Capacity Limits for a Cognitive Radio Network under Fading Channel [⋆]

Yuehong Gao[1,2], Jinxing Yang[1], Xin Zhang[1], and Yuming Jiang[2]

[1] Wireless Theories and Technologies (WT&T) Lab
Beijing University of Posts and Telecommunications (BUPT), 100876 Beijing, China
[2] Center for Quantifiable Quality of Service in Communication Systems (Q2S)
Norwegian University of Science and Technology (NTNU), 7491 Trondheim, Norway

Abstract. In this paper, performance evaluation of a cognitive radio network is conducted. The analysis is based on stochastic network calculus. The system is supposed to work in a Time Division Multiple Access (TDMA) mode with fixed slot length. The wireless channel is modeled as a Gilbert-Elliott (GE) fading channel, where the channel quality transits between state *ON* and state *OFF* according to a Markov chain. Spectrum sensing errors, which can be classified into mis-detection and false-alarm, are taken into consideration. Particularly, a stochastic arrival curve for spectrum sensing error process, and a stochastic service curve for GE channel, are derived. In addition, performance distribution bounds are obtained based on stochastic network calculus. Furthermore, numerical calculations are made to show the capacity limits under delay constraints.

Keywords: Capacity, Cognitive Radio, GE Channel, Performance Bound, Stochastic Network Calculus

1 Introduction

Nowadays, cognitive radio has become a promising technology, since it provides a solution to improve the spectrum utilization efficiency. In a cognitive radio network, the secondary users (SUs) sense the spectrum before transmitting on it, and if they find available spectrum holes, they will make use of those resources. The sensing results, however, may not exactly match with the real condition. In other words, spectrum sensing error happens sometimes, which leads to collision between the primary transmission and the secondary transmission or waste of transmission opportunities for secondary users. Therefore, physical layer re-transmission is needed in order to deal with such collisions.

In this paper, we consider a cognitive radio network with two classes of input traffic, the aggregated flow from primary users and the one from secondary users, as shown in Fig.1. The system works in a slotted mode with fixed slot length

---

T. The flow from PUs has higher priority over SUs flow to be served. Secondary users try to sense the channel and act based on the sensing results. Sensing errors may happen and will affect the performance.
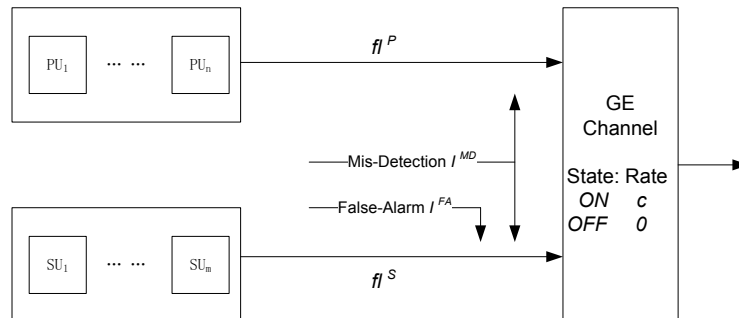

Fig. 1. System Model

How to analyze the performance guarantee for each class of users is a key issue in cognitive radio networks. Some queueing theory based studies have been made such as in [1], where delay and queue related parameters are derived. In [2–4], we made some analysis by using network calculus to derive the backlog and delay distribution bounds for a simplified system model. Network calculus is an approach to deal with flow problems in communication networks, which was introduced by Cruz in 1991 [5]. After about 20-year development, network calculus has evolved into two branches: the deterministic branch and the stochastic branch. In this study, stochastic network calculus is employed. In stochastic network calculus, stochastic arrival curve (sac) and stochastic service curve (ssc) are used to describe the characteristics of a flow and a server, respectively. Based on sac and ssc, performance bounds can be derived.

