with $I_{[L_j, R_j)}$ denoting the indicator function of $[L_j, R_j)$ and $-0.5 \leq L_1 < R_1 \leq \dots \leq L_K < R_K \leq 0.5$, $\alpha_j = R_j - L_j$, $P_j > 0$ and $\sum_{j=1}^K P_j = 1$. For $a, b \in [0, 1)$, define

$$\langle a, b \rangle = \begin{cases} [a, b) & \text{if } a < b \\ [0, 1) - [b, a) & \text{if } b < a \\ \{a\} & \text{if } a = b \end{cases} \quad (1.8)$$

and

$$C(x, L, R) = \langle \langle x + L \rangle, \langle x + R \rangle \rangle \quad (1.9)$$

where $\langle \cdot \rangle$ is the modulo-1 operator. By (3.3) in [3], the one-step transition density of $\{U_n^+\}$ is

$$g^+(y|x) = \frac{\partial}{\partial y} P(U_{n+1}^+ \leq y | U_n^+ = x) = \sum_{j=1}^K \frac{P_j}{\alpha_j} I[y \in C(x, L_j, R_j)] \quad (1.10)$$

where $I[y \in C(x, L_j, R_j)]$ denotes the indicator function of $C(x, L_j, R_j)$.

It is of interest to compare TES processes and their time-reversed counterparts which are also stationary Markovian processes (see [26] and [31]). Let $\{U_n^+\}_{n=0}^\tau$ be a TES⁺ subsequence and define its time-reversed version $\{\tilde{U}_n^+\}_{n=0}^\tau$ by $\tilde{U}_n^+ = U_{\tau-n}^+$. Now, from (1.5) we get

$$U_\tau^+ = \langle U_0 + S_\tau^V \rangle \quad (1.11)$$

Applying modulo-1 arithmetic we can isolate U_0 as $U_0 = \langle U_\tau^+ - S_\tau^V \rangle$, or equivalently in terms of the time-reversed subsequence

$$\tilde{U}_\tau^+ = \langle \tilde{U}_0^+ - S_\tau^V \rangle \quad (1.12)$$

A comparison of (1.11) and (1.12) reveals that a time-reversed TES⁺ sequence is also a TES⁺ sequence with an innovation sequence $\{\tilde{V}_n\}$, where $\tilde{V}_n = -V_{\tau-n}$. We refer to the $\tilde{V}_n$ as *reverse innovations*. Indeed, the one-step transition density $\tilde{g}^+$ of $\{\tilde{U}_n^+\}$ is related to $g^+(y|x)$ by $\tilde{g}^+(y|x) = g^+(x|y)$ on their common support, since the marginals of $\{U_n^+\}$ and $\{\tilde{U}_n^+\}$ are both uniform on $[0, 1)$.

National Library
of Canada

Acquisitions and
Bibliographic Services Branch

395 Wellington Street
Ottawa, Ontario
K1A 0N4

Bibliothèque nationale
du Canada

Direction des acquisitions et
des services bibliographiques

395, rue Wellington
Ottawa (Ontario)
K1A 0N4

Microfilm and microfiche editions

Éditions microfilm et microfiche

NOTICE

The quality of this microform is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us an inferior photocopy.

Reproduction in full or in part of this microform is governed by the Canadian Copyright Act, R.S.C. 1970, c. C-30, and subsequent amendments.

AVIS

La qualité de cette microforme dépend grandement de la qualité de la thèse soumise au microfilmage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a conféré le grade.

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de qualité inférieure.

La reproduction, même partielle, de cette microforme est soumise à la Loi canadienne sur le droit d'auteur, SRC 1970, c. C-30, et ses amendements subséquents.

TES/GI/1 Queues:
Exponential Bounding Techniques for
the Delay
(Distribution) and Stability Issues

Fang, Youjian

A Master Thesis

Submitted to

the Faculty of Graduate Studies and Research

in Partial Fulfilment of the Requirement for the Degree of

Master's in Information and System Science

Carleton University

Ottawa

Canada

April, 1995

National Library
of Canada

Acquisitions and
Bibliographic Services Branch

395 Wellington Street
Ottawa, Ontario
K1A 0N4

Bibliothèque nationale
du Canada

Direction des acquisitions et
des services bibliographiques

395, rue Wellington
Ottawa (Ontario)
K1A 0N4

Your file Votre référence

Our file Notre référence

THE AUTHOR HAS GRANTED AN
IRREVOCABLE NON-EXCLUSIVE
LICENCE ALLOWING THE NATIONAL
LIBRARY OF CANADA TO
REPRODUCE, LOAN, DISTRIBUTE OR
SELL COPIES OF HIS/HER THESIS BY
ANY MEANS AND IN ANY FORM OR
FORMAT, MAKING THIS THESIS
AVAILABLE TO INTERESTED
PERSONS.

