pmatrix} \sigma_{f_l}^2 & \rho_l \sigma_{f_l} \sigma_{s_l} \\ \rho_l \sigma_{f_l} \sigma_{s_l} & \sigma_{s_l}^2 \end{pmatrix}$; and $\sigma_{f_l}^2$, $\sigma_{s_l}^2$, and ρ_l, respectively, denote the variance of f_l, the variance of s_l in the Gaussian distribution, and the correlation coefficient between two variances.

After obtaining the result of the detected peaks, greater quantization error may exist if the corresponding parameter of the peaks with the largest amplitude is directly taken as the chirp signal parameter. In addition, the accuracy of parameter estimation is limited and the model order is unlikely to be estimated. Thus, after obtaining the peaks through the MDCFT, the PMC must be utilized to estimate the model order and signal parameters.

When initializing the method, we assume the selected D IFs $g\left(\boldsymbol{\theta}; \boldsymbol{\xi}_1^{(0)}\right), \ldots, g\left(\boldsymbol{\theta}; \boldsymbol{\xi}_D^{(0)}\right)$, where $\boldsymbol{\theta} \triangleq \{f, s\}$. Each IF is composed of the corresponding bivariate normal distribution of $n (n \leq L)$ different peaks in the DCFT domain, and the components of each IF are different. The dth IF can be expressed as the product of n bivariate Gaussian distribution as follows:

$$g\left(\boldsymbol{\theta}; \boldsymbol{\xi}_d^{(0)}\right) = p\left(\varphi\left(\boldsymbol{\theta}_{c_1}^{(0)}\right)\right) \cdot \cdots \cdot p\left(\varphi\left(\boldsymbol{\theta}_{c_n}^{(0)}\right)\right), \tag{15}$$

where $\{c_1, \ldots, c_n\}$ denotes the index number of the peak that constitutes IFs $g\left(\boldsymbol{\theta}; \boldsymbol{\xi}_d^{(0)}\right)$, $1 \leq c_1 < \cdots < c_n \leq L$, and $n \geq 1$. $\boldsymbol{\xi}_d^{(0)}$ represents the mean and covariance of all bivariate Gaussian distributions of IF $g\left(\boldsymbol{\theta}; \boldsymbol{\xi}_d^{(0)}\right)$, that is, $\boldsymbol{\xi}_d^{(0)} = \left(\boldsymbol{\mu}_{c_1}^{(0)}, \boldsymbol{\Sigma}_{c_1}^{(0)}, \ldots, \boldsymbol{\mu}_{c_n}^{(0)}, \boldsymbol{\Sigma}_{c_n}^{(0)}\right)$. The overall IF constructed according to Eq. (12) can be expressed as $q^{(0)}(\boldsymbol{\theta}) = \sum_{d=1}^{D} \alpha_d^{(0)} g\left(\boldsymbol{\theta}; \boldsymbol{\xi}_d^{(0)}\right)$ and $\alpha_d^{(0)} = 1/D$ for $d = 1, 2, \ldots, D$.

3.2.2 Update of IFs

According to the general flow of the PMC method in Sect. 2.4, $g\left(\boldsymbol{\theta}; \boldsymbol{\xi}_d^{(t)}\right)$ represents the dth IF in the tth iteration, which can be expressed as

$$g\left(\boldsymbol{\theta}; \boldsymbol{\xi}_d^{(t)}\right) = p\left(\varphi\left(\boldsymbol{\theta}_{c_1}^{(t)}\right)\right) \cdot \cdots \cdot p\left(\varphi\left(\boldsymbol{\theta}_{c_n}^{(t)}\right)\right), \tag{16}$$

where $\boldsymbol{\xi}_d^{(t)} = \left(\boldsymbol{\mu}_{c_1}^{(t)}, \boldsymbol{\Sigma}_{c_1}^{(t)}, \ldots, \boldsymbol{\mu}_{c_n}^{(t)}, \boldsymbol{\Sigma}_{c_n}^{(t)}\right)$. According to Eq. (12), the overall IF can be constructed as

$$q^{(t)}(\boldsymbol{\theta}) = \sum_{d=1}^{D} \alpha_d^{(t)} g\left(\boldsymbol{\theta}; \boldsymbol{\xi}_d^{(t)}\right). \tag{17}$$

The ith weight of the sample drawn from the IFs in the tth iteration are calculated as follows:

$$\tilde{\omega}^{(i,t)} = p\left(\boldsymbol{\theta}^{(i,t)}\right) / q\left(\boldsymbol{\theta}^{(i,t)}\right). \tag{18}$$

The IF parameters are updated as follows (Please refer to the literature [26,28,29] and see the detailed derivation in "Appendix 2"):

$$\alpha_d^{(t+1)} = \sum_{i=1}^{R} \omega^{(i,t)} h_d\left(\boldsymbol{\theta}^{(i,t)}; \alpha^{(t)}, \boldsymbol{\xi}^{(t)}\right) \tag{19}$$

$$\boldsymbol{\mu}_d^{(t+1)} = \frac{\sum_{i=1}^{R} \omega^{(i,t)} \boldsymbol{\theta}^{(i,t)} h_d\left(\boldsymbol{\theta}^{(i,t)}; \alpha^{(t)}, \boldsymbol{\xi}^{(t)}\right)}{\alpha_d^{(t+1)}} \tag{20}$$

$$\boldsymbol{\Sigma}_d^{(t+1)}$$
$$= \frac{\sum_{i=1}^{R} \omega^{(i,t)} \left(\boldsymbol{\theta}^{(i,t)} - \boldsymbol{\mu}_d^{(t+1)}\right) \left(\boldsymbol{\theta}^{(i,t)} - \boldsymbol{\mu}_d^{(t+1)}\right)^{\mathrm{T}} h_d\left(\boldsymbol{\theta}^{(i,t)}; \alpha^{(t)}, \boldsymbol{\xi}^{(t)}\right)}{\alpha_d^{(t+1)}} \tag{21}$$

where $h_d\left(\boldsymbol{\theta}^{(i,t)}; \alpha^{(t)}, \boldsymbol{\xi}^{(t)}\right) = \frac{\alpha_d^{(t)} g\left(\boldsymbol{\theta}^{(i,t)}; \boldsymbol{\xi}_d^{(t)}\right)}{\sum_{d=1}^{D} \alpha_d^{(t)} g\left(\boldsymbol{\theta}^{(i,t)}; \boldsymbol{\xi}_d^{(t)}\right)}$, and $\omega^{(i,t)}$ denotes the normalized weight of the ith sample in the tth iteration.