In [2–4], the wireless channel was assumed to be a constant error-free channel. However, this assumption is not practical in real systems, because the essential characteristic of a wireless channel is its fading nature. In this paper, this shortage is overcome by considering the Gilbert-Elliott (GE) channel model [6, 7], which has two states, Good (G) and Bad (B). In the discrete time model, these two states transit between each other according to a Markov chain. In addition, the sensing error process is re-modeled compared with the model in [2], so that tighter performance bounds can be obtained.

The paper is organized as follows. In Section 2, stochastic network calculus basics are introduced, where the stochastic arrival curve for sensing error process and the stochastic service curve for GE fading channel are derived, and the delay bound for each flow is also obtained. Section 3 discusses the numerical results and capacity limits, followed by a summary in Section 4.

2 Stochastic Network Calculus Analysis

2.1 Traffic Modeling

Stochastic arrival curve can be defined from different aspects [8]. Here, we explore the *virtual-backlog-centric* based definition as follows.

Definition 1. (Stochastic Arrival Curve). *A flow $A(t)$ is said to have a virtual-backlog-centric (v.b.c) stochastic arrival curve $\alpha \in F^3$ with bounding function $f \in \bar{F}^4$, denoted by $A(t) \sim_{sac} \langle f, \alpha \rangle$, if for all $t \geq 0$ and all $x \geq 0$ there holds:*

$$P\left\{ \sup_{0 \leq s \leq t} \{A(s,t) - \alpha(t-s)\} > x \right\} \leq f(x), \tag{1}$$

where $A(s,t)$ denotes the cumulative amount of traffic during period $(s,t]$, $A(0,t)$ is written as $A(t)$ for short, and $\alpha(t)$ is a non-decreasing function.

Stochastic arrival curves of many traffic models have been derived, such as in [8]. Therefore, we just employ the models directly in this paper, and put our efforts on other aspects.

2.2 Modeling of Spectrum Sensing Error Process

In cognitive radio networks, secondary users sense the spectrum and utilize the available white spaces for their transmissions. However, sensing errors may happen, which lead to *transmission collision* or *opportunity waste*. To be specific, sensing errors can be classified into two types, mis-detection (MD) and false-alarm (FA). Mis-detection means that the spectrum is occupied by PUs but the spectrum sensing result says it is available for SUs, which will result in transmission collision and influence both PUs' and SUs' current transmission. However, false alarm occurs in the opposite way, when SUs believe that the spectrum is being used by PUs but actually the spectrum is idle. As a result, SUs will miss those transmission opportunities.

Based on the facts described above, the error process can be considered as a special type of input traffic, which also competes for the transmission resource and has the highest priority.In this part, the stochastic arrival curve for sensing error process will be derived.

Here, we consider a slotted system with fixed slot length T, and the probability that sensing error happens in one time slot is supposed to be p. By further assuming the independency between the appearances of sensing errors in adjacent slots, the impairment arrival process $I(t)$ is a Lévy process, where $I(t)$ denotes the number of sensing errors during slot $(0,t]$. Then, according to Lemma 1 in the appendix, process $I(t)$ has a v.b.c stochastic arrival curve, denoted

---

[3] F: the set of non-negative wide-sensing increasing functions
[4] $\bar{F}$: the set of non-negative wide-sensing decreasing functions

by $I(t) \sim_{sac} \langle f^I, \alpha^I \rangle$, where

$$f^I(x) = e^{-\theta\theta_1} e^{-\theta x} \tag{2}$$

$$\alpha^I(t) = \left[\frac{1}{\theta} \log E[e^{\theta I(1)}] + \theta_1 \right] \cdot t \equiv [\rho^I(\theta) + \theta_1] \cdot t \tag{3}$$

for free parameters $\forall \theta_1 \geq 0$ and $\forall \theta > 0$.

