This article has been accepted for publication in IEEE Transactions on Vehicular Technology. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TVT.2022.3183578

Learning-aided Markov Chain Monte Carlo Scheme For Spectrum Sensing In Cognitive Radio

Zheng Wang, *Member, IEEE,* Shanxiang Lyu, Ling Liu, *Member, IEEE* and Yili Xia, *Senior Member, IEEE*

***Abstract*—In this paper, a learning-aided stochastic strategy is studied for the non-cooperative spectrum sensing in cognitive radio (CR) networks. By sampling from the channel availability distribution in a Markov chain Monte Carlo (MCMC) way, the proposed learning-aided Metropolis-Hastings (LMH) algorithm generates the target channel sequence for fine sensing. The flexible proposal distribution in MH sampling is fully exploited, and a learning mechanism based on the multiple sampling stages within one Markov move is proposed, which tries to take advantages of the samples obtained by the previous stage. The reversibility of the Markov chain in the proposed LMH sampling is studied in detail while its faster convergence rate in the Markov mixing is demonstrated as well, which leads to better spectrum sensing performance and efficiency in cognitive radio.**

Keywords: Learning mechanism, spectrum sensing, Markov chain Monte Carlo, cognitive radio networks

I. Introduction

As a promising technique in wireless communications, cognitive radio (CR) enables an opportunistic access to the available spectrum resources without affecting the primary user networks. Specifically, spectrum sensing is performed firstly to identify the available channels in cognitive radio via cooperative or non-cooperative methods [1]–[4], where the following channel access and data transmission are carried out for a better spectrum utilization given these available channels.

In particular, non-cooperative spectrum sensing aims to return a set of available channels so that the probability of obtaining the desired available channels for fine sensing can be maximized. However, a pressing challenge of non-cooperative schemes comes from the fact that their sensing performance chiefly depends on the parametric environment model by assumption. Since there is a substantial difference between the ideal traffic model and the one in real world, the demanding spectrum sensing performance is severely limited in practice. Due to the rapid development of heterogeneous CR ad hoc networks with high environment dynamics, the assumption of accurate parametric environment model turns out to be more difficult, which leads to the non-parametric-based sensing scheme for the non-cooperative spectrum sensing [5].

In [6], a non-parametric sensing scheme was proposed, where the statistical inference by randomly sampling over channel availability distributions is performed for a better sensing performance. By formulating the problem of seeking the optimal sequence of available channels as to maximize the probability of channel availabilities, the choices of the desired available channels are sampled by the way of Markov chain Monte Carlo (MCMC), where the channel availability distribution is approximated by the mixing of Markov chain. By doing so, the sensing performance is guaranteed due to the fact that the candidate channels with large availabilities will more likely be returned. In [7], another non-parametric spectrum sensing named as the refined Metropolis-Hastings (RMH) is proposed. Based on the MH sampling in MCMC, the conditional proposal distribution is introduced to improve the convergence rate of the Markov mixing, which leads to a better spectrum sensing in cognitive radio networks.

In this paper, the learning-aided Metropolis-Hastings (LMH) algorithm is proposed for non-parametric spectrum sensing to improve its performance and efficiency. Typically, due to the flexibility of the proposal distribution in Metropolis-Hastings (MH) sampling, multiple sampling stages within one Markov move are proposed, where convergence acceleration can be realized by learning from the past samples in the previous stages. More specifically, the reversibility of the designed Markov chain is guaranteed. Based on it, the convergence superiority in Markov mixing is also demonstrated.

II. System Model

Considering spectrum resources with M non-overlapping channels, with respect to the secondary user, a fast sensing is carried out periodically in the way of energy detection. Based on fast sensing, a channel sequence $\mathcal{S} = [s_1, \ldots, s_T]$ is generated and a fine sensing over the spectrum is then performed to facilitate the subsequent channel access and data transmission. In particular, the problem of seeking the optimal channel sequence $\widehat{\mathcal{S}}$ for fine sensing can be reformulated as to maximize the probability of channel availabilities [6]

$$\widehat{\mathcal{S}} = \arg\max_{s_1,\ldots,s_T} P(X_{s_1} = 1, \ldots, X_{s_T} = 1), \quad (1)$$

where the available coefficient $X_{s_i} \in \{0, 1\}$ serves as a binary indicator of the channel s_i. Clearly, an effective way to obtain the above target channel sequence is to randomly sample each

This work was supported in part by the National Natural Science Foundation of China under Grant 61801216 and 61771124, in part by the Natural Science Foundation of Jiangsu Province under Grant BK20180420, in part by the State Key Laboratory of Integrated Services Networks (Xidian University) under Grant No. ISN21-31, in part by the Zhi Shan Young Scholar Program of Southeast University, in part by the Fundamental Research Funds for the Central Universities 2242022k30002.