L'AUTEUR A ACCORDE UNE LICENCE
IRREVOCABLE ET NON EXCLUSIVE
PERMETTANT A LA BIBLIOTHEQUE
NATIONALE DU CANADA DE
REPRODUIRE, PRETER, DISTRIBUER
OU VENDRE DES COPIES DE SA
THESE DE QUELQUE MANIERE ET
SOUS QUELQUE FORME QUE CE SOIT
POUR METTRE DES EXEMPLAIRES DE
CETTE THESE A LA DISPOSITION DES
PERSONNE INTERESSEES.

THE AUTHOR RETAINS OWNERSHIP
OF THE COPYRIGHT IN HIS/HER
THESIS. NEITHER THE THESIS NOR
SUBSTANTIAL EXTRACTS FROM IT
MAY BE PRINTED OR OTHERWISE
REPRODUCED WITHOUT HIS/HER
PERMISSION.

L'AUTEUR CONSERVE LA PROPRIETE
DU DROIT D'AUTEUR QUI PROTEGE
SA THESE. NI LA THESE NI DES
EXTRAITS SUBSTANTIELS DE CELLE-
CI NE DOIVENT ETRE IMPRIMES OU
AUTREMENT REPRODUITS SANS SON
AUTORISATION.

ISBN 0-612-03046-6

Name Youjian Fang

Dissertation Abstracts International is arranged by broad, general subject categories. Please select the one subject which most nearly describes the content of your dissertation. Enter the corresponding four-digit code in the spaces provided.

Mathematics and Statistics

0 4 0 5

U-M-I

SUBJECT TERM

SUBJECT CODE

Subject Categories

THE HUMANITIES AND SOCIAL SCIENCES

COMMUNICATIONS AND THE ARTS

Architecture 0729
Art History 0377
Cinema 0900
Dance 0378
Fine Arts 0357
Information Science 0723
Journalism 0391
Library Science 0399
Mass Communications 0708
Music 0413
Speech Communication 0459
Theater 0465

EDUCATION

General 0515
Administration 0514
Adult and Continuing 0516
Agricultural 0517
Art 0273
Bilingual and Multicultural 0282
Business 0688
Community College 0275
Curriculum and Instruction 0727
Early Childhood 0518
Elementary 0524
Finance 0277
Guidance and Counseling 0519
Health 0680
Higher 0745
History of 0520
Home Economics 0278
Industrial 0521
Language and Literature 0279
Mathematics 0280
Music 0522
Philosophy of 0998
Physical 0523

Psychology 0525
Reading 0535
Religious 0527
Sciences 0714
Secondary 0533
Social Sciences 0534
Sociology of 0340
Special 0529
Teacher Training 0530
Technology 0710
Tests and Measurements 0288
Vocational 0747

LANGUAGE, LITERATURE AND LINGUISTICS

Language
General 0679
Ancient 0289
Linguistics 0290
Modern 0291
Literature
General 0401
Classical 0294
Comparative 0295
Medieval 0297
Modern 0298
African 0316
American 0591
Asian 0305
Canadian (English) 0352
Canadian (French) 0355
English 0593
Germanic 0311
Latin American 0312
Middle Eastern 0315
Romance 0313
Slavic and East European 0314

PHILOSOPHY, RELIGION AND THEOLOGY

Philosophy 0422
Religion
General 0318
Biblical Studies 0321
Clergy 0319
History of 0320
Philosophy of 0322
Theology 0469

SOCIAL SCIENCES

American Studies 0323
Anthropology
Archaeology 0324
Cultural 0326
Physical 0327
Business Administration
General 0310
Accounting 0272
Banking 0770
Management 0454
Marketing 0338
Canadian Studies 0385
Economics
General 0501
Agricultural 0503
Commerce-Business 0505
Finance 0508
History 0509
Labor 0510
Theory 0511
Folklore 0358
Geography 0366
Gerontology 0351
History
General 0578