In each slot, the happening of sensing error has a Bernoulli distribution with parameter p. Therefore, $\rho^I(\theta)$ in Eq.(3) can be expressed as

$$\rho^I(\theta) = \frac{1}{\theta} \log(1 - p + pe^{\theta\sigma}), \tag{4}$$

where σ denotes the number of packets that are not transmitted successfully in a slot due to a sensing error. Furthermore, mis-detection process and false-alarm process have the same characteristic as the sensing error process, and the only difference is the happening probability. In later parts, the following notations are used to represent the stochastic arrival curves of mis-detection process and false-alarm process:

$$I^{MD}(t) \sim_{sac} \langle f^{MD}, \alpha^{MD} \rangle, \ I^{FA}(t) \sim_{sac} \langle f^{FA}, \alpha^{FA} \rangle, \tag{5}$$

where f^{MD} and f^{FA} have the same form as in Eq.(2), α^{MD} and α^{FA} can be obtained by replacing the probability p in Eq.(3) with p^{MD} and p^{FA}, respectively.

2.3 Server Modeling

Similar to the concept of stochastic arrival curve, stochastic service curve is defined to describe the service guarantee that a server can provide, and several different definitions have been proposed. Here, we employ the following one [8].

Definition 2. (Stochastic Service Curve). *A system S is said to provide a stochastic service curve $\beta \in F$ with bounding function $g \in \bar{F}$, denoted by $S \sim_{ssc} \langle g, \beta \rangle$, if for all $t \geq 0$ and all $x \geq 0$ there holds:*

$$P\{A \otimes \beta(t) - A^*(t) > x\} \leq g(x). \tag{6}$$

Here, $A \otimes \beta(t) \equiv \inf_{0 \leq s \leq t} \{A(s) + \beta(t - s)\}$, and $A^*(t)$ denotes the cumulative amount of output traffic up to time t.

The Gilbert-Elliott channel model is named after the originators, which can be further classified into discrete-time and continuous-time model. In this paper, the discrete time model is considered, since it matches well with the slotted system model. Fig.2 shows a two-state GE channel, where the channel can either be in ON state (state 1), in which data can be decoded error-free (if no collision happens during the transmission), or in state OFF (state 0), in which the channel quality is too bad to transmit any data. The channel state transits among the two states as a Markov process with transition matrix of Q, where q_{ij} denotes the transition probability from state i to state j ($i, j \in \{0, 1\}$).

Let $S(t)$ denote the service provided by the channel during $(0, t]$. Then, there are two cases.
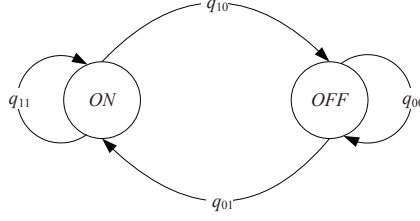
Fig. 2. Discrete-time two-state Gilbert-Elliott channel model

– Case 1: t is not within any backlogged period. In this case, there is no backlog in the system at time t, which means that all traffic that arrived up to time t has left the server. Hence, $A^*(t) = A(t)$ and consequently $A \otimes \beta(t) - A^*(t) = A(t) + \beta(0) - A^*(t) = 0$.

– Case 2: t is within a backlogged period $(t_0, t_b]$, where t_0 is the start point of the backlogged period. Then, $A^*(t_0) = A(t_0)$ and

$$A \otimes \beta(t) - A^*(t) \leq A(t_0) + \beta(t - t_0) - A^*(t) \tag{7}$$

$$= \beta(t - t_0) + A^*(t_0) - A^*(t) = \beta(t - t_0) - S(t_0, t) \tag{8}$$

Then, we have:

$$P\{A \otimes \beta(t) - A^*(t) > x\} \leq P\{\beta(t - t_0) - S(t_0, t) > x\}$$
$$\leq P\{e^{\theta[\beta(t-t_0)-S(t_0,t)]} > e^{\theta x}\} \leq e^{-\theta x} E[e^{\theta[\beta(t-t_0)-S(t_0,t)]}]$$
$$\leq e^{-\theta x} E[e^{\theta[\mu(\theta)\cdot(t-t_0)-S(t_0,t)]}]$$
$$= e^{-\theta x} E[e^{\theta[\mu(\theta)\cdot\tau-S(\tau)]}] \leq e^{-\theta x}$$