As a typical AIS methodology, the PMC method has shown superior performance in solving high-dimensional problems, which has been proven theoretically and experimentally. In the PMC method, samples are drawn according to the importance weight in each iteration, such that the convergence problem is resolved [17,18].

3.2.3 Joint model selection and parameter estimation

In case the number of chirp components in the data is a priori unknown, the idea in the literature [18] was employed in this paper to realize the joint model selection and parameter estimation. We assume that $\boldsymbol{\theta} \in S_K$, S_K denotes the parameter space of M_K. The total parameter space $\boldsymbol{\Omega}$ is assumed as $\boldsymbol{\Omega} \triangleq \cup_{K=0}^{K_{\max}} \{K\} \times S_K$, where $K_{\max}$ denotes the maximum number of the chirp components. We adopted the Bayesian inference strategy. The marginal maximum a posterior estimate of K is given by

$$
\begin{aligned}
K^* &= \arg \max_K \left\{ p\left(K|\boldsymbol{y}\right) \right\} \\
&= \arg \max_K \left\{ \int_{\boldsymbol{\theta} \in S_K} p\left(K, \boldsymbol{\theta}|\boldsymbol{y}\right) \mathrm{d}\boldsymbol{\theta} \right\} \\
&= \arg \max_K \left\{ \int_{(\tilde{K}, \boldsymbol{\theta}) \in \boldsymbol{\Omega}} \prod_K \left(\tilde{K}\right) p\left(\tilde{K}, \boldsymbol{\theta}|\boldsymbol{y}\right) \mathrm{d}\boldsymbol{\theta} \right\}
\end{aligned} \tag{22}
$$

where the indicator function $\prod_K \left(\tilde{K}\right)$ takes the value of 1 when $\tilde{K} = K$ and is 0 otherwise.

R denotes the sample size. The samples $\left\{ \left(K^{(i)}, \boldsymbol{\theta}^{(i)}\right) \right\}_{i=1}^{R}$ can be drawn from the IF $q\left(\boldsymbol{\theta}\right)$, and the normalized weights $\omega^{(i)}$ can be calculated. According to the principle of the IS approach, a Monte Carlo integration can be performed by drawing R samples from the IF to approximate the integral in Eq. (22) with the sample average [18]

$$
\begin{aligned}
K^* &= \arg \max_K \left\{ \int_{(\tilde{K}, \boldsymbol{\theta}) \in \boldsymbol{\Omega}} \prod_K \left(\tilde{K}\right) p\left(\tilde{K}, \boldsymbol{\theta}|\boldsymbol{y}\right) \mathrm{d}\boldsymbol{\theta} \right\} \\
&= \arg \max_K \left\{ \int_{(\tilde{K}, \boldsymbol{\theta}) \in \boldsymbol{\Omega}} \prod_K \left(\tilde{K}\right) \frac{p\left(\tilde{K}, \boldsymbol{\theta}|\boldsymbol{y}\right)}{q\left(\boldsymbol{\theta}\right)} q\left(\boldsymbol{\theta}\right) \mathrm{d}\boldsymbol{\theta} \right\} \\
&\simeq \arg \max_K \left\{ \frac{1}{R} \sum_{i=1}^{R} \prod_K \left(K^{(i)}\right) \cdot \frac{p\left(K^{(i)}, \boldsymbol{\theta}^{(i)}|\boldsymbol{y}\right)}{q\left(\boldsymbol{\theta}^{(i)}\right)} \right\} \\
&= \arg \max_K \left\{ \frac{1}{R} \sum_{i=1}^{R} \prod_K \left(K^{(i)}\right) \cdot \omega^{(i)} \right\}
\end{aligned} \tag{23}
$$

As R goes to infinity, the result in Eq. (23) approximates the true posterior by the law of large numbers [29]. In the parameter estimation, the sum of sample weights approaching the true values usually has a large value λ. When the value λ is less than a specific threshold value (for example,

1e−2), the IF belongs to an unreasonable prior and can be abandoned. Therefore, after several iterations, the IF with the largest λ can be determined and its model order is the true value $\hat{K} = K$.

The tth iteration in the proposed method comprises two steps. Multinomial distribution $M\left(\alpha_1^{(t)}, \ldots, \alpha_D^{(t)}\right)$ is used to denote the candidate probability of IFs, which is usually determined by the sum of all sample weights generated from some IF. One step is to select the IF according to $M\left(\alpha_1^{(t)}, \ldots, \alpha_D^{(t)}\right)$, whereas the other is to draw samples from the selected IF.

If the model order is a priori known, the detection process above can be skipped and the signal parameter can be directly estimated using Eqs. (16)–(21). In the aforementioned model, order estimation method, the pre-detection estimation method first detects peaks according to the chirpogram transformed using the MDCFT method and then constructs IFs. The PMC method is used to estimate the model order and signal parameters at the same time. The method performance is closely related to IF selection. By utilizing the characteristics of chirp signals, the proposed method efficiently selects IFs and the generated samples occur around the true value with a great possibility. The method adapts low SNR and has high estimation accuracy in case of a small sample size. Therefore, the proposed method has efficient detection and parameter estimation performance.
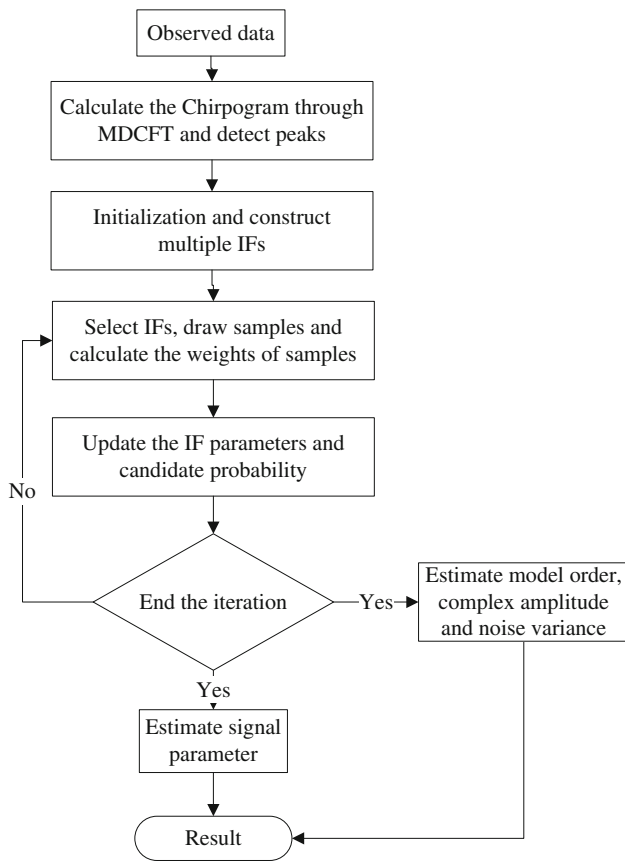
3.3 Flow of the proposed method