Ancient 0579
Medieval 0581
Modern 0582
Black 0328
African 0331
Asia, Australia and Oceania 0332
Canadian 0334
European 0335
Latin American 0336
Middle Eastern 0337
United States 0337
History of Science 0585
Law 0398
Political Science
General 0615
International Law and Relations 0616
Public Administration 0617
Recreation 0814
Social Work 0452
Sociology
General 0626
Criminology and Penology 0627
Demography 0938
Ethnic and Racial Studies 0631
Individual and Family Studies 0628
Industrial and Labor Relations 0629
Public and Social Welfare 0630
Social Structure and Development 0700
Theory and Methods 0344
Transportation 0704
Urban and Regional Planning 0999
Women's Studies 0453

THE SCIENCES AND ENGINEERING

BIOLOGICAL SCIENCES

Agriculture
General 0473
Agronomy 0285
Animal Culture and Nutrition 0475
Animal Pathology 0476
Food Science and Technology 0359
Forestry and Wildlife 0478
Plant Culture 0479
Plant Pathology 0480
Plant Physiology 0817
Range Management 0777
Wood Technology 0746
Biology
General 0306
Anatomy 0287
Biostatistics 0308
Botany 0309
Cell 0379
Ecology 0329
Entomology 0353
Genetics 0369
Limnology 0793
Microbiology 0410
Molecular 0307
Neuroscience 0317
Oceanography 0416
Physiology 0433
Radiation 0821
Veterinary Science 0778
Zoology 0472
Biophysics
General 0786
Medical 0760

EARTH SCIENCES

Biogeochemistry 0425
Geochemistry 0996

Geodesy 0370
Geology 0372
Geophysics 0373
Hydrology 0388
Mineralogy 0411
Paleobotany 0345
Paleoecology 0426
Paleontology 0418
Paleozoology 0985
Palynology 0427
Physical Geography 0368
Physical Oceanography 0415

HEALTH AND ENVIRONMENTAL SCIENCES

Environmental Sciences 0768
Health Sciences
General 0566
Audiology 0300
Chemotherapy 0992
Dentistry 0567
Education 0350
Hospital Management 0769
Human Development 0758
Immunology 0982
Medicine and Surgery 0564
Mental Health 0347
Nursing 0569
Nutrition 0570
Obstetrics and Gynecology 0380
Occupational Health and Therapy 0354
Ophthalmology 0381
Pathology 0571
Pharmacology 0419
Pharmacy 0572
Physical Therapy 0382
Public Health 0573
Radiology 0574
Recreation 0575

Speech Pathology 0460
Toxicology 0383
Home Economics 0386

PHYSICAL SCIENCES

Pure Sciences
Chemistry
General 0485
Agricultural 0749
Analytical 0486
Biochemistry 0487
Inorganic 0488
Nuclear 0738
Organic 0490
Pharmaceutical 0491
Physical 0494
Polymer 0454
Radiation 0754
Mathematics 0405
Physics
General 0605
Acoustics 0986
Astronomy and Astrophysics 0606
Atmospheric Science 0608
Atomic 0748
Electronics and Electricity 0607
Elementary Particles and High Energy 0798
Fluid and Plasma 0759
Molecular 0609
Nuclear 0610
Optics 0752
Radiation 0756
Solid State 0611
Statistics 0463
Applied Sciences
Applied Mechanics 0346
Computer Science 0984

Engineering
General 0537
Aerospace 0538
Agricultural 0539
Automotive 0540
Biomedical 0541
Chemical 0542
Civil 0543
Electronics and Electrical 0544
Heat and Thermodynamics 0348
Hydraulic 0545
Industrial 0546
Marine 0547
Materials Science 0794
Mechanical 0548
Metallurgy 0743
Mining 0551
Nuclear 0552
Packaging 0549
Petroleum 0765
Sanitary and Municipal System Science 0790
Geotechnology 0428
Operations Research 0796
Plastics Technology 0795
Textile Technology 0994

PSYCHOLOGY

General 0621
Behavioral 0384
Clinical 0622
Developmental 0620
Experimental 0623
Industrial 0524
Personality 0625
Physiological 0989
Psychobiology 0349
Psychometrics 0632
Social 0451

The undersigned recommend to the Faculty of Graduate Studies
and Research acceptance of the thesis