where the third step is known as Chernoff bound, and the fourth step is due to that $S(t)$ is stationary, and

$$\mu(\theta) \equiv -\frac{1}{\theta\tau} \log E[e^{-\theta S(\tau)}]$$

which is known as the *effective bandwidth* of process S in the literature [9] [10]. For the two-state Markov chain of the considered GE channel, its effective bandwidth has an explicit form [10], which is adopted in this paper as

$$\hat{\mu}(\theta) = \frac{1}{-\theta} \log \left(\frac{q_{00} + q_{11}e^{-c\theta} + \sqrt{(q_{00} + q_{11}e^{-c\theta})^2 - 4(q_{11} + q_{00} - 1)e^{-c\theta}}}{2} \right)$$

By combining both cases, a stochastic service curve of GE channel has been found as:

$$S(t) \sim_{ssc} \langle g(x) = e^{-\theta x}, \beta(t) = \mu(\theta) \cdot t \rangle. \tag{9}$$

2.4 Delay Bound

Previous work in [2] has discussed how to obtain the stochastic service curve that can be effectively provided to each input traffic. Here, by using the following notations for input traffic,

$$fl^P: A^P(t) \sim_{sac} \langle f^P(x), \alpha^P(t) \rangle \tag{10}$$

$$fl^S: A^S(t) \sim_{sac} \langle f^S(x), \alpha^S(t) \rangle, \tag{11}$$

and by further assuming Re-Transmission until Success (RT-S) scheme, the stochastic service curve for PUs' traffic and SUs' traffic can be expressed as:

$$fl^P: \quad S^P(t) \sim_{ssc} \langle g^P(x), \beta^P(t) \rangle \tag{12}$$

$$\text{with} \quad g^P(x) = f^{MD} \otimes g(x), \beta^P(t) = \beta - \alpha^{MD}(t) \tag{13}$$

$$fl^S: \quad S^S(t) \sim_{ssc} \langle g^S(x), \beta^S(t) \rangle \tag{14}$$

$$\text{with} \quad g^S(x) = f^I \otimes g \otimes f^P(x), \beta^S(t) = \beta - \alpha^I - \alpha^P(t) \tag{15}$$

Then, based on the performance bound theorem in [2], the delay distribution bound can be summarized as:

Theorem 1. *(Delay Bound)*

$$P\{D^U(t) > h(\alpha^U + x, \beta^U)\} \leq [f^U \otimes g^U(x)]_1, \tag{16}$$

where $U \in \{P, S\}$, $h(\alpha + x, \beta) = sup_{s \geq 0}\{inf\{\tau \geq 0 : \alpha(s) + x \leq \beta(s + \tau)\}\}$ and $[\cdot]_1$ denotes $max(min(\cdot, 1), 0)$

3 Numerical Results

In previous sections, traffic model, server model as well as the considered cognitive radio network model are described with the delay bound theorem as an ending. In this section, specific parameters and configurations will be substituted into the deduction above in order to obtain the capacity limit under certain delay constraints.

The input packet arrival, fl^P and fl^S, are assumed to be Poisson flow. And the stochastic arrival curve for Poisson traffic is defined as follows.

Definition 3. *(Poisson Traffic)*. *Suppose all packets of a flow have the same size L and they arrive according to a Poisson process with mean arrival rate λ. Then the flow has a stochastic arrival curve $A(t) \sim_{sac} \langle f_{Pois}, rt \rangle$ for any $r > \lambda L$ with bounding function [8]:*

$$f_{Pois}(x) = 1 - (1 - a) \sum_{i=0}^{k} \left[\frac{[a(i-k)]^i}{i!} e^{-a(i-k)} \right]$$

where $a = \frac{\lambda L}{r}$ and $k = \lceil \frac{x}{L} \rceil$.