Considering the detailed description of the principles and implementation procedures of the proposed method, the flow chart of the MDCFT–PMC method is shown in Fig. 1.

The specific implementation process of the proposed method is shown in Table 1. The estimation process of the model order will be given as supplementary explanation.

4 Numerical results

To validate the parameter estimation performance of multi-component chirp signals based on the MDCFT–PMC method, this section performs simulations in case of changing SNR and sample size and compares the results with that of the MCMC [12–14] and Gibbs–MPMC methods [17,20]. This section also uses simulations to indicate the convergence of the proposed method.

The parameters of the simulations are set as follows: In the Bayesian model of multi-component chirp signals, the hyper-parameter of the prior noise variance is $\upsilon_0 = 2$, $\gamma_0 > 0$. The iteration times of the Gibbs–MPMC and MDCFT–PMC methods are both 10. The size of the samples drawn in each iteration is R. In the initialization of the method, the

Fig. 1 Flow chart of the MDCFT–PMC method

$$\mathrm{MSE} = \sum_{w=1}^{W} \sum_{k=1}^{K} \left(\hat{\boldsymbol{\theta}}_k^{(w)} - \boldsymbol{\theta}_k^{(w)} \right)^2 \Big/ W K, \qquad (24)$$

where W denotes the simulation times in a certain simulation realization. $\boldsymbol{\theta}_k^{(w)}$ and $\hat{\boldsymbol{\theta}}_k^{(w)}$, respectively, correspond to the true and estimated values of the chirp parameter of the kth component in the wth simulation. Accordingly, the CRLB [30], which is used to compare performance, also takes the mean square error of the CRLB of each chirp component.

4.1 Example

With the aid of simulation in typical realization, the implementation process of the proposed method is illustrated. This subsection provides a comparison of the chirpograms based on the DCFT and MDCFT methods.

Similar to the simulation realization in [16,20], two chirp components with SNR = 10 dB are assumed in the environment. The normalized center frequencies are 0.302 and 0.328; the normalized chirp rates are 0.001 and 0.002; the complex amplitudes are 1; the number of observed data is 50; the calculation formula of the SNR is SNR = $10 \log 10 \left[A_1^2 / \sigma_\varepsilon^2 \right]$; the observation noise σ_ε^2 obeys complex Gaussian distribution; and the components in the real and imaginary parts are independent and satisfy Gaussian distribution $N \left(0, \sigma_\varepsilon^2 / 2 \right)$. In single simulation experiment, the divided lattice space based on the MDCFT is 100×100. The iteration times based on the PMC method are 10, and the size of the samples drawn in each iteration is 1,000. The estimation outputs of the center frequency and chirp rate of two chirp components are shown in Fig. 2.

In Fig. 2, the MDCFT–PMC method can guarantee that the generated samples will occur around the true value with great probability and can reach very high estimation accuracy after 10 iterations. After the relatively acute fluctuation of the center frequency and chirp rate, the estimation value can stabilize around the true value and can reach very high estimation accuracy after 10 iterations. The estimated value

variances of the center frequency and chirp rate of IFs are 0.01 and 0.001, respectively. The correlation coefficient between both variances is 0.1. The setting of the MCMC method is consistent with the literature [12]. The selected length of the Markov chain is $10 \times R$ when comparing the performance. In each simulation realization, W simulations are implemented. When evaluating the parameter estimation accuracy of each method, the variances of the center frequency and chirp rate of each chirp are taken as the criterion, which is defined as

Table 1 Implementation procedures of the MDCFT–PMC method

Parameter estimation based on the MDCFT–PMC method
Step 1. Initialization. In view of the observed data, utilize Eq. (11) to calculate chirpogram and perform peak detection, retain top L spectral peaks with the largest amplitudes. According to Eq. (15), construct D IFs and note the candidate probability α_d as $1/D$ for $d = 1, 2, \ldots, D$
In tth iteration:
Step 2. Draw the index numbers of IFs from multinomial distribution $M \left(\alpha_1^{(t)}, \ldots, \alpha_D^{(t)} \right)$
Step 3. According to the IFs appointed by the index numbers, draw samples and calculate weights according to Eq. (18)
Step 4. After all samples are completely generated, calculate the normalized weights
Step 5. According to Eq. (19) to (21), update the parameters and candidate probability of IFs
Repeat Steps 2–5 until the end of iteration, obtain the signal parameters and model order of multi-component chirp, and estimate the complex amplitudes and noise variance at the same time

Note If the model order of chirp signals is known, the estimation of model order in the implementation procedure can be omitted