**TES/GI/1 Queues:
Exponential Bounding Techniques for
the Delay
(Distribution) and Stability Issues**

submitted by Youjian Fang

in partial fulfilment of the requirements for
the degree of Master of Science

I. Lambadaris, Thesis Supervisor

V. Dlab, Chair, Department of Mathematics and Statistics

Carleton University
May 15, 1995

**TES/GI/1 QUEUES:
EXPONENTIAL BOUNDING TECHNIQUES
FOR THE DELAY
(DISTRIBUTION) AND STABILITY ISSUES**

Fang, Youjian

April, 1995

Acknowledgment

I would like to thank Prof. I. Lambadaris, Dr. M. Devetsikiotis and Prof. A. R. Kaye for their direction of my thesis research. I would like to thank Prof. A. R. Kaye, the Department of Systems and Computer Engineering and the Department of Mathematics and Statistics for their financial support for my study and research. Finally I would like to thank the graduate students of Broadband Network Laboratory in the Department of System and Computer Engineering of Carleton University for helping me whenever I encountered difficulties related to the computing facilities of the laboratory.

Abstract

The *TES* (Transform-Expand-Sample) technique has emerged as a general methodology for modeling autocorrelated random processes with arbitrary marginal distributions. Because of their generality and practical applicability, TES models can be readily used to accurately characterize bursty traffic streams in ATM networks (e.g., variable-bit-rate compressed video).

Although TES models can be easily implemented for simulation studies, the need still exists for analytical results on the performance of queueing systems under TES traffic. Of particular interest are the tails of the waiting time distribution in queues driven by TES-modeled bursty traffic. Such tail probabilities, when they become exceedingly small, may be difficult to obtain via conventional simulation.

The main purpose of this thesis is to establish exponential bounds of waiting times of TES/GI/1 queueing systems. Towards to this goal we first consider a general G/GI/1 queue in which the interarrival times form a stationary Markov chain with continuous state space being $[0, \infty)$ and the service times are independent and identically distributed. We then derive a general result for exponential upper bounds for the waiting time. Furthermore, we apply these result to TES/GI/1 queueing systems, and compare our bounds with simulation results. In order to obtain accurate estimates of extremely low probabilities we used fast simulation techniques based on *importance sampling*.

In this thesis we also consider the stability of queueing systems with TES traffic. We derived a Strong Law of Large Numbers for TES sequence. We subsequently use our results to prove stability theorems for the TES/GI/1 and TES/TES/1 queueing systems.

Contents

1	Introduction	5
1.1	Introduction to TES	5
1.2	Distributions of Waiting Times and Queue Lengths	9
1.3	Thesis Overview	16
2	Exponential Bounds of Waiting Times	19
2.1	Exponential Bounds for the Waiting Time Distribution in G/GI/1 Queues .	19
2.2	Application to TES/GI/1 Queues	22
2.3	Proofs of Theorem 2.1.1 and its Corollaries	28
2.3.1	Preliminaries	29
2.3.2	Introductory Lemmas	33
2.3.3	Proof of Theorem 2.1.1	38
3	Strong Law of Large Numbers of TES and Stability of Queueing Systems with TES Traffic	41
3.1	Strong Law of Large Numbers for TES Processes	41
3.2	Stability Results of Queueing Systems with TES Traffic	46

Chapter 1

Introduction

In this chapter we will present the Transform-Expanded-Sample (TES) methodology, some known results on the tail probabilities of waiting time or queue length of queueing systems along with techniques to obtain these results. We will conclude with an overview of this thesis.

1.1 Introduction to TES

The TES (Transform-Expand-Sample) technique was introduced by B. Melamed in 1991 [34], as a versatile methodology for modeling and generating stationary correlated random processes. TES was successfully used to model the highly autocorrelated compressed VBR video traffic (see [35]).

We now introduce a brief outline of TES processes. For a real number x , define the modulo-1 operation, denoted by $\langle \cdot \rangle$, as

$$\langle x \rangle = x - \max\{\text{integer } n; n \leq x\} \quad (1.1)$$

Let $\{V_n\}$ be independent identically distributed (i.i.d.) random variables (r.v.) with a common density $f_V(t)$. Further let U_0 be a r.v. uniformly distributed on $[0, 1)$ and independent

of $\{V_n\}$. Two *background* TES sequences may be defined as follows:

$$U_n^+ = \begin{cases} U_0 & \text{if } n = 0 \\ \langle U_{n-1}^+ + V_n \rangle & \text{if } n \geq 1 \end{cases} \quad (1.2)$$

and

$$U_n^- = \begin{cases} U_n^+ & \text{if } n \text{ even} \\ 1 - U_n^+ & \text{if } n \text{ odd} \end{cases} \quad (1.3)$$