Table 1. QoS Requirements for Different Services in LTE System

Traffic	Delay Budget	Packet Loss Prob
VoIP	$50ms$	10^{-2}
TCP	$250ms$	10^{-4}

The network is supposed to be a LTE system using OFDM technology with slot length of $0.5ms$. In each slot, there are 7 OFDM symbols in time domain, 50 resource blocks (RB) in frequency domain with 12 sub-carriers in each RB. $16QAM$ and $1/3 - rate$ Turbo code are used as the modulation and coding scheme. Then, the packet length for Poisson arrival is set as the effective bits transmitted in an LTE slot, i.e., $5.6kbits$. Based on this assumption, the parameter σ in error process and c in channel model are all equal to 1 packet per slot. State transition probability q_{01} and q_{10} for GE channel are set as 1 and 0.11, respectively. The free parameters, such as θ, are optimized numerically with a tradeoff between acceptable accuracy and tolerable complexity.

Primary traffic flow is supposed to be a VoIP session, which belongs to the Guaranteed Bit Rate (GBR) bearer in LTE system. While secondary traffic flow is set as TCP interaction service, which is non-GBR bearer because of the lower priority in the whole network. Table.1 lists the QoS Class Identifier (QCI) requirements.

In the system model considered here, re-transmission until success mechanism is employed, which means no packet is dropped because of collision or deep channel fading. Packet loss only happens when the sojourn delay exceeds the delay budget. Therefore, the delay constraints can be written as:

$$Constraint\ 1: \qquad P(Delay^P > 100slots) \leq 10^{-2} \qquad (17)$$

$$Constraint\ 2: \qquad P(Delay^S > 500slots) \leq 10^{-4} \qquad (18)$$

In order to fulfill the delay constraints, there exist an *upper bound* on the *arrival rate* λ of input traffic, which is defined as the *capacity limit* in this paper.

The capacity limit of PUs flow can be expressed as:

$$C^P = max\{\lambda^P, \text{subject to } Constraint\ 1\} \qquad (19)$$

Fig.3(a) shows the delay distribution of PUs input flow calculated from Theorem 1. We can notice that, there is still some capacity margin when the arrival rate of PUs traffic is 1600 packets per second; while delay constraint cannot be met when the arrival rate is increased to 1720 packets per second. The maximum arrival rate of primary traffic, also called capacity limit C^P, is 1690 packets per second when the delay constraint can be guaranteed at the same time.

As for the secondary traffic, it can be transmitted when there is no primary traffic. Therefore, the maximum arrival rate of the secondary network has close relationship with the load η of primary network, which is defined as the ratio of

actual arrival rate over the capacity limit, i.e., $\eta = \lambda^P/C^P$. Then, the capacity limit of SUs flow can be expressed as:

$$C^S = max\{\lambda^S | \eta, \text{subject to } \textit{Constraint 2}\} \tag{20}$$

Fig.3(b) shows three delay distribution bounds when η is set as 0. We can notice that $\lambda^S = 1738$ packets per second is the capacity limit C^S under delay constraint 2.

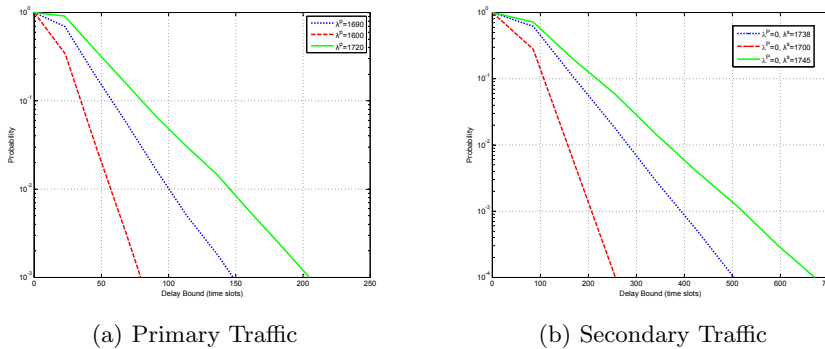(a) Primary Traffic   (b) Secondary Traffic