Fig. 2 Iteration estimation results of the center frequency and chirp rate (*up*), and the samples histogram after the 10th iteration (*down*)
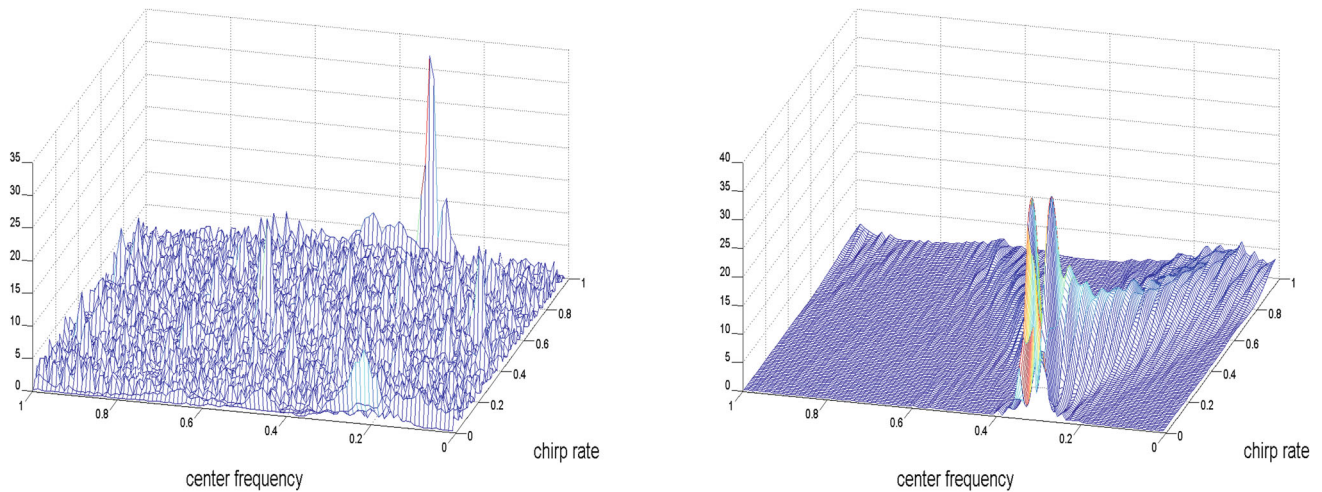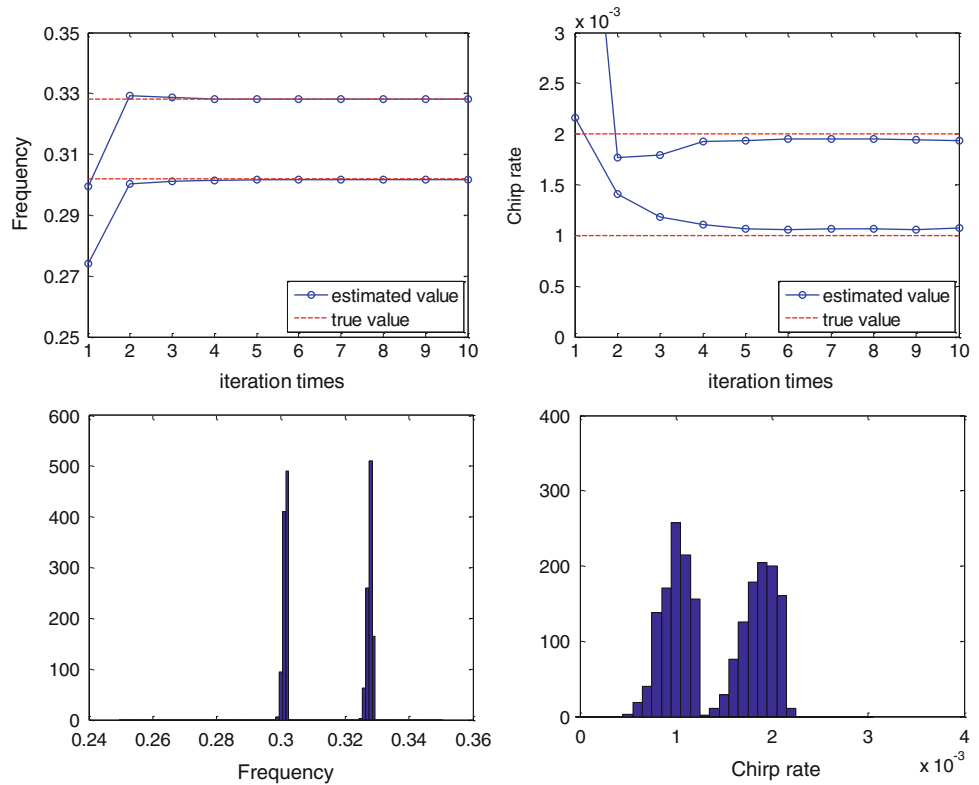
Fig. 3 Chirpogram calculated by the DCFT (*left*) and the MDCFT (*right*)

can accurately reflect the true value. The iteration process of the MDCFT–PMC method can be interrupted at any time, and an increase in iteration times can improve the parameter estimation accuracy.

Figure 3 shows the chirpogram of the signals calculated by the DCFT and MDCFT. The divided lattice space is 100×100. Thus, the center frequency and chirp rate both do not lie in the divided lattice point. Figure 3 (left) shows that the chirpogram calculated by the DCFT is affected by the "picket

fence," and the true peak is submerged in numerous peaks in the DCFT domain. Consequently, the detection accuracy of peaks decreased. Figure 3 (right) shows that the MDCFT can avoid the "picket fence" effect and easily detect the correct peaks.

Therefore, when constructing the IFs of the PMC method with the information provided by the peaks in the chirpogram, the detection results can include the true peak with more retained peaks in the chirpogram calculated by the DCFT,

which undoubtedly increases the processing difficulty and computational complexity of the follow-up method. By contrast, the chirpogram calculated by the MDCFT evidently increases the detection rate of the real peaks and better constructs more reasonable IFs using the detection results.

4.2 Estimation performance of the MDCFT–PMC method on the model order

The estimation performance of the MDCFT–PMC method in terms of the model order of multi-component chirp signals is verified through simulation. The variation range of the SNR falls between $[-5, 10]$ dB, and the variation range of the sample size falls between $[600, 1,000]$, with a step length of 200. The other simulation parameters are similar to those in Sect. 4.1. The simulation experiments are repeated 50 times.
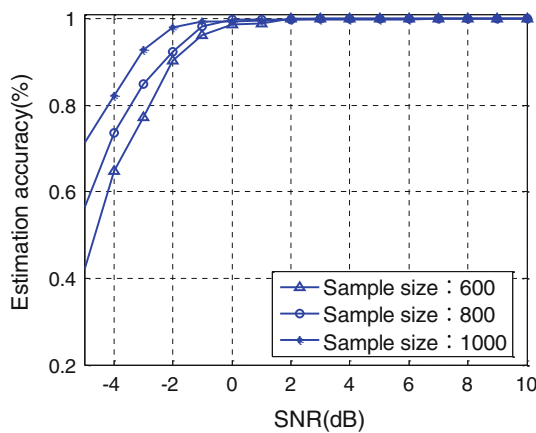
**Fig. 4** Estimation accuracy of the MDCFT–PMC method on the model order in case of different SNR

The estimation accuracy of the MDCFT–PMC method on the model order is shown in Fig. 4.

