JOINT MODEL SELECTION AND PARAMETER ESTIMATION OF GTD MODEL USING RJ-MCMC ALGORITHM

Shi Zhiguang, Zhou Jianxiong, Zhao Hongzhong, Fu Qiang

ATR Lab, National University of Defence Technology, Changsha, P.R.China. 410073

ABSTRACT

The Bayes principle is applied to the joint model selection and parameter estimation of GTD model to explore the prior information. An algorithm using RJ-MCMC is designed. It not only has better model selection and parameter estimation performance than the non-Bayes algorithms, but also solves the mixed parameter estimation problem in GTD model effectively. The advantage of this algorithm is especially evident at low SNR, for short data and with closely-spaced components. Simulations verify the effectiveness of this algorithm.

Key words: Bayes principle, RJ-MCMC, GTD model, model order selection, parameter estimation.

1. INTRODUCTION

Scattering center model describe the electromagnetic characteristic of radar target in optical region, and is widely used in radar data compression, target analysis and signal processing. The scattering center model based on Geometric Theory of Diffraction (GTD model) is a basic scattering center model [1]. Extracting the parameters of GTD model from radar measurements includes two problems: one is to decide how many scattering centers there are on the target, i.e., model order selection or model selection, the other is to estimate the parameters of each scattering center, i.e., parameter estimation. The commonly used model selection criteria such as AIC, MDL and MAP are all asymptotically effective and their performance degenerates under low SNR and short data [2]. Meanwhile, the parameter estimation problem of GTD model involves optimizing a high dimensional non-linear cost function with mixed (both discrete and continuous) parameters and is very difficult to solve.

In this paper, Bayes method [3, 4] is applied in joint model selection and parameter estimation problem of GTD model. Bayes methods always use high dimensional non-linear integration which has no closed-form analytical solution. The reversible jump Markov Chain Monte Carlo (RJ-MCMC) method [5, 6] will be used to solve this problem.

The Bayes based joint model selection and parameter estimation of GTD model using RJ-MCMC algorithm has many advantages. ①There are prior information and constraints on the parameters of GTD model, which are vital in the Bayes method. ② Bayes method achieves better performance than non-Bayes methods such as ESPRIT and MUSIC by exploring the prior information. ③ RJ-MCMC algorithm avoids simplification when calculating the *a posteriori* probability and therefore is more accurate in model selection than the asymptotically effective criteria. ④ RJ-MCMC algorithm can estimate both the discrete and continuous parameters in GTD model at the same time, so it avoids the error brought by taking the type parameter as a continuous one.

2. JOINT MODEL SELECTION AND PARAMETER ESTIMATION OF GTD MODEL BASED ON BAYES PRINCIPLE

The target response stimulated by a stepped frequency radar can be expressed as follows according to GTD theory [1].

$$y(n) = \sum_{i=1}^{k} a_i \left(j \frac{f_c + n\Delta f}{f_c} \right)^{\alpha_i} \exp(j \frac{4\pi(f_c + n\Delta f)}{c} r_i) + w_n \ , \quad n = 0, 2, ..., N-1 \quad (1)$$

where k is the number of scattering centers, i.e., the order of the model. a_i, r_i and α_i are respectively the scattering complex intensity, projective location and geometric type parameter of the ith scattering center. We have α_i=0.5k with $k \in \{-2,-1,0,1,2\}$ for most dominant scattering centers. f_c is the starting frequency, Δf the frequency step and N the frequency stepping number. w_n is the measured noise in the nth pulse which is assumed to be white Gaussian. Equation (1) can also be expressed in the matrix form.

$$Y = D_k \vec{A}_k + \vec{W}_k \quad (2)$$

where $Y = [y(0), \cdots, y(N-1)]^T$, $\vec{A}_k = [A_1, \cdots, A_k]^T$, $\left[D_k\right]_{n+1,i} = \left(j\left(1 + n\Delta f / f_c\right)\right)^{\alpha_i} \exp\left(jn\omega_i\right)$, $\vec{W}_k = [w_0, \cdots, w_{N-1}]^T$, $\vec{\omega}_k = [\omega_1, \cdots, \omega_k]^T$, $\vec{\alpha}_k = [\alpha_1, \cdots, \alpha_k]^T$, $\vec{W}_k \sim N(0, \sigma_k^2 I_{NN})$, $\omega_i = 4\pi r_i \Delta f / c$, $A_i = a_i \exp(j4\pi r_i f_c / c)$.