Z. Wang and Y. Xia are with School of Information Science and Engineering, and with Frontiers Science Center for Mobile Information Communication and Security, Southeast University, Nanjing 210096, China; Z. Wang is also with the State Key Laboratory of Integrated Services Networks, Xidian University, Xian 710071, China (e-mail: z.wang@ieee.org). S. Lyu is with the College of Cyber Security, Jinan University, Guangzhou, 510632, China; Ling Liu is with the College of Computer Science and Software Engineering, Shenzhen University, Shenzhen 518060, China.

element of $\mathcal{S}$ from the following distributions

$$\overline{P}(s_i) = f(s_i)/K, \quad 1 \leq s_i \leq M. \tag{2}$$

Here, $K > 0$ represents a normalized constant, $f(s_i) = \frac{1}{\varpi}\int_0^\infty \int_\lambda^\infty l(s_i)e^{-\frac{\omega}{\varpi}} d_{s_i} d_\omega$ indicates the channel availability outputted by fast sensing, ω and ϖ are the instantaneous signal-to-interference-plus-noise ratio (SINR) and the average SINR respectively, λ is decided by the probability of false alarm, and $l(s_i)$ denotes the probability density function (PDF) of the test statistics of channel s_i [8]

$$l(s_i) = \begin{cases} \frac{1}{2^{k/2}\Gamma(k/2)} s_i^{(k/2)-1} e^{-s_i/2}, & H_0; \\ \frac{1}{2}e^{-(s_i/2+\mu)}(\frac{s_i}{2\mu})^{k/4-0.5} I_{(k/2)-1}(\sqrt{2\mu s_i}), & H_1, \end{cases} \tag{3}$$

where H_0 and H_1 represent the absence and presence of the primary network signals on channel i, k denotes the degrees of freedom, Γ standards for *Gamma* function and I indicates *Bessel* function.

In order to sample from (2), the Metropolis-Hastings (MH) sampling algorithm from MCMC is introduced, which yields the channel sequence $\mathcal{S}$ by a Markov chain $\mathcal{M} = [M_1 = m_1, M_2 = m_2, \ldots]$. Theoretically, the designed Markov chain exponentially converges to the target distribution $\overline{P}(s_i)$

$$\|P^t(m, \cdot) - \overline{P}(\cdot)\|_{TV} \leq \varrho^t, \tag{4}$$

where samples from $\overline{P}(s_i)$ can be obtained after the convergence mixing. Here, t represents the number of Markov move, $\|\cdot\|_{TV}$ indicates the total variation distance, $0 < \varrho < 1$ denotes the convergence rate, M_t and m_t stand for the Markov state at move t and its realization. Specifically, MH sampling applies a proposal distribution $Q(\cdot)$ to invoke an acceptance-rejection mechanism during each Markov move [9]. To summarize, given the channel sequence $[s_1, \ldots, s_{i-1}]$, $1 \leq i \leq T$, the operations of MH sampling algorithm to obtain the channel choice s_i in one Markov move can be described as follows:

Markov mixing:

1) *Given $M_{t+1} = m_t$, sample a candidate state m^* from the proposal distribution $Q(m_{t+1}|m_t)$.*

2) *Compute the acceptance ratio $\alpha(m^*|m_t)$*

$$\alpha(m^*|m_t) = \min\left\{1, \frac{\overline{P}(m^*)Q(m_t|m^*)}{\overline{P}(m_t)Q(m^*|m_t)}\right\}. \tag{5}$$

3) *Accept m^* by the Markov state M_{t+1} (i.e. $M_{t+1} = m^*$) with probability $\alpha(m^*|m_t)$, otherwise let $M_{t+1} = m_t$.*

Channel selection:

4) *If m^* is accepted by M_{t+1} and if $m^* \notin \{s_1, \ldots, s_{i-1}\}$, add $s_i = m^*$ to the channel sequence as $\mathcal{S} = [s_1, \ldots, s_i]$, otherwise try to generate a valid s_i in the next Markov move.*

Overall, the above sampling process is carried out subsequently (i.e., from $i = 1$ to $i = T$) to get the desired channel sequence $\mathcal{S} = [s_1, \ldots, s_T]$ for fine sensing. It is clear to see that the sensing performance is mainly determined by the convergence of the Markov chain as a better sampling approximation of the target distribution $\overline{P}(\cdot)$ can be achieved by a faster Markov mixing.