Figure 4 shows that in terms of different sample sizes, the MDCFT–PMC method does not have approximately the same SNR threshold. When the sample size is 1,000, the estimation accuracy of the MDCFT–PMC method can reach 95 % or more in case of -2 dB. When the sample size is 600 or 800, the estimation accuracy of the MDCFT–PMC method can reach 95 % in case of -1 dB. An increase in sample size can improve the estimation accuracy of the MDCFT–PMC method in terms of the model order and enable the method to have the adaptive capacity for low SNR and to adapt to the follow-up processing requirements.

4.3 Convergence of the MDCFT–PMC method

The MDCFT–PMC method adopts the strategies based on the PMC and has good convergence. By the aid of the simulation experiments in typical simulation realizations, the method is illustrated in detail. The simulation in Sect. 4.1 was repeated for 50 times, and the update speed of the center frequency and chirp rate in 10 iterations was obtained as $\sum_{w=1}^{W} \left\| \hat{\boldsymbol{\theta}}_{k+1} - \hat{\boldsymbol{\theta}}_k \right\|_2 / \left\| \hat{\boldsymbol{\theta}}_k \right\|_2 / W$. Figure 5a corresponds to the center frequency parameters, whereas Fig. 5b corresponds to the chirp rate parameters.

Figure 5a, b shows that the convergence of the estimation of the MDCFT–PMC method on the center frequency and chirp rate is good. The convergence rate is initially rapid. After four iteration processes, the sample results of the center frequency and chirp rate quickly converged. In the follow-up iteration processes, the convergence rate tends to be stable. During method initialization, the constructed IFs tend to select IFs with a larger search parameter space. Significant randomness exists during sample generation. In the follow-
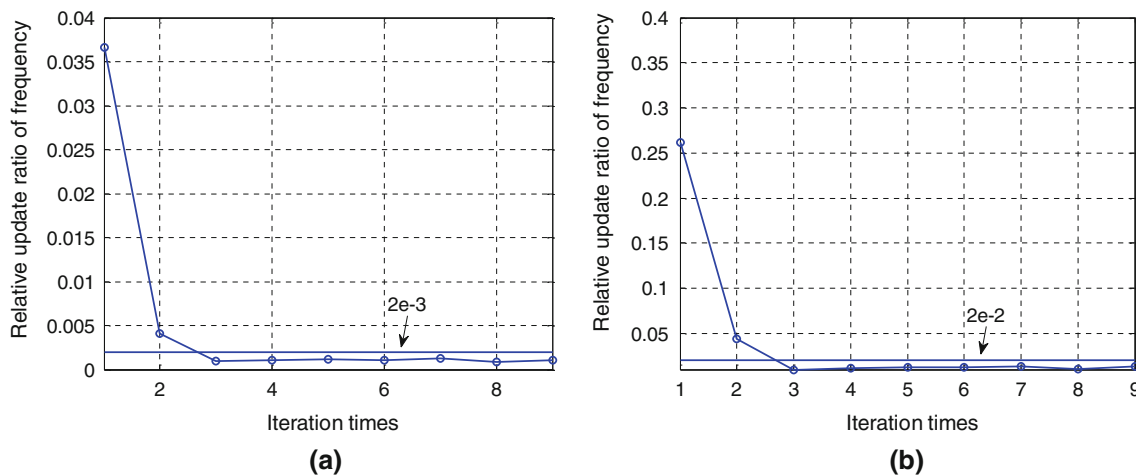
Fig. 5 Convergence of the MDCFT–PMC method. **a** Center frequency; **b** chirp rate

up stage, the MDCFT–PMC method adaptively selects and adjusts IFs in the iteration process. In this manner, the IFs with better local character are likely to be selected and the sample can occur around the true value with greater possibility.
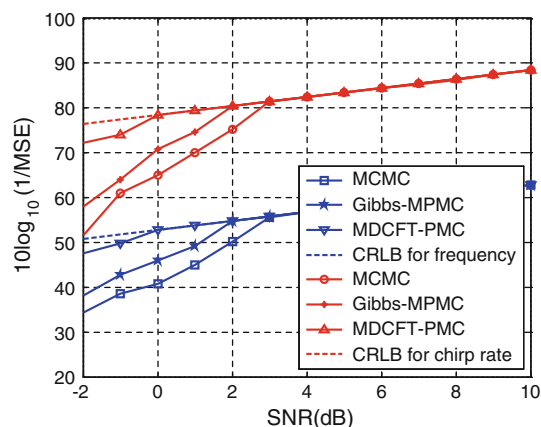
4.4 Parameter estimation performance of the MDCFT–PMC method

The parameter estimation performance of the MDCFT–PMC method on multi-component chirp signals is verified. To prevent the influences of other factors, the model order of chirp signals is assumed as known. This section investigates the MDCFT–PMC method performance along with the SNR and sample size variation and compares this method with the classical MCMC and Gibbs–MPMC methods.

4.4.1 Performance comparison in noisy environment

To verify the robustness of the MDCFT–PMC method in a noisy environment, this subsection evaluates the parameter estimation performance of the proposed method through simulation experiments and compares with the MCMC and the Gibbs–MPMC method. We assume that the SNR variation range of chirp signals falls between [−2, 10] dB. The simulation experiments are repeated 100 times, whereas the other parameters are the same as those in Sect. 4.1. The parameter estimation performance chart of the center frequency and chirp rate based on the MCMC, Gibbs–MPMC, and MDCFT–PMC methods is shown in Fig. 6.

Figure 6 shows that the parameter estimation accuracy of the three methods to reach the CRLB differs in terms of the SNR thresholds. The MCMC and Gibbs–MPMC methods adopt the DCFT method to calculate the chirpogram and

perform spectral peak detection. Compared with the MDCFT method, the chirpogram calculated using the DCFT method does not have evident peaks, which results in a decrease in detection performance. The MCMC method performance depends on the proposal distribution of the initial selection. If the selected proposal distribution is not ideal, the performance will exhibit a sharp reduction. The Gibbs–MPMC method selects spectral peaks according to the chirpogram calculated using the DCFT method and constructs the IF to generate samples. Resampling is adopted to approach the true parameters constantly. The performance of this method can be easily affected by the SNR. The estimation accuracy of the center frequency and chirp rate of the MDCFT–PMC method can reach the CRLB in case of 0 dB, whereas those of the Gibbs–MPMC method can reach the CRLB in case of 2 dB. The effect of the SNR is greater on the MPMC method. In the case of low SNR, the proposed method has better parameter estimation performance.