Now the parameters to be estimated are converted to $\theta_k \triangleq \left(\vec{A}_k^T, \vec{\alpha}_k^T, \vec{\omega}_k^T, \sigma_k^2\right)^T$. We note the constraints that $\omega_i \in [-\pi, \pi]$ and $\alpha_i \in [-1,-0.5,0,0.5,1]$ in GTD model. This prior information will be explored by the Bayes method in the following to improve the model selection and parameter estimation accuracy.

In the Bayes method, all the unknown parameters are supposed to be random variables. The best choice of k should maximize the *a posteriori* probability and the estimate of θ_k should minimize the mean square error (MMSE estimates).

$$\hat{k} = \arg\max_{k \in Z_K} p\left(k \mid Y\right) \qquad Z_K = \left\{0, 1, \cdots, k_{\max}\right\} \tag{3}$$

$$\hat{\theta}_k = \mathrm{E}\left(\theta_k \mid Y, k\right) = \int \theta_k p\left(\theta_k \mid Y, k\right) d\theta_k \tag{4}$$

We assume σ_k^2 comes from a conjugate inverse-Gamma distribution, i.e., $\sigma_k^2 \sim Ig\left(v_0/2, \gamma_0/2\right)$ [6]. The distribution of $\left(k, \vec{A}_k, \vec{\omega}_k, \vec{\alpha}_k\right)$ can be assumed according to the physical meaning of GTD model.

$$\begin{aligned} & p\left(k, \vec{A}_k, \vec{\omega}_k, \vec{\alpha}_k \mid \sigma_k^2\right) \\ & = p\left(\vec{A}_k \mid \vec{\omega}_k, \vec{\alpha}_k, k, \sigma_k^2\right) p\left(\vec{\omega}_k \mid k, \sigma_k^2\right) p\left(\vec{\alpha}_k \mid k, \sigma_k^2\right) p\left(k \mid \sigma_k^2\right) \\ & \propto \left|\pi\sigma_k^2 \Sigma_k\right|^{-k} \exp\left(-\frac{\vec{A}_k^H \Sigma_k^{-1} \vec{A}_k}{\sigma_k^2}\right) \bullet U_{\Omega_k}\left(\vec{\omega}_k\right) \bullet DU_{\Psi_k}\left(\vec{\alpha}_k\right) \bullet \frac{\Lambda^k}{k!} \exp\left(-\Lambda\right) \end{aligned} \tag{5}$$

where $\Sigma_k^{-1} = \delta^{-2} D_k^H D_k$, δ^2 is the expected SNR The first term is the *a priori* distribution of $\vec{A}_k$ which is zero mean Gaussian with covariance $\sigma_k^2 \Sigma_k$; the second term is the *a priori* distribution of $\vec{\omega}_k$ which is uniformly distributed on $\Omega_k \triangleq \left(-\pi, \pi\right)^k$; the third term is the *a priori* distribution of $\vec{\alpha}_k$ which is also uniformly distributed on $\Psi_k \triangleq \left[-1, -0.5, 0, 0.5, 1\right]^k$; the last term is the *a priori* distribution of k which is a truncated Poisson distribution with expected order Λ. If k=0 we have $\vec{A}_0^H \Sigma_0^{-1} \vec{A}_0 \triangleq 0$ and $\left|\pi\sigma_0^2 \Sigma_0\right| \triangleq 1$.

The *a posteriori* distribution is

$$\begin{aligned} & p\left(k, \theta_k \mid Y\right) \propto p\left(Y \mid k, \vec{A}_k, \vec{\omega}_k, \vec{\alpha}_k, \sigma_k^2\right) p\left(k, \vec{A}_k, \vec{\omega}_k, \vec{\alpha}_k \mid \sigma_k^2\right) p\left(\sigma_k^2\right) \\ & \propto \left(\pi\sigma_k^2\right)^{-N} \exp\left(-\frac{\left(\vec{A}_k - m_k\right)^H M_k^{1}\left(\vec{A}_k - m_k\right)}{\sigma_k^2}\right) \left|\pi\sigma_k^2 \Sigma_k\right|^{-k} \exp\left(-\frac{Y_k^H Q_k Y_k}{\sigma_k^2}\right) \\ & \bullet U\left(\vec{\omega}_k; \Omega_k\right) \bullet DU\left(\vec{\alpha}_k; \Psi_k\right) \bullet \frac{\Lambda^k}{k!} \exp\left(-\Lambda\right) \bullet Ig\left(\sigma_k^2; v_0/2, \gamma_0/2\right) \end{aligned} \tag{6}$$