$\{U_n^+\}$ is called TES⁺ process and $\{U_n^-\}$ is called TES⁻ process, and the random variables $\{V_n\}$ are referred to as *innovations*. It can be shown that if U is a r.v. uniformly distributed on $[0,1)$ and V is any r.v independent of U then $U' = \langle U + V \rangle$ is also uniformly distributed on $[0,1)$ (see Lemma 1 of Part I in [25]). It is easy to see that both $\{U_n^+\}$ and $\{U_n^-\}$ form stationary Markov processes of correlated identically distributed (c.i.d.) random variables with uniform marginals on $[0,1)$.

Using simple properties of modulo-1 addition, we can equivalently represent U_n^+ in (1.2) as

$$U_n^+ = \langle U_0 + S_n^V \rangle \quad (1.4)$$

where

$$S_n^V = \sum_{j=1}^n V_j, \quad n \geq 0 \quad (1.5)$$

is the sum of n i.i.d. random variables ($S_0^V = 0$). From (1.3), an analogous representation for U_n^- is

$$U_n^- = \begin{cases} \langle U_0 + S_n^V \rangle & n \text{ even} \\ \langle 1 - U_0 - S_n^V \rangle & n \text{ odd} \end{cases} \quad (1.6)$$

In practice we usually take the densities of innovation to be step functions, i.e.

$$f_V(t) = \sum_{j=1}^K \frac{P_j}{\alpha_j} I_{[L_j, R_j)}(t) \quad (1.7)$$

with $I_{[L_j, R_j)}$ denoting the indicator function of $[L_j, R_j)$ and $-0.5 \leq L_1 < R_1 \leq \dots \leq L_K < R_K \leq 0.5$, $\alpha_j = R_j - L_j$, $P_j > 0$ and $\sum_{j=1}^K P_j = 1$. For $a, b \in [0, 1)$, define

$$\langle a, b \rangle = \begin{cases} [a, b) & \text{if } a < b \\ [0, 1) - [b, a) & \text{if } b < a \\ \{a\} & \text{if } a = b \end{cases} \quad (1.8)$$

and

$$C(x, L, R) = \langle \langle x + L \rangle, \langle x + R \rangle \rangle \quad (1.9)$$

where $\langle \cdot \rangle$ is the modulo-1 operator. By (3.3) in [3], the one-step transition density of $\{U_n^+\}$ is

$$g^+(y|x) = \frac{\partial}{\partial y} P(U_{n+1}^+ \leq y | U_n^+ = x) = \sum_{j=1}^K \frac{P_j}{\alpha_j} I[y \in C(x, L_j, R_j)] \quad (1.10)$$

where $I[y \in C(x, L_j, R_j)]$ denotes the indicator function of $C(x, L_j, R_j)$.

It is of interest to compare TES processes and their time-reversed counterparts which are also stationary Markovian processes (see [26] and [31]). Let $\{U_n^+\}_{n=0}^{\tau}$ be a TES⁺ subsequence and define its time-reversed version $\{\tilde{U}_n^+\}_{n=0}^{\tau}$ by $\tilde{U}_n^+ = U_{\tau-n}^+$. Now, from (1.5) we get

$$U_{\tau}^+ = \langle U_0 + S_{\tau}^V \rangle \quad (1.11)$$

Applying modulo-1 arithmetic we can isolate U_0 as $U_0 = \langle U_{\tau}^+ - S_{\tau}^V \rangle$, or equivalently in terms of the time-reversed subsequence

$$\tilde{U}_{\tau}^+ = \langle \tilde{U}_0^+ - S_{\tau}^V \rangle \quad (1.12)$$

A comparison of (1.11) and (1.12) reveals that a time-reversed TES⁺ sequence is also a TES⁺ sequence with an innovation sequence $\{\tilde{V}_n\}$, where $\tilde{V}_n = -V_{\tau-n}$. We refer to the $\tilde{V}_n$ as *reverse innovations*. Indeed, the one-step transition density $\tilde{g}^+$ of $\{\tilde{U}_n^+\}$ is related to $g^+(y|x)$ by $\tilde{g}^+(y|x) = g^+(x|y)$ on their common support, since the marginals of $\{U_n^+\}$ and $\{\tilde{U}_n^+\}$ are both uniform on $[0, 1)$.