A salient feature of MH sampling is that its proposal distribution $Q(\cdot)$ can be any fixed distribution from which one can conveniently draw samples. To this end, the conditional proposal distribution is applied by refined Metropolis-Hastings (RMH) in [7] to enable a faster convergence rate with

$$Q'(m^*|m_t) = \frac{Q(m^*|m_t)}{1 - Q(m^* = m_t|m_t)} \tag{6}$$

and

$$\alpha'(m^*|m_t) = \min\left\{1, \frac{\overline{P}(m^*)Q'(m_t|m^*)}{\overline{P}(m_t)Q'(m^*|m_t)}\right\}, \tag{7}$$

where the previous Markov state $M_t = m_t$ is excluded from the candidate state space of m^* in the current sampling for M_{t+1}. Note that each channel choice s_i in sequence $\mathcal{S}$ should be different from each other, and this requirement results in the channel selection shown above. To this end, the selection judgement $m^* \notin \{s_1, \ldots, s_{i-1}\}$ at step (4) is also taken into account as well by the usage of conditional proposal distribution $Q'(m^*|m_t)$.

III. Learning-aided MH Sampling Algorithm

To improve the convergence of the Markov chain for a better spectrum sensing, one can reduce the probability of having the same state over two consecutive Markov moves i.e., $P(M_{t+1} = M_t)$. Intuitively, with a small probability $P(M_{t+1} = M_t)$ (i.e., large $P(M_t \neq M_{t+1})$), the state space of the Markov chain will be explored more efficiently, so as to a faster convergence. Meanwhile, considering the operation of channel selelction, the case $M_{t+1} = M_t$ during Markov moves should also be avoided by failing to satisfy the selection judgement, which is harmful to the sensing efficiency.

A. Algorithm Description

The proposed LMH algorithm employs multiple sampling stages in the generation of the sample for s_i, where the sample candidate rejected by the previous stage can be utilized in a learning way to serve the following sampling in the current stage. From it, a faster convergence rate can be achieved by effectively reducing the probability $P(M_{t+1} = M_t)$.

Basically, $Q'(m^*|m_t)$ in (6) can be viewed as the proposal distribution at the first stage. If m^* generated by $Q'(m^*|m_t)$ is accepted by the Markov state M^{t+1}, then the Markov chain continues the next move in the traditional way. However, once m^* is rejected by the acceptance-rejection mechanism, instead of letting $M^{t+1} = m_t$, a second stage of the proposal distribution is employed to obtain another sample candidate m' as

$$Q''(m'|m_t,m^*) = \frac{Q(m'|m_t)}{1-Q(m'=m_t|m_t)-Q(m'=m^*|m_t)}. \tag{8}$$

Then, another judgement with respect to the acceptance ratio $\alpha''(m'|m_t, m^*)$ should be carried out to decide whether accept the new sample candidate m' by M_{t+1} or not. Because of the second stage of the proposal distribution, there is an extra chance to produce the sample candidate for Markov state M_{t+1}. Meanwhile, besides m_t, m^* is taken into account as well in the sampling of m' for M_{t+1}, which corresponds to a learning mechanism during the Markov mixing.

However, the second stage of the proposal distribution $Q''(m'|m_t, m^*)$ may destroy the *detailed balance* of the