where

$$\begin{cases} M_k^{-1} = D_k^H D_k + \Sigma_k^{-1} \\ m_k = M_k D_k^H Y \\ Q_k = I_N - D_k^H M_k D \end{cases} \tag{7}$$

The nuisance parameters $\vec{A}_k$ and σ_k^2 can be integrated out from (6).

$$p\left(k, \vec{\omega}_k, \vec{\alpha}_k \mid Y\right) \propto \left(\gamma_0/2 + Y^H Q_k Y\right)^{-(N+v_0/2)} \bullet \frac{\left(\Lambda / \left(1+\delta^2\right)\right)^k}{k!} \bullet U_{\Omega_k}\left(\vec{\omega}_k\right) \bullet DU_{\Psi_k}\left(\vec{\alpha}_k\right) \tag{8}$$

Expression (8) can not be used in (3) and (4) to get the solution directly because: ①the *a posteriori* probability is nonlinear with the unknown parameters and can not be integrated analytically; ②there is still an unknown coefficient in (8) so that it can not be directly used in (4). RJ-MCMC algorithm provides solution to these problems. It sets an ergodic Markov chain $\left(k^{(i)}, \theta_{k^{(i)}}\right)_{i \in \mathbb{N}}$ whose stable distribution is equal to $p\left(k, \vec{\omega}_k, \vec{\alpha}_k \mid Y\right)$ and then obtains the statistical inference of the parameters from the P samples of the chain.

$$\hat{k} = \arg\max_{k \in Z_K}\left(\hat{p}(k \mid Y)\right) = \arg\max_{k \in Z_K}\left(\frac{1}{P}\sum_{i=1}^{P} \mathrm{II}_{\{k\}}\left(k^{(i)}\right)\right) \tag{9}$$

$$\hat{\theta}_k = \hat{E}(\theta_k \mid Y, k) = \frac{\sum_{i=1}^{M} \theta_{k^{(i)}}^i II_{\{k\}}\left(k^{(i)}\right)}{\sum_{i=1}^{M} II_{\{k\}}\left(k^{(i)}\right)} \tag{10}$$

3. RJ-MCMC ALGORITHM FOR JOINT MODEL SELECTION AND PARAMETER ESTIMATION OF GTD MODEL

RJ-MCMC differs from the MCMC algorithm in that it permits the sampling process to jump among different subspaces with different model order. Suppose the *a posteriori* probability is $p(\theta, k|Y)$ and the collect of all the possible parameter subspaces is $\Theta = \bigcup_{k=0}^{k_{\max}} k \times \theta_k$, where $k_{\max}$ is the maximum of the possible model order. In each step of the iteration, we first decide the proposal distribution and the candidate samples from this distribution; then we determine whether the candidate samples are acceptable according to the accepting probability [5]

$$\vartheta\left(\theta, \theta^*\right) = \min\left\{1, r\left(\theta, \theta^*\right)\right\} \tag{11}$$

where

$$r\left(\theta, \theta^*\right) = \frac{p\left(\theta^*, k^* \mid Y\right)}{p\left(\theta, k \mid Y\right)} \frac{q\left(\theta, k \mid \theta^*, k^*\right)}{q\left(\theta^*, k^* \mid \theta, k\right)} \tag{12}$$

θ^*, k^* are candidate samples and candidate order; θ, k are the samples and order in the previous iteration. $q(\bullet)$ is the proposal distribution; it should be easy to sample and be nonzero in the support region of $p(\bullet)$.