Background sequences are only an intermediate step in obtaining the transformed *foreground* TES sequences $\{X_n^+\}$ and $\{X_n^-\}$ constructed as

$$X_n^+ = D(U_n^+), \quad X_n^- = D(U_n^-) \quad (1.13)$$

where the function $D(x)$ is called *distortion* function. During the modeling process if X_n^+ or X_n^- are required to have distribution function $F(x)$, the distortion is taken as the inverse function of $F(x)$, i.e. $D(x) = F^{-1}(x)$, because it is well known that if U is a r.v uniformly distributed on $[0,1)$ then $X = F^{-1}(U)$ has distribution function $F(x)$. For further details see [34, 35, 25, 19] and references therein.

TES models have the following advantages: (1) They are designed to fit arbitrary empirical marginal distributions and to simultaneously approximate the empirical autocorrelation up to a specified number of leading lags. (2) The sample paths generated by Monte Carlo simulation of the models can be made to “resemble” the sample paths of empirical data. (3) TES sequences are easily generated on a computer. (4) A software package called *TESool* is available, which provides a visual interactive environment for fitting TES models to empirical data (see [19]).

Although there are some techniques (e.g. Markov Modulated Arrival Process) which can be used to model some time series they cannot capture both marginal distribution and autocorrelation function simultaneously. For example, autoregression can fit a variety of autocorrelation function but not arbitrary marginal distribution. Conversely the minification/maxification approach [32] can fit general marginal distributions, but has limitations in fitting autocorrelation functions.

1.2 Distributions of Waiting Times and Queue Lengths

Of particular importance for the performance analysis of queueing systems, modeling modern computer communication networks is the derivation of the probability distribution of queue sizes or queueing delays. In this thesis we concentrate on the derivation and characterization of the tail of waiting time distributions. In this section we provide a literature survey and we briefly outline the research that was performed in this field.

Consider a stable GI/GI/1 queueing system. Let $\{A_k\}$ denote the independent and identically distributed (i.i.d.) interarrival times and $\{B_k\}$ denote the i.i.d. service times. Let $F_A(t)$ and $F_B(t)$ denote the distribution functions of A_k and B_k respectively, and for $\alpha > 0$ let $A^*(\alpha)$ and $B^*(\alpha)$ denote the Laplace-Stieltjes transforms of $F_A(t)$ and $F_B(t)$ respectively, i.e.

$$A^*(\alpha) = \int_0^\infty e^{-\alpha t} F_A(t) dt$$

and

$$B^*(\alpha) = \int_0^\infty e^{-\alpha t} F_B(t) dt.$$

Let Q_k and W_k denote the queue length and waiting time observed by the k 'th job upon its arrival. Define

$$\Lambda(\eta) = \log(A^*(\eta)) + \log(B^*(-\eta)) \quad (1.14)$$

If there is $\theta > 0$ such that $\Lambda(\theta) = 0$ (which is necessarily unique because of the convexity of $\Lambda(\eta)$), and if $\Lambda'(\theta) < \infty$, Feller in [16] showed that for a stationary queue,

$$P(Q_k \geq n) \sim K_Q \omega^n \quad (1.15)$$

where $\omega = A^*(\theta)$, and for FIFO (first in - first out) queueing discipline

$$P(W_k \geq x) \sim K_W e^{-\theta x} \quad (1.16)$$

where K_Q and K_W are constants and $\sim$ means “asymptotically equal to”.

The classical approach by which one can obtain these results is presented by Feller [16]. This approach is based on the FIFO waiting time relationship

$$W_k = (W_{k-1} + B_{k-1} - A_k)^+ \quad (k \geq 1) \quad (1.17)$$

where $x^+ = \max(x, 0)$. From (1.17) it follows that for $x > 0$,

$$P(W_k \geq x) = P\left(\max_{1 \leq l \leq k} \sum_{i=l}^k (B_{i-1} - A_i) \geq x\right) = P\left(\max_{1 \leq l \leq k} \sum_{i=1}^l (B_{i-1} - A_i) \geq x\right) \quad (1.18)$$