This article has been accepted for publication in IEEE Transactions on Vehicular Technology. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TVT.2022.3183578

underlying MCMC, which also refers to the *reversibility* as an important Markovian property [10]. If reversibility is not satisfied, the global convergence of the underlying chain to the target distribution will never hold, which means the acceptance ratio $\alpha''(m'|m_t, m^*)$ at the second stage should be carefully designed. Therefore, we give the following solution to preserve the stationary distribution

$$\begin{aligned}&\alpha''(m'|m_t,m^*)\\&=\min\left\{1,\frac{P(m')Q'(m^*|m')Q''(m_t|m',m^*)[1-\alpha'(m^*|m')]}{P(m_t)Q'(m^*|m_t)Q''(m'|m_t,m^*)[1-\alpha'(m^*|m_t)]}\right\}\\&=\min\left\{1,\beta\cdot\frac{1-\alpha'(m^*|m')}{1-\alpha'(m^*|m_t)}\right\},\end{aligned}\tag{9}$$

with

$$\beta\triangleq\frac{P(m')Q'(m^*|m')Q''(m_t|m',m^*)}{P(m_t)Q'(m^*|m_t)Q''(m'|m_t,m^*)}.\tag{10}$$

Proposition 1. *With the designed acceptance ratio $\alpha''(m'|m_t, m^*)$ at the second sampling stage, the underlying Markov chain induced by the proposed LMH sampling algorithm is reversible by satisfying*

$$P(m_t)P(m_t, m') = P(m')P(m', m_t).\tag{11}$$

Proof: First of all, we omit the proof of the reversibility with respect to the sampling at the first stage, which is straightforward to confirm. If sample candidate m^* is rejected at the first stage, then the sampling at the second stage is invoked while the transition probability $P(M_t = m_t, M_{t+1} = m')$ of the Markov chain can be written as

$$\begin{aligned}&P(m_t, m')\\&=Q'(m^*|m_t)\cdot[1-\alpha'(m^*|m_t)]\cdot Q''(m'|m_t, m^*)\cdot\alpha''(m'|m_t, m^*),\end{aligned}\tag{12}$$

where the reversibility can be verified by

$$\begin{aligned}&P(m_t)P(m_t, m')\\&=\min\{P(m_t)Q'(m^*|m_t)Q''(m'|m_t, m^*)[1-\alpha'(m^*|m_t)],\\&\qquad P(m')Q'(m^*|m')Q''(m_t|m', m^*)[1-\alpha'(m^*|m')]\}\\&=P(m')P(m', m_t).\end{aligned}\tag{13}$$

This completes the proof. ■

In order to obtain the numerator of the acceptance ratio at the second stage, it seems that one needs to compute the acceptance probability $\alpha'(m^*|m')$. However, this is only a mental trial, which is not implemented in fact. Clearly, if the sampling candidate at the second stage is also rejected, it is possible to move on to a third stage and so on to further exploit the convergence potential. For simplicity, only the second stage in LMH is described through the context while the operations of the extension stages can be designed in the same way. Thanks to the flexible setting of the proposal distributions in MH samplings, the proposal distribution at each stage could be different [11]. Therefore, given the incomplete channel sequence $[s_1, \dots, s_{i-1}]$, the sampling operations in the proposed LMH algorithm for the target choice s_i within one Markov move M_t can be summarised as follows:

***Markov Mixing**:*

Stage 1:

1) *Sample from the proposal distribution $Q'(m^*|m_t)$ in (6) to obtain the candidate m^*.*

2) *Make a judgement based on $\alpha'(m^*|m_t)$ in (7) about whether accept m^* by M_{t+1}.*

a) *if m^* is accepted: go to 5)*

b) *if m^* is rejected: go to the operations at stage 2.*

Stage 2:

3) *Sample from the proposal distribution $Q''(m'|m_t, m^*)$ in (8) to obtain the candidate m'.*

4) *Make a decision for S_{t+1} based on $\alpha''(m'|m_t, m^*)$ in (9) to accept m' or not.*

c) *if m' is accepted: go to 5)*

d) *if m' is rejected: let $M_{t+1} = m_t$, try to obtain a valid s_i at the next Markov move.*

***Channel selection**:*

5) *If m^* or $m' \notin \{s_1, \dots, s_{i-1}\}$, add $s_i = m^*$ or m' into the channel sequence as $\mathcal{S} = [s_1, \dots, s_i]$, otherwise try to obtain a valid s_i at the next Markov move..*

B. Convergence Analysis and Performance Enhancement

Compared to MH and RMH sampling algorithms, LMH sampling learns the prior knowledge obtained at different stages within the same Markov move, which still retains the Markovian property of the Markov mixing. As for the proposed learning mechanism, we show that LMH sampling is able to achieve a faster convergence rate than both RMH and MH samplings. Therefore, a better sampling accuracy can be achieved by LMH sampling in the approximation of the target distribution $\overline{P}(s_t)$, which accounts for an enhanced spectrum sensing performance.