In each iteration of the RJ-MCMC algorithm, the parameters are updated, dead or born according to the probability u_k, d_k, b_k, and $u_k+d_k+b_k=1$ for all ks .

$$b_k = c\min\left\{1, \frac{p(k+1)}{p(k)}\right\},\quad d_{k+1} = c\min\left\{1, \frac{p(k)}{p(k+1)}\right\} \quad (13)$$

where c adjusts the probability between jump and update, and usually c=0.5. The flow chart of RJ-MCMC algorithm is shown in fig.1 and the update, birth and death move for the GTD model will be explained in the following.

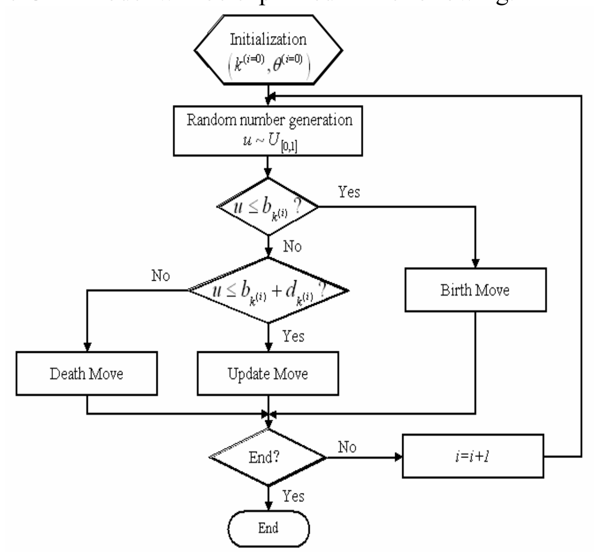

Fig.1 Flow chart of RJ-MCMC algorithm

Suppose we now come to the ith iteration and the current model order is k.

1. Birth Move. In effect if $k<k_{\max}$. A new scattering center is created whose parameters obey $\omega \sim U(-\pi,\pi)$ and $\alpha \sim DU(\{-1,-0.5,0,0.5,1\})$. The collective parameters update to $(k+1, \theta_{k+1})$. The proposal distribution is

$$q\left(\theta_{k+1}^{'}, k+1 \mid \theta_k, k\right) = p\left(k+1\right)\frac{1}{2\pi}\frac{1}{5} \propto \frac{\Lambda^{k+1}}{(k+1)!}\frac{1}{2\pi}\frac{1}{5} \quad (14)$$

and

$$q\left(\theta_k, k \mid \theta_{k+1}^{'}, k+1\right) = p\left(k\right)\Big/\binom{k+1}{1} \propto \frac{\Lambda^k}{k!}\frac{1}{k+1} \quad (15)$$

The accepting probability is:

$$\vartheta_b = \min\left(1, \left(\frac{\gamma_0/2 + Y^H Q_k Y}{\gamma_0/2 + Y^H Q_{k+1} Y}\right)^{(N+v_0/2)} \bullet \frac{1}{\left(1+\delta^2\right)\left(1+k\right)}\right) \quad (16)$$

2. Death Move. In effect if k>0. A scattering center is deleted randomly. The collective parameters become $(k-1, \theta_{k-1})$. The accepting probability is

$$\vartheta_d = \min\left(1, \left(\frac{\gamma_0/2 + Y^H Q_{k-1} Y}{\gamma_0/2 + Y^H Q_k Y}\right)^{-(N+v_0/2)} \left(1+\delta^2\right)k\right) \quad (17)$$

3. Update Move. The model order remains the same while the scattering center parameters are updated using the mixed sampling algorithm [6]. For frequency parameter, the accepting probability for the candidate samples is

$$\vartheta_u = \min\left(1, \left(\frac{\gamma_0/2 + Y^H Q_k Y}{\gamma_0/2 + Y^H Q_k^{'} Y}\right)^{(N+v_0/2)}\right) \quad (18)$$

where $M_k^{'} = M_k \Big|_{\vec{\omega}_k = \vec{\omega}_k^{'}}, m_k^{'} = m_k \Big|_{\vec{\omega}_k = \vec{\omega}_k^{'}}, Q_k^{'} = Q_k \Big|_{\vec{\omega}_k = \vec{\omega}_k^{'}}$ and $\vec{\omega}_k^{'} \triangleq \left[\omega_1, \cdots, \omega_{j-1}, \omega_j^{'}, \cdots, \omega_k\right]^T$.