Feller obtained (1.16) by using an exponential change in distribution of form

$$dF_A^\theta(t) = e^{-\theta t} dF_A(t) / A^*(\theta)$$

and

$$dF_B^\theta(t) = e^{-\theta t} dF_B(t) / B^*(\theta)$$

In [39] Sadowsky and Szpankowski generalized the classical approach presented in [16] to GI/GI/m (i.e. independent queue arrival and service processes, m servers) queues. Assume that the queueing system is stable and operating under the stationary distribution. Then

$$P(Q_k \geq n) \sim K_Q \omega^{-n} \quad (1.19)$$

and

$$P(W_k > x) \sim K_W e^{-\theta x} \quad (1.20)$$

where K_Q , K_W , θ and ω are constants. If all m servers have the same service time distribution and if

$$\Lambda_m(\eta) = \log(A^*(m\eta)) + \log(B^*(-\eta)) \quad (1.21)$$

then

$$\theta = \sup\{\eta : \Lambda_m(\eta) \leq 0\}$$

and

$$\omega = \inf_{0 \leq \eta \leq \theta} A^*(\eta).$$

In [41] Takahashi considered the waiting times in PH/PH/ m (i.e. phase-type arrival and service processes, m servers) queues in which all m servers have the same service time distributions. Applying a different approach based on the method of phases he proved that

$$1 - W(x) \simeq K e^{-m\theta x} \quad (1.22)$$

where $W(x)$ is the steady-state waiting time distribution of the PH/PH/ m queue, K is a constant and θ is the solution in η of the equation $A^*(m\eta)B^*(-\eta) = 1$.

In [36] Neuts and Takahashi considered a GI/PH/ m system in which arrival process is independent, service process is of phase-type and there are m servers, and moreover the various servers are allowed to be heterogenous, i.e. the service time distributions may be different for different servers. A result similar to (1.22) was obtained.

In [1] Addie and Zukerman proposed a new traffic model based on a stationary Gaussian process which can be used as a model of superposition of a quite large variety of processes. Consider a FIFO single-server queue with infinite buffer. Let time be divided into fixed length sampling intervals. α_n will denote the amount of work entering the system during the n th sampling interval and β_n will denote the work which can be processed by the server per sampling interval. Assume that the service takes place at the end of the each of the intervals. Let

$$Y_n = \alpha_n - \beta_n \quad n \geq 0$$

Then the system unfinished work process satisfies the recurrence relation:

$$V_{n+1} = (V_n + Y_n)^+ \quad (1.23)$$

Suppose that $\{Y_n\}_{n=-\infty}^{\infty}$ is a stationary, ergodic, Gaussian discrete-time process. They proved that

$$P(V_{\infty} \geq t) e^{-s^* t} \longrightarrow c \quad (t \rightarrow \infty) \quad (1.24)$$

where $s^* < 0$ and c are some constants. Methods for the estimation of s^* and c are given in [1].

The theory of Large Deviations may be used to show results similar to those presented up to now. We consider a general single-server queue and for $t \in T$ ($T = R_+$ or Z_+), denote by A_t the amount of work which arrives to be processed in the interval $[-t, 0)$ and by S_t the amount of work which can be processed in the same interval. The workload process $\mathcal{W}$ is defined by $\mathcal{W}_0 = 0$ and

$$\mathcal{W}_t = A_t - S_t \quad (1.25)$$

and the queue size of the unprocessed work at time zero is

$$Q = \sup_{t \geq 0} \mathcal{W}_t \quad (1.26)$$

Glynn and Whitt [20] showed that for $T = Z_+$ that if the pair $(\mathcal{W}_t/t, t)$ satisfies a large deviation principle, then

$$\lim_{b \rightarrow \infty} b^{-1} \log P(Q > b) = -\delta \quad (1.27)$$

where

$$\delta = \sup\{\theta : \lambda(\theta) \leq 0\} \quad (1.28)$$

and λ is the cumulative generating function of the work load process defined by

$$\lambda(\theta) = \lim_{t \rightarrow \infty} t^{-1} \log E(e^{\theta \mathcal{W}_t}) \quad (1.29)$$

Duffield and O'Connell extended these results in two directions [14]. First the same result is shown to hold when $T = R_+$, subject to a local growth condition on $\mathcal{W}$. Then an analogous result is shown with large deviation scaling more general than b^{-1} , t^{-1} in (1.27) and (1.29). In [5] Botvich and Duffield considered a general workload process $\{\mathcal{W}_t^L\}_{t \in T}$ ($T = Z_+$ or R_+), where $L \in N$ and $\mathcal{W}_0^L = 0$. They defined the cumulate generating function to be

$$\lambda_t^L(\theta) = (Lt)^{-1} \log E(e^{\theta \mathcal{W}_t^L}) \quad (1.30)$$