Fig. 3. Delay Tail Distribution

If we define $C^P = 1690$ and $C^S = 1738$ packets per second as 100% load of the primary network and secondary network, respectively, Fig.4 provides the admissible capacity region of the system, given Poisson arrivals. It is shown that the maximum arrival rate of secondary flow decreases when the load of primary flow increases. Particularly, for any point below the curve, which corresponds to a load of primary traffic and a load of secondary traffic, the system can guarantee the delay requirement and the required loss probability.

4 Conclusion and Discussion

In this paper, capacity limits, defined as the maximum arrival rate, for both primary and secondary traffics in a cognitive radio network are obtained under delay constraints. Stochastic network calculus is relied on to derive the delay distribution bounds, which includes two fundamental concepts: stochastic arrival curve and stochastic service curve. The spectrum sensing error process is analyzed with stochastic arrival curve first. Secondly, Gilbert-Elliott is used to model the fading channel, and its stochastic service curve is derived. And then, specific expressions for delay distribution bounds are obtained. Parameters and configurations in LTE network are used to calculate the numerical results, where the capacity limit of primary traffic and the capacity limit of secondary traffic under different traffic load are discussed.
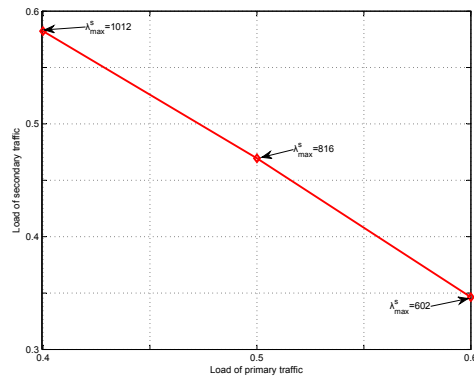
Fig. 4. Capacity Region

In this paper, we have only considered Poisson arrival due to space limitation. However, the analysis can be easily extended to other types of arrivals. Particularly, for many types of traffic, their stochastic arrival curves have been found (e.g. see [11]), with which, the corresponding delay bounds and capacity/throughput regions are readily obtained.

References

1. Suliman, I., Lehtomaki, J.: Queueing analysis of opportunistic access in cognitive radios. In: 2nd International Workshop on Cognitive Radio and Advanced Spectrum Management, pp. 153–157. Aalborg (2009)
2. Gao, Y., Jiang, Y.: Performance Analysis for a Cognitive Radio with Imperfect Spectrum Sensing. In: IEEE INFOCOM 2010 Workshop on Cognitive Wireless Communications and Networking. IEEE Press, San Diego (2010)
3. Gao, Y., Jiang, Y., Lin, T., Zhang, X.: Performance Bounds for a Cognitive Radio Network with Network Calculus Analysis. In: International Conference on Network Infrastructure and Digital Content. Beijing (2010)
4. Gao, Y., Jiang, Y., Zhang, X.: On the Backlog and Delay Distributions of Primary and Secondary Users in a Cognitive Radio Network. China Communications, vol.7, no.5, pp 23-28. (2010)
5. Cruz, R. L.: A calculus for network delay, Part I: Network elements in isolations, and Part II: Network Analysis. In: IEEE Transaction of Information Theory, vol. 37(1), pp: 114–141. IEEE Press (1991)
6. Gilbert, E. N.: Capacity of a bursty-noise channel. Bell System Technical Journal, vol. 39(5), pp. 1253–1265. (1960)
7. Elliott, E. O.: Estimates of error rates for codes on bursty-noise channels. Bell System Technical Journal, vol. 42(9), pp. 1977–1997. (1963)
8. Jiang, Y., Liu, Y.: Stochastic Network Calculus. Springer. (2008)
9. Kelly, F.: Notes on effective bandwidths. In: Stochastic Networks: Theory and Applications, Royal Statistical Society Lecture Notes Series, 4. Oxford University Press (1996)