Theorem 1. *Given the target distribution $\overline{P}(\cdot)$, LMH sampling achieves a better convergence performance than MH and RMH samplings by a smaller convergence rate, i.e,*

$$\varrho_{\text{LMH}} < \varrho_{\text{RMH}} < \varrho_{\text{MH}}.\tag{14}$$

Proof: To perform the convergence analysis about the Markov mixing, the probability $P(M_t = M_{t+1})$ between two consecutive Markov states is evaluated.

Specifically, given $M_t = m_t$, the probability $P(M_t = M_{t+1})$ in the original MH sampling consists of two parts. On one hand, m_t will be accepted by M_{t+1} if it is sampled from the proposal distribution Q while the acceptance ratio α in (5) equals to 1 in this condition. On the other hand, m_t will be accepted by M_{t+1} if the sampled candidate $m^* \neq m_t$ from the proposal distribution Q is rejected. Therefore, it follows that

$$\begin{aligned}&P_{\text{MH}}(M_t = M_{t+1}) = P_{\text{MH}}(M_{t+1} = m_t|M_t = m_t)\\&=Q(m_t|m_t)\alpha(m_t|m_t)+\sum_{m^*\neq m_t}Q(m^*|m_t)(1-\alpha(m^*|m_t))\\&=Q(m_t|m_t)+\sum_{m^*\neq m_t}Q(m^*|m_t)(1-\alpha(m^*|m_t))\\&=1-\sum_{m^*\neq m_t}Q(m^*|m_t)\alpha(m^*|m_t)\end{aligned}\tag{15}$$

As for RMH sampling, according to the conditional proposal distribution Q' in (6), the choice m_t is removed from the candidate list of M_{t+1}. For this reason, its probability $P(M_t = M_{t+1})$ can be expressed as

$$\begin{aligned} P_{\text{RMH}}(M_t = M_{t+1}) &= \sum_{m^*} Q'(m^*|m_t)(1 - \alpha'(m^*|m_t)) \\ &= \sum_{m^*} Q'(m^*|m_t) - \sum_{m^*} Q'(m^*|m_t)\alpha'(m^*|m_t) \\ &= 1 - \sum_{m^* \neq m_t} \min\left\{ \frac{Q(m^*|m_t)}{1 - Q(m^* = m_t|m_t)}, \frac{\overline{P}(m^*)Q(m_t|m^*)}{\overline{P}(m_t)(1 - Q(m_t = m^*|m^*))} \right\} \\ &< 1 - \sum_{m^* \neq m_t} \min\left\{ Q(m^*|m_t), \frac{\overline{P}(m^*)Q(m_t|m^*)}{\overline{P}(m_t)} \right\} \\ &= P_{\text{MH}}(M_t = M_{t+1}). \end{aligned} \tag{16}$$

Next, regarding to $P(M_t = M_{t+1})$ of the proposed LMH sampling, we have

$$\begin{aligned} P_{\text{LMH}}(M_t = M_{t+1}) &= \sum_{m^*} [Q'(m^*|m_t)(1 - \alpha'(m^*|m_t)) \cdot \gamma] \\ &< \sum_{m^*} Q'(m^*|m_t)(1 - \alpha'(m^*|m_t)) \\ &= P_{\text{RMH}}(M_t = M_{t+1}), \end{aligned} \tag{17}$$

where the inequality holds because of the coefficient γ

$$\gamma = \sum_{m'} [Q''(m'|m_t, m^*)(1 - \alpha''(m'|m_t, m^*))] < 1. \tag{18}$$

Therefore, it follow that

$$P_{\text{LMH}}(M_t{=}M_{t+1}) < P_{\text{RMH}}(M_t{=}M_{t+1}) < P_{\text{MH}}(M_t{=}M_{t+1}) \tag{19}$$

which corresponds to

$$P_{\text{LMH}}(M_t{\neq}M_{t+1}) > P_{\text{RMH}}(M_t{\neq}M_{t+1}) > P_{\text{MH}}(M_t{\neq}M_{t+1}). \tag{20}$$