4.4.2 Performance comparison in case of changing sample size

A simulation in case of changing sample size is performed. Only the Gibbs–MPMC and MDCFT–PMC methods are considered. The variation range of the sample size falls between [100, 1,400], with the step length of 100. The simulation is repeated 100 times. The chart of the parameter estimation performance of the center frequency and chirp rate based on the Gibbs–MPMC and MDCFT–PMC methods is shown in Fig. 7.

Figure 7 shows that when the sample size is 400, the parameter estimation accuracy of the MDCFT–PMC method can reach the CRLB. When the sample size is 700, the parameter estimation accuracy of the Gibbs–MPMC method can reach the CRLB. The high quality of the samples generated in the parameters' importance region of the Gibbs–MPMC method

Fig. 6 Parameter estimation performance of the MCMC, Gibbs–MPMC, and MDCFT–PMC methods along with the variation of SNR
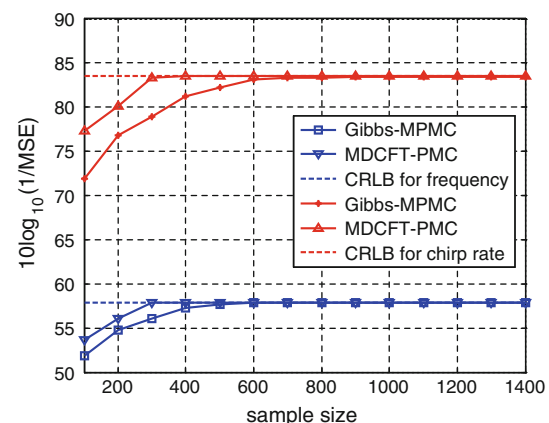
Fig. 7 Parameter estimation performance of Gibbs–MPMC and MDCFT–PMC methods with varied sample size

is appropriate to the true parameters. In other words, a larger size of samples results in a greater possibility of involving the true value. Therefore, the performance of the Gibbs–MPMC method is significantly affected by the sample size. In the iteration process of the MDCFT–PMC method, the samples with high quality are utilized to update the IFs, which are used to generate samples in the next iteration. In this way, even in the case of a small sample size, the parameter and target function of multi-component chirp signals can be estimated. In conclusion, the proposed method exhibits better parameter estimation performance in case of a small sample size.

5 Conclusion

Through the combination and development of the previous achievements on the MDCFT and PMC methodologies, this paper applied the relevant progress into the joint Bayesian model selection and parameter estimation of multi-component chirp signals with white Gaussian noise. To enable the samples drawn from the IF to occur around the true value with a greater possibility, this paper adopted the MDCFT to calculate the chirpogram and detect peaks in the DCFT domain and constructed multiple IFs on the basis of peaks. Meanwhile, this paper used the PMC method to update the IFs with an adaptive learning process. This method can estimate not only the signal parameters but also the posterior distribution of the observed data on unknown parameters.

The simulation results show that compared with two existing methods based on the Monte Carlo concept, the proposed method exhibits better performance in case of lower SNR or small sample size and good convergence. The proposed method also has strong significance in solving related problems in spectral analysis, wideband radar imaging field, and so on.

6 Appendix 1

The posterior distribution of $\left(\boldsymbol{\alpha}, \boldsymbol{f}, \boldsymbol{s}, \sigma_\varepsilon^2, K \right)$ can be expressed as follows according to the likelihood function of $\boldsymbol{y}$ given in Eq. (3) and the a priori of each parameter given in Eqs. (6) and (7):

$$p\left(\boldsymbol{\alpha}, \boldsymbol{f}, \boldsymbol{s}, \sigma_\varepsilon^2, K | \boldsymbol{y} \right) \propto p\left(\boldsymbol{y} | \boldsymbol{\alpha}, \boldsymbol{f}, \boldsymbol{s}, \sigma_\varepsilon^2, K \right) p\left(\boldsymbol{\alpha} | \boldsymbol{f}, \boldsymbol{s}, \sigma_\varepsilon^2, K \right)$$
$$\times p\left(\boldsymbol{f}, \boldsymbol{s} | K \right) p\left(\sigma_\varepsilon^2 \right) p\left(K \right)$$
$$\propto \frac{1}{\left(\pi \sigma_\varepsilon^2 \right)^N} \exp\left\{ -\frac{1}{\sigma_\varepsilon^2} \left(\boldsymbol{y} - \boldsymbol{D}\boldsymbol{\alpha} \right)^{\mathrm{H}} \left(\boldsymbol{y} - \boldsymbol{D}\boldsymbol{\alpha} \right) \right\}$$

$$\times \frac{\left| \boldsymbol{D}^{\mathrm{H}}\boldsymbol{D} \right|}{\left(\pi \delta^2 \sigma_\varepsilon^2 \right)} \exp\left\{ -\frac{1}{\delta^2 \sigma_\varepsilon^2} \boldsymbol{\alpha}^{\mathrm{H}} \left(\boldsymbol{D}^{\mathrm{H}} D \right) \boldsymbol{\alpha} \right\}$$
$$\times \left(\frac{1}{2} \right)^K \times \sigma_\varepsilon^{2(-\upsilon_0 - 1)} \exp\left[\frac{-\gamma_0}{\sigma_\varepsilon^2} \right] \times \frac{1}{K} \tag{25}$$