For $\vec{\alpha}_k$, the accepting probability is

$$\vartheta_u = \min\left(1, \left(\frac{\gamma_0/2 + Y^H Q_k Y}{\gamma_0/2 + Y^H Q_k^{'} Y}\right)^{(N+v_0/2)}\right) \quad (19)$$

where $M_k^{'} = M_k \Big|_{\vec{\alpha}_k = \vec{\alpha}_k^{'}}, m_k^{'} = m_k \Big|_{\vec{\alpha}_k = \vec{\alpha}_k^{'}}, Q_k^{'} = Q_k \Big|_{\vec{\alpha}_k = \vec{\alpha}_k^{'}}$ and $\vec{\alpha}_k^{'} \triangleq \left[\alpha_1, \cdots, \alpha_{j-1}, \alpha_j^{'}, \cdots, \alpha_k\right]^T$.

For the nuisance parameters, we have

$$\sigma_k^2 \mid Y, \vec{\omega}_k, \vec{\alpha}_k \sim Ig\left(v_0/2 + N, \gamma_0/2 + Y^H Q_k Y\right) \quad (20)$$

$$\vec{A}_k \mid Y, \sigma_k^2, \vec{\omega}_k, \vec{\alpha}_k \sim N\left(m_k, \sigma_k^2 M_k\right) \quad (21)$$

It can be proved that no matter what the initial point of the Markov chain is, its distribution will converge to $p\left(k, \theta_k \mid Y\right)$ at a uniform geometric rate [6].But there exists a transient state before the Markov chain converges, and the samples in this transient state should be discarded.

4. SIMULATION RESULTS

The data is produced according to (1) and we set three scattering centers as shown in Table 1 where the first and second scattering centers are in the same Fourier resolution bin. Radar parameters are set to be N=64, Δf=30MHz, f_c=9GHz, $\sigma^2_{RW}=0.2N$. Hyper-parameters are set to be Λ=3, δ^2=1000. SNR is defined as

$$SNR = 10\log_{10}\left(\left|A_1\right|^2 / \sigma^2\right) \quad (22)$$

Table 1 Parameters of the Three Scattering Centers

i	$\lvert A_i\rvert^2$	$-\arg(A_i)$	$\omega_i/2\pi$	α_i
1	20	0	0.2	0.5
2	20	$\pi/4$	0.2+1/N	0
3	20	$\pi/3$	0.2+3/N	-0.5

200 Monte-Carlo simulations are carried out in each of the conditions.

a) Model selection performance

The RJ-MCMC algorithm is compared with other three algorithms using TLS-ESPRIT for parameter estimation and AIC, MDL and MAP criteria for model selection separately.

Fig.2 is the correct model selection rate at different SNR levels. When SNR <2dB, the three algorithms using TLS-ESPRIT all have very poor performance. This is because both the ESPRIT algorithm and the three model selection criteria fail when SNR is low. However, the RJ-MCMC algorithm achieves the correct selection rate above 90% even when SNR=-4dB. This verifies its advantage at low SNR. When SNR level is high, both the RJ-MCMC and the EMAP(Esprit-MAP) algorithm achieves very good performance. This is because both of them are based on the MAP criteria except that the E-MAP algorithm uses an approximation of the *a posteriori* probability which is only valid at high SNR whereas the RJ-MCMC algorithm solves the *a posteriori* probability without approximation.

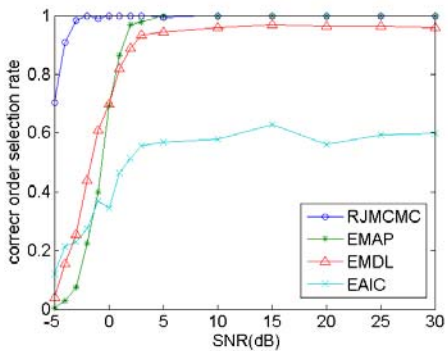

Fig.2 Model order selection performance

b) Parameter estimation performance

Only the realizations with the correct model order decision are used in evaluating the estimation performance.