Intuitively, this means the Markov states in LMH are more dynamic than those in RMH and MH so as to a better convergence performance. More precisely, each off-diagonal element in the transition matrix $\mathbf{P}_{\text{LMH}}$ is always larger than those of $\mathbf{P}_{\text{RMH}}$ and $\mathbf{P}_{\text{MH}}$ while accordingly each diagonal element in the transition matrix $\mathbf{P}_{\text{LMH}}$ is always smaller than those of $\mathbf{P}_{\text{LMH}}$ and $\mathbf{P}_{\text{MH}}$. According to literatures of MCMC, this is named as *Peskun ordering* written by [12]

$$P_{\text{LMH}}(M_t, M_{t+1}) \succeq P_{\text{RMH}}(M_t, M_{t+1}) \succeq P_{\text{MH}}(M_t, M_{t+1}). \tag{21}$$

We now recall the following Lemma to specify the relation between Peskun ordering and convergence rate.

Lemma 1 ([12]). *Given reversible Markov chains P and G with stationary distribution π, if $P \succeq G$, then their convergence rates satisfy $\varrho_P \leq \varrho_G$.*

The definition of Peskun ordering $P(M_t, M_{t+1}) \succeq G(M_t, M_{t+1})$ is based on the inequality $P(M_t, M_{t+1}) \geq G(M_t, M_{t+1})$, where the equality $\varrho_P = \varrho_G$ holds only if $P(M_t, M_{t+1}) = G(M_t, M_{t+1})$. Since the case of equality is excluded here, we can immediately obtain $\varrho_{\text{LMH}} < \varrho_{\text{RMH}} < \varrho_{\text{MH}}$ based on (21) and Lemma 1. ∎

From Peskun ordering, a Markov chain has smaller probability of staying in the same state (i.e., $P(M_t = M_{t+1})$) will explore the state space more efficiently, which results in a better Markov mixing [13]. Therefore, the performance of spectrum sensing will be enhanced by the proposed LMH sampling algorithm accordingly. Meanwhile, with the increasing probability $P(M_t \neq M_{t+1})$, the sensing efficiency will also be strengthened as the risk of failing to satisfy the selection judgement m^*or $m' \notin \{s_1, \ldots, s_{i-1}\}$ is reduced.

C. The Choice of the Proposal Distribution Q

To reduce the computational cost of each Markov move, the conditional symmetric Gaussian proposal distribution Q is applied in proposed LMH sampling algorithm as

$$Q(m^*|m_t) = \frac{e^{-\frac{1}{2\sigma^2}|m^* - m_t|^2}}{\sum_{m^*} e^{-\frac{1}{2\sigma^2}|m^* - m_t|^2}} = Q(m_t|m^*), \tag{22}$$

where the standard deviation $\sigma > 0$ is flexible to choose to adjust the Markov mixing. In this way, the proposal distribution $Q'(m^*|m_t)$ in (6) becomes symmetric as well, which leads to a simplified acceptance ratio at the first sampling stage

$$\alpha'(m^*|m_t) = \min\left\{1, \frac{\overline{P}_d(m^*)}{\overline{P}_d(m_t)}\right\}. \tag{23}$$

Clearly, m^* is accepted by M^{t+1} if $\overline{P}_d(m^*) > \overline{P}_d(m_t)$, otherwise it will be accepted with probability $\frac{\overline{P}_d(m^*)}{\overline{P}_d(m_t)}$. If m^* is rejected, then another candidate m' will be sampled from the proposal distribution $Q''(m'|m_t, m^*)$ with $m' \neq m^* \neq m_t$, m_t and m^* are learnt as the prior knowledge at the second stage to enhance the convergence. Due to the symmetry, the acceptance ratio $\alpha''(m'|m_t, m^*)$ at the second stage becomes

$$\alpha''(m'|m_t, m^*) = \min\left\{1, \frac{\overline{P}_d(m')\left[1 - \min\left[1, \frac{\overline{P}_d(m^*)}{\overline{P}_d(m')}\right]\right]}{\overline{P}_d(m_t) - \overline{P}_d(m^*)}\right\}, \tag{24}$$

which can be further expressed as

$$\begin{aligned} \alpha''(m'|m_t, m^*) &= \min\left\{1, \frac{\max\{0, \overline{P}_d(m') - \overline{P}_d(m^*)\}}{\overline{P}_d(m_t) - \overline{P}_d(m^*)}\right\} \\ &= F\left(\frac{\overline{P}_d(m') - \overline{P}_d(m^*)}{\overline{P}_d(m_t) - \overline{P}_d(m^*)}\right), \end{aligned} \tag{25}$$

where F denotes the *cumulative distribution function* (CDF) of a uniform random variable over the interval $(0, 1)$. To make it specific, the following three cases are summarized:

1) *if $\overline{P}_d(m') > \overline{P}_d(m_t)$, then $\alpha''(m'|m_t, m^*) = 1$, accept m' and let $M^{t+1} = m'$.*

2) *if $\overline{P}_d(m_t) > \overline{P}_d(m') > \overline{P}_d(m^*)$, accept m' with probability $\frac{\overline{P}_d(m') - \overline{P}_d(m^*)}{\overline{P}_d(m_t) - \overline{P}_d(m^*)}$.*

3) *if $\overline{P}_d(m') < \overline{P}_d(m^*)$, then $\alpha''(m'|m_t, m^*) = 0$, reject m' and let $M^{t+1} = m_t$.*

Based on the symmetric Gaussian distribution, the generations of candidate samples m^* and m' from Q' and Q'' respectively are significantly simplified while the acceptance ratio $\alpha''(m'|m_t, m^*)$ of the second stage in (25) becomes much more straightforward than that in (9). From this, the sampling with respect to the average energy detection probability $\overline{P}_d(\cdot)$ can be easily carried out. In addition, if m' is also rejected, it is possible to move on to the next stage and obtain another

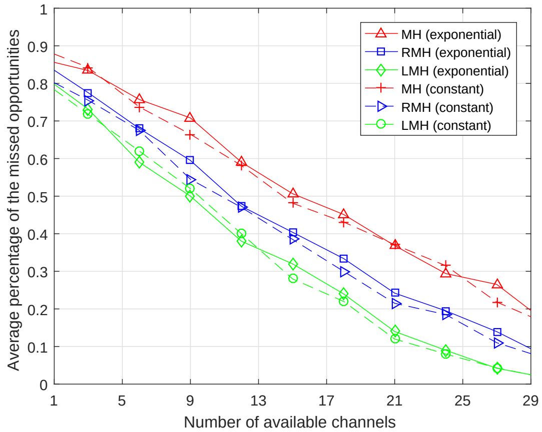

Fig. 1. Average percentage of the missed opportunities under different traffic models.

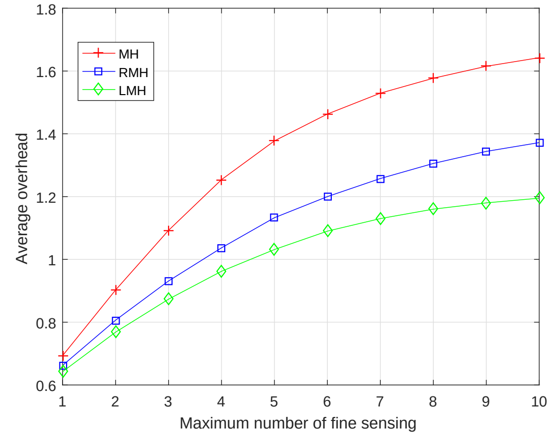

Fig. 2. Average overhead of obtaining an available channel with different maximum numbers of fine sensing.

candidate from the proposal distribution $Q'''(\cdot)$ based on the previous samples m_t, m^* and m'. We emphasize that the coefficient γ in (18) will be getting smaller with the extension of multiple sampling stages, implying a better convergence if more sampling stages are applied.

IV. Simulation Results

To confirm the performance and efficiency gains of the proposed LMH sampling algorithm for spectrum sensing, the following simulations are carried out over $M = 40$ channels in the licensed spectrum with SINR= 20 dB. Meanwhile, the the probability of false alarm is set as 0.01.

Fig. 1 evaluates the average percentages of the missed spectrum opportunities for different spectrum sensing schemes, where two different traffic models (i.e., exponential packet arrival with an average arrival rate of 10 arrival/sec and constant packet arrival with an average arrival rate of 10 arrival/sec) are applied for a better comparison with $n_{\max} = 8$. Here, $n_{\max}$ indicates the maximum number of channels for fine sensing, and the percentage of the missed spectrum opportunities is defined as $p_m = 1 - \frac{n^*}{n_{\max}}$, where n^* is the obtained number of available channels, n_{ava} denotes the average number of available channels, n_{req} stands for the number of request available channels [6]. Clearly, under different traffic models, the proposed LMH sampling achieves the smallest probabilities of missing opportunities than both MH and RMH sampling schemes, which is accordance with the results of convergence.