Referring to the literature [12,13,18], the expression after integrating terms relevant to amplitudes $\boldsymbol{\alpha}$ in Eq. (25) is given as

$$p\left(\boldsymbol{\alpha}, \boldsymbol{f}, \boldsymbol{s}, \sigma_\varepsilon^2, K | \boldsymbol{y} \right)$$
$$\propto \frac{\left| \boldsymbol{D}^{\mathrm{H}}\boldsymbol{D} \right|}{\left(\pi \delta^2 \sigma_\varepsilon^2 \right)} \exp\left\{ -\frac{1}{\sigma_\varepsilon^2} \left(\boldsymbol{\alpha} - \boldsymbol{m}_{\boldsymbol{\alpha}} \right)^{\mathrm{H}} \boldsymbol{\Sigma}_{\boldsymbol{\alpha}}^{-1} \left(\boldsymbol{\alpha} - \boldsymbol{m}_{\boldsymbol{\alpha}} \right) \right\}$$
$$\times \frac{1}{\left(\pi \sigma_\varepsilon^2 \right)^N} \exp\left\{ -\frac{1}{\sigma_\varepsilon^2} \boldsymbol{y}^{\mathrm{H}}\boldsymbol{P}\boldsymbol{y} \right\}$$
$$\times \sigma_\varepsilon^{2(-\upsilon_0 - 1)} \exp\left[\frac{-\gamma_0}{\sigma_\varepsilon^2} \right] \times \frac{1}{K \cdot 2^K} \tag{26}$$

where

$$\boldsymbol{\Sigma}_{\boldsymbol{\alpha}}^{-1} = \boldsymbol{D}^{\mathrm{H}}\boldsymbol{D} \left(1 + \delta^{-2} \right) \tag{27}$$

$$\boldsymbol{m}_{\boldsymbol{\alpha}} = \boldsymbol{\Sigma}_{\boldsymbol{\alpha}} \boldsymbol{D}^{\mathrm{H}}\boldsymbol{y} \tag{28}$$

$$\boldsymbol{P} = \boldsymbol{I}_N - \boldsymbol{D} \left(\boldsymbol{D}^{\mathrm{H}}\boldsymbol{D} \right)^{-1} \boldsymbol{D}^{\mathrm{H}} / \left(1 + \delta^{-2} \right) \tag{29}$$

Equation (26) has a Gaussian function of $\boldsymbol{\alpha}$ and inverse Gamma function of σ_ε^2. After the integration of these two variables in (26), the posterior distribution of the observed data $\boldsymbol{y}$ on $K, \boldsymbol{f}$, and $\boldsymbol{s}$ is

$$p\left(\boldsymbol{f}, \boldsymbol{s}, K | \boldsymbol{y} \right) \propto \left(\gamma_0 + \boldsymbol{y}^{\mathrm{H}}\boldsymbol{P}\boldsymbol{y} \right)^{-(N + \upsilon_0)} \times \frac{\left(1 + \delta^2 \right)}{K \cdot 2^K} \tag{30}$$

7 Appendix 2

We assume that $p\left(\boldsymbol{\theta} \right)$ is the target function; $q^{(t)}\left(\boldsymbol{\theta} \right)$ represents the IF in the tth iteration; and $\boldsymbol{\theta}$ represents the unknown parameters. When applying the PMC methodology to the parameter estimation problem, whether the adopted IF can be approximated to the target function is important. Our objective is to minimize the Kullback divergence, which can be expressed as $K\left(p || q^{(t)} \right) = \int \log\left(\frac{p(\boldsymbol{\theta})}{q^{(t)}(\boldsymbol{\theta})} \right) p\left(\boldsymbol{\theta} \right) \mathrm{d}\boldsymbol{\theta}$.

Assume $h_d\left(\boldsymbol{\theta}; \alpha, \boldsymbol{\xi} \right) = \frac{\alpha_d g(\boldsymbol{\theta}; \boldsymbol{\xi}_d)}{\sum_{d=1}^{D} \alpha_d g(\boldsymbol{\theta}; \boldsymbol{\xi}_d)}$. In tth iteration, the intermediate quantity is constructed as

$$L^{(t)}\left(\alpha, \boldsymbol{\xi} \right)$$
$$= \int \sum_{d=1}^{D} h_d\left(\boldsymbol{\theta}; \alpha^{(t)}, \boldsymbol{\xi}_d^{(t)} \right) \log\left(\alpha_d^{(t)} g\left(\boldsymbol{\theta}; \boldsymbol{\xi}_d^{(t)} \right) \right) p\left(\boldsymbol{\theta} \right) \mathrm{d}\boldsymbol{\theta} \tag{31}$$

The derivation in [26,28,29] indicates that for any α and $\boldsymbol{\xi}$, when the intermediate quantity (31) increases, the target function also increases. The maximum $L^{(t)}\left(\alpha, \boldsymbol{\xi} \right)$ canobtain

a closed solution. In the multivariate Gaussian distribution, the parameters include mean value and covariance matrix. The intermediate quantity can be expressed as

$$
\begin{aligned}
L^{(t)}&(\alpha,\xi)\\
&=\int\sum_{d=1}^{D}h_d\left(\theta;\alpha^{(t)},\xi_d^{(t)}\right)\left\{\log\left(\alpha_d^{(t)}\right)\right.\\
&\quad\left.-\frac{1}{2}\left[\log\left|\Sigma_d^{(t)}\right|+\left(\theta-\mu_d^{(t)}\right)^{\mathrm{T}}\left(\Sigma_d^{(t)}\right)^{-1}\left(\theta-\mu_d^{(t)}\right)\right]\right\}p\left(\theta\right)d\theta
\end{aligned}
$$
(32)

up to terms that do not depend on α,ξ.

When the above formula reaches the minimum, the following equations are satisfied:

$$
\alpha_d^{(t+1)}=\int h_d\left(\theta;\alpha^{(t)},\xi_d^{(t)}\right)p\left(\theta\right)\mathrm{d}\theta
$$
(33)

$$
\mu_d^{(t+1)}=\frac{\int\theta^{(t)}h_d\left(\theta^{(t)};\alpha^{(t)},\xi_d^{(t)}\right)p\left(\theta\right)\mathrm{d}\theta}{\alpha_d^{t+1}}
$$
(34)

$$
\begin{aligned}
&\Sigma_d^{(t+1)}\\
&=\frac{\int\left(\theta^{(t)}-\mu_d^{(t+1)}\right)\left(\theta^{(t)}-\mu_d^{(t+1)}\right)^{\mathrm{T}}h_d\left(\theta;\alpha^{(t)},\xi_d^{(t)}\right)p\left(\theta\right)\mathrm{d}\theta}{\alpha_d^{t+1}}
\end{aligned}
$$
(35)

In practice, the numerator and the denominator are integral. By utilizing the samples and weights in each iteration for approximation, Eqs. (19)–(21) can be obtained.