10. Chang, C.-S.: Performance Guarantees in Communication Networks. Springer-Verlag. (2000)
11. Jiang, Y., Yin, Q., Liu, Y., Jiang, S.: Fundamental Calculus on Generalized Stochastically Bounded Bursty Traffic for Communication Networks. Computer Networks, vol. 53, no. 12, pp. 2011-2021. Elsevier Science. (2009)

Appendix

Lemma 1. *(v.b.c Stochastic Arrival Curve) If an arrival process $A(t)$ has independent stationary increments, then it has a v.b.c stochastic arrival curve $\alpha(t) = [\rho(\theta) + \theta_1] \cdot t$ with bounding function $f(x) = e^{-\theta\theta_1}e^{-\theta x}$ for $\forall \theta_1 \geq 0$, and for $\forall \theta > 0$ and $\rho(\theta) = \frac{1}{\theta}\log E\left[e^{\theta A(1)}\right].$*

Proof. Define a sequence of non-negative random variables $\{V_s\}$ as

$$V_s = e^{\theta A(t-s,t) - \theta[\rho(\theta)+\theta_1]\cdot s}. \tag{21}$$

Since $A(t)$ has independent stationary increments, we then have,

$$V_{s+1} = e^{\theta A(t-s-1,t) - \theta[\rho(\theta)+\theta_1]\cdot(s+1)} \tag{22}$$

$$= e^{\theta \sum_{k=t-s}^{t} X_k - \theta[\rho(\theta)+\theta_1]\cdot(s+1)} \tag{23}$$

$$= V_s \cdot e^{\theta X_{t-s} - \theta[\rho(\theta)+\theta_1]} \tag{24}$$

where $X_k = A(k-1,k)$ is used to simplify the notations. In addition, it is easy to know that X_{t-s} is independent of $X_t, X_{t-1}, ..., X_{t-s+1}$, and it has stationary increments, there holds:

$$E[V_{s+1}|V_1, V_2, ..., V_s] = E[V_{s+1}|X_t, X_{t-1}, ..., X_{t-s+1}] \tag{25}$$

$$= E[V_s \cdot e^{\theta X_{t-s} - \theta[\rho(\theta)+\theta_1]}|X_t, X_{t-1}, ..., X_{t-s+1}] \tag{26}$$

$$= E[V_s|X_t, X_{t-1}, ..., X_{t-s+1}] \cdot E[e^{\theta X_{t-s} - \theta[\rho(\theta)+\theta_1]}] \tag{27}$$

$$= V_s \cdot \frac{E[e^{\theta X_1}]}{\theta\rho(\theta) + \theta\theta_1} \leq V_s \tag{28}$$

Hence, $V_1, V_2, ..., V_t$ form a non-negative supermartingale. Then based on an inequality for supermartingale, Doob's inequality, the definition of $\rho(\theta)$ and $A(t)$ has stationary increments, there holds:

$$P\left\{ \sup_{0 \leq s \leq t} \{A(s,t) - [\rho(\theta) + \theta_1] \cdot (t-s)\} > x \right\} \tag{29}$$

$$= P\left\{ \sup_{0 \leq s \leq t} \{e^{A(s,t) - [\rho(\theta)+\theta_1]\cdot(t-s)}\} > e^x \right\} \tag{30}$$

$$= P\left\{ \sup_{1 \leq s \leq t} V_s > e^{\theta x} \right\} \leq P\{V_1 > e^{\theta x}\} \tag{31}$$

$$\leq e^{-\theta x} E[e^{\theta A(t-1,t) - \theta[\rho(\theta)+\theta_1]}] \leq e^{-\theta x}e^{-\theta\theta_1} \tag{32}$$

which ends the proof.