As for sensing efficiency, the sensing overhead in obtaining the available channels is applied, which is defined as $o = n_{\max} - n^*$ for $n^* = n_{\text{req}}$ and $o = n_{\max}$ for $n^* \neq n_{\text{req}}$. According to it, the average overhead to obtain an available channel (i.e., $n^* = 1$) for fine sensing is presented in Fig. 2. In particular, the average sensing overhead of the proposed LMH sampling is smaller than those of MH and RMH sampling schemes, which comes from the a higher probability $P(M_t \neq M_{t+1})$ in channel selection. To make it more specific, Table I is given to show the acceptance rates in selection judgement for these MCMC-based sampling schemes, and the proposed LMH sampling entails a most efficient sensing process by the highest acceptance ratio for all cases of i.

TABLE I
Average Acceptance Probabilities by Channel Sequence $\mathcal{S}$.

	$i=1$	$i=2$	$i=3$	$i=4$	$i=5$	$i=6$
MH	1	0.71	0.51	0.41	0.36	0.27
RMH	1	0.84	0.75	0.69	0.61	0.56
LMH	1	0.91	0.83	0.75	0.69	0.63

V. Conclusion

In this paper, the learning-aided Metropolis-Hastings (LMH) sampling algorithm is proposed for non-cooperative spectrum sensing in cognitive radio networks. By learning the knowledge from the designed multiple sampling stages within one Markov move, the probability that the Markov chain stays at the same state will be reduced significantly. Therefore, a faster convergence rate in Markov mixing can be achieved, which brings considerable performance and efficiency gains to spectrum sensing.

References

[1] A. Ali and W. Hamouda, "Advances on spectrum sensing for cognitive radio networks: Theory and applications," *IEEE Communications Surveys Tutorials*, vol. 19, no. 2, pp. 1277–1304, 2017.
[2] A. Patel *et al.*, "Non-antipodal signaling based robust detection for cooperative spectrum sensing in MIMO cognitive radio networks," *IEEE Signal Processing Letters*, vol. 20, no. 7, pp. 661–664, 2013.
[3] L. Arienzo and D. Tarchi, "Statistical modeling of spectrum sensing energy in multi-hop cognitive radio networks," *IEEE Signal Processing Letters*, vol. 22, no. 3, pp. 356–360, 2015.
[4] H. Sun, A. Nallanathan, S. Cui, and C.-X. Wang, "Cooperative wideband spectrum sensing over fading channels," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 3, pp. 1382–1394, 2016.
[5] F. Shen, G. Ding, Z. Wang, and Q. Wu, "UAV-based 3D spectrum sensing in spectrum-heterogeneous networks," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 6, pp. 5711–5722, 2019.
[6] X. Y. Wang, A. Wong, and P. Ho, "Spectrum sensing in cognitive radio using a Markov-chain Monte-Carlo scheme," *IEEE Communications Letters*, vol. 14, no. 9, pp. 830–832, 2010.
[7] Z. Wang *et al.*, "Dynamic Markov chain Monte Carlo-based spectrum sensing," *IEEE Signal Processing Letters*, vol. 27, pp. 1380–1384, 2020.
[8] C. Vladeanu, C. Nastase, and A. Martian, "Energy detection algorithm for spectrum sensing using three consecutive sensing events," *IEEE Wireless Communications Letters*, vol. 5, no. 3, pp. 284–287, 2016.
[9] D. A. Levin, Y. Peres, and E. L. Wilmer, *Markov Chains and Mixing Time*, American Mathematical Society, 2008.

[10] S. P. Meyn and R. L. Tweedie, *Markov chains and stochastic stability*. UK, Cambridge University Press, 2009.

[11] J. S. Liu, *Monte Carlo Strategies in Scientific Computing*, New York: Springer-Verlag, 2001.

[12] A. Mira, “Ordering and improving the performance of Monte Carlo Markov chains,” *Statistical Science*, vol. 16, no. 4, pp. 340–350, 2001.

[13] Z. Wang, Y. Huang, and S. Lyu, “Lattice-reduction-aided Gibbs algorithm for lattice Gaussian sampling: Convergence enhancement and decoding optimization,” *IEEE Transactions on Signal Processing*, vol. 67, no. 16, pp. 4342–4356, Aug 2019.