References

1. Djuric, P.M., Kay, S.M.: Parameter estimation of chirp signals. IEEE Trans. Acoust. Speech Signal Process. **38**(12), 2118–2126 (1999)
2. Li, H.S., Djuric, P.M.: MMSE estimation of nonlinear parameters of multiple linear & quadratic chirps. IEEE Trans. Signal Process. **46**(3), 796–800 (1998)
3. Abatzoglou, T.J.: Fast maximum likelihood joint estimation of frequency and frequency rate. In: IEEE International Conference on Aerospace Electronic Systems, Tokyo, Japan, April 7–11 (1986)
4. Besson, O., Ghogho, M., Swami, A.: Parameter estimation for random amplitude chirp signals. IEEE Trans. Signal Process. **47**(12), 208–3219 (1999)
5. Ikram, M.Z., Meraim-Abed, K., Hua, Y.: Iterative parameter estimation of multiple chirp signals. Electron. Lett. **33**(8), 657–659 (1997)
6. Ikram, M.Z., Meraim-Abed, K., Hua, Y.: Estimating the parameters of chirp signals: an iterative approach. IEEE Trans. Signal Process. **46**(12), 3436–3441 (1998)
7. Barbarossa, S.: Analysis of multicomponent LFM signals by a combined Wigner–Hough transform. IEEE Trans. Signal Process. **43**(6), 1511–1515 (1995)
8. Guo, J., Zou, H., Yang, X., Liu, G.: Parameter estimation of multicomponent chirp signals via sparse representation. IEEE Trans. Aerosp. Electron. Syst. **47**(3), 2261–2268 (2011)
9. Applebaum, L., Howard, S.D., Searle, S., Calderbank, R.: Chirp sensing codes: deterministic compressed sensing measurements for fast recovery. Appl. Comput. Harmon. Anal. **26**(2), 283–290 (2009)
10. Millioz, F., Davies, M.: Sparse detection in the chirplet transform: application to FMCW radar signals. IEEE Trans. Signal Process. **60**(6), 2800–2813 (2012)
11. Mallat, S.G., Zhang, Z.: Matching pursuits with time-frequency dictionaries. IEEE Trans. Signal Process. **41**(12), 3397–3415 (1993)
12. Andrieu, C., Doucet, A.: Joint Bayesian model selection and estimation of noisy sinusoids via reversible jump MCMC. IEEE Trans. Signal Process. **47**(10), 2667–2676 (1999)
13. Davy, M., Doncarli, C., Tourneret, J.Y.: Classification of chirp signals using hierarchical Bayesian learning and MCMC methods. IEEE Trans. Signal Process. **50**(2), 377–388 (2002)
14. Lin, C.C., Djuric, P.M.: Estimation of chirp signals by MCMC. in: ICASSP '00. Proceedings, June 5–9, pp. 265–268 (2000)
15. Lin, Y., Peng, Y., Wang, X.: Maximum likelihood parameter estimation of multiple chirp signals by a new Markov Chain Monte Carlo approach. In: IEEE Radar Conference Proceedings, Philadelphia, USA, Apr. 26–29, pp. 559–562 (2004)
16. Saha, S., Kay, S.M.: Maximum likelihood parameter estimation of superimposed chirps using Monte Carlo importance sampling. IEEE Trans. Signal Process. **50**(2), 224–230 (2002)
17. Cappe, O., Guillin, A., Marin, J.M., Robert, C.P.: Population Monte Carlo. J. Comput. Graph. Stat. **13**(4), 907–929 (2004)
18. Hong, M., Bugallo, M.F., Djuric, P.: Joint model selection and parameter estimation by population Monte Carlo simulation. IEEE J. Sel. Top. Signal. Process. **4**(3), 526–539 (2010)
19. Bugallo, M.F., Hong, M., Djuric, P.M.: Marginalized population Monte Carlo. In: ICASSP '09. Proceedings, Taipei, Taiwan, April 19–24, pp. 2925–2928 (2009)
20. Shen, B., Bugallo, M.F., Djuric, P.M.: Estimation of multimodal posterior distributions of chirp parameters with population Monte Carlo sampling. In: ICASSP '12. Proceedings, Kyoto, Japan, March 25–30, pp. 3861–3864 (2012)
21. Xia, X.G.: Discrete chirp-Fourier transform and its application to chirp rate estimation. IEEE Trans. Signal Process. **48**(11), 3122–3133 (2000)
22. Fan, P., Xia, X.G.: Two modified discrete chirp-Fourier transform schemes. Sci. China, Ser. F **44**(5), 329–341 (2001)
23. Guo, X., Sun, H.B., Wang, S.L.: Comments on discrete chirp-Fourier transform and its application to chirp rate estimation. IEEE Trans. Signal Process. **50**(12), 3115–3116 (2002)
24. Xia, X.G.: Response to comments on discrete chirp-Fourier transform and its application to chirp rate estimation. IEEE Trans. Signal Process. **50**(12), 3116 (2002)
25. Wu, L., Wei, X.Z., Yang, D.G., Wang, H.Q., Li, X.: ISAR imaging of targets With complex motion based on discrete chirp Fourier transform for cubic chirps. IEEE Trans. Geosci. Remote Sens. **50**(10), 4201–4212 (2012)
26. Darren, W., Martin, K., Karim, B., Olivier, C., Jean, F.C., Gersende, F., Simon, P., Christian, P.R.: Estimation of cosmological parameters using adaptive importance sampling. Phys. Rev. D **80**(2), 1511–1539 (2009)
27. Cheng, H., Zeng, D., Zhu, J., Tang, B.: Maximum likelihood estimation of co-channel multicomponent polynomial phase signals using importance sampling. Progress Electromagn. C **23**, 111–122 (2011)
28. Kilbinger, M., Wraith, D., Robert, C.P., Benabed, K., Cappé, O., Cardoso, J.F., Fort, G., Prunet, S., Bouchet, F.R.: Bayesian model

comparison in cosmology with population Monte Carlo. MNRAS **405**(4), 2381–2390 (2010)

29. Douc, R., Guillin, A., Marin, J.M., Robert, C.P.: Convergence of adaptive mixtures of importance sampling schemes. Ann. Appl. Stat. **35**(1), 420–448 (2007)

30. Friedlander, B., Francos, J.M.: Estimation of amplitude and phase parameter of multicomponent signals. IEEE Trans. Signal Process. **43**(4), 917–926 (1995)