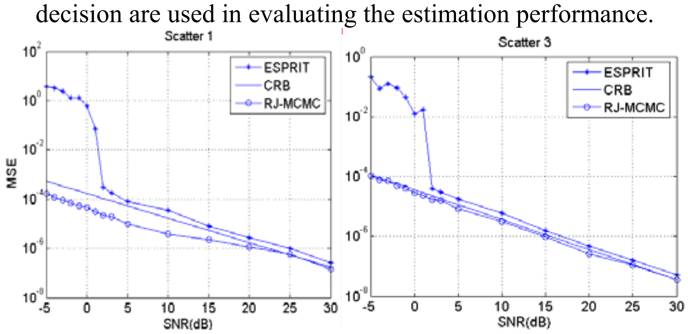

Fig.3 Estimation performance of frequency parameter

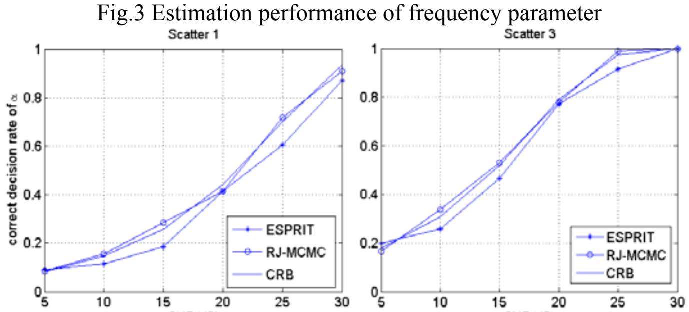

Fig.4 .Correct decision rate of type parameter

Fig. 3 is the Mean Square Error (MSE) of the frequency estimates under different SNR levels. The estimation accuracy of RJ-MCMC is better than TLS-ESPRIT, especially under low SNR and for closely spaced scattering center, e.g., scattering center 1. The performance of RJ-MCMC is even better than the Cramér-Rao Bound (CRB). This is because the CRB is deduced by assuming all the unknowns to be non-random parameters [7] whereas the RJ-MCMC algorithm takes these unknowns to be random variables with known prior distribution. On the contrary, the ESPRIT algorithm which takes the unknowns to be non-random parameters never performs better than the CRB.

Fig.4 gives the correct decision rate of the type parameter. We see that the performance of the RJ-MCMC algorithm is close to the lower bound and is better than that of TLS-ESPRIT. This simulation is carried out at high SNR (since high SNR is necessary for the correct decision of the type parameter [7]) so that the advantage of the RJ-MCMC algorithm is not as evident as in Fig.3.

5. DISCUSSION

When the SNR level is high, the advantage of the RJ-MCMC algorithm over the ESPRIT-MAP algorithm becomes insignificant whereas its computation complexity is still high. So the ESPRIT-MAP algorithm is recommended at high SNR level.

The RJ-MCMC algorithm can also be applied to non-Gaussian noise and colored noise [8, 9]. This generalization will be one of the topics in our future work.

6. REFERENCES

[1] Potter L C, Chiang D M. A GTD-based parametric model for radar scattering, IEEE Trans AP, vol.43, no.10, pp.1058~1067. 1995.

[2] Stoica P. Model-order selection: a review of information criterion rules. IEEE SP Magazine. pp.36~47. 2004.

[3] Bretthorst G L. Bayesian spectrum analysis and parameter estimation. Lecture Notes in Statistics (vol. 48), New York: Springer-Verlag, 1988.

[4] Djuric P M, Li H. Bayesian spectrum estimation of harmonic signals. IEEE SP Lett, vol.2, no.11, pp. 213–215. 1995.

[5] William N, Reilly J P. Wideband array signal processing using MCMC methods. IEEE Trans SP, vol.53, no.2, pp. 411~426. 2005.

[6] Andrieu C, Doucet A. Joint Bayesian model selection and estimation of noisy sinusoids via reversible jump MCMC. IEEE Trans SP, vol.47, no.10, pp.2667-2676. 1999.

[7] Zhou Jianxiong. Theory and Technology on Reconstructing 3D Scattering Centers of Radar Targets in Optical Region, doctoral dissertation (in Chinese), Changsha: NUDT. 2006.

[8] Larocque J R, Reilly J P. Reversible jump MCMC for joint detection and estimation of sources in colored noise. IEEE Trans SP, vol.50, no.2, pp.231~240. 2002.

[9] Kannan B, Fztzgerald W J, Kuruoglu E E. Joint DOA frequency and model order estimation in additive α-stable noise. ICASSP'00 Proc 6, pp.3798~3801.2000.