Under some conditions, one of which is that $(\mathcal{W}_t^L/L, L)$ satisfies a large deviation principle, they proved that

$$\lim_{L \rightarrow \infty} \log P\left(\sup_{t > 0} \mathcal{W}_t^L > Lb\right) = -I(b) \quad (1.31)$$

where $I(b)$ denote a rate function (for details see [5]).

In certain cases finding lower or upper bounds for the distributions of waiting times or queue lengths may be of particular importance. There are certain methods to obtain the bounds, one of which is based on martingale techniques. In his classical note, Kingman [28] used techniques from martingale theory in order to obtain exponential upper bounds for the queue lengths in the GI/GI/1 queue. We briefly present a slightly modified version of Kingman's method. Let messages labeled 1, 2, ... arrive at the queue, and be served according to FIFO (first in first out) discipline. Let $Y_n = B_n - A_n$, where B_n is the service time of the n^{th} message, and A_n is the interarrival time of the n^{th} message. If the queue is initially empty, then the waiting time of the $(n+1)^{\text{th}}$ message is

$$W_{n+1} = \max\{0, Y_n, Y_n + Y_{n-1}, \dots, Y_n + \dots + Y_1\} \quad (1.32)$$

Suppose that $E(Y_n) < 0$ which implies that the queueing system is stable. We can find $y > 1$ such that $E(y^{Y_j}) = 1$. Since Y_n are i.i.d., then for $x \geq 0$ the event $\{W_n \geq x\}$ has the same probability as the event $\{\max_{1 \leq j \leq n} Z_j \geq y^x\}$ where $Z_j = y^{Y_1 + Y_2 + \dots + Y_j}$. We can easily show that

$$E(Z_{j+1} | Y_1, \dots, Y_j) = Z_j.$$

In other word, $\{Z_j\}_{j \geq 1}$ is a martingale. Thus by the Doob-Kolmogorov inequality for positive martingales [43], we have the upper bound

$$P(W_n \geq x) = P(\max_{1 \leq j \leq n} Z_j \geq y^x) \leq y^{-x}$$

The bound is optimized by taking $y = \bar{y} =: \max_{y \geq 1} \{y : E(y^{Y_1}) = 1\}$. The bound is independent of n and so is also a bound for the distribution of $W = \lim_{n \rightarrow \infty} W_n$.

In [8] Buffet and Duffield considered a single-server queue in ATM (asynchronous transfer mode) with L independent sources, each of which is represented by a copy of Markov chain $\{X_t\}_{t \in \mathbb{Z}}$ taking values in the state space $\{0, 1\}$ with transition matrix

$$T = \begin{pmatrix} 1-a & a \\ d & 1-d \end{pmatrix} \quad (1.33)$$

The queue operates as follows. Let X_t^l , $l = 1, 2, \dots, L$, denote the l -th input at time t . Then at each integral time t , all active lines deliver one cell into the buffer of the queue. The queue then serves at most s cells from the buffer, s being a fixed quantity. Define

$$Y_t = \sum_{1 \leq l \leq L} X_{-t}^l.$$

Then under some conditions the steady-state waiting time at time 0 is

$$v = \max \left\{ 0, \sup_{t \geq 1} \{Y_1 + Y_2 + \dots + Y_t - ts\} \right\}$$

Let

$$\alpha(x) = (ax + 1 - a)^L, \quad \beta(x) = \frac{(1 - d)x + d}{ax + 1 - a} \quad (1.34)$$

where a, d are given in 1.33 and let $x, y > 0$. They defined a martingale as follows.

$$\begin{aligned} Z_1 &= x^{Y_1} \\ Z_k &= \alpha(x)^{1-k} y^{Y_1 + Y_2 + \dots + Y_{k-1}} x^{Y_k} \quad \text{for } k > 1 \end{aligned}$$

and $\mathcal{F}_k$ is the σ -algebra generated by the random variables $Y_1, Y_2, \dots, Y_k$. If $1 \leq y \leq x$, $x = y\beta(x)$, and $y^s/\alpha(x) \geq 1$, applying Doob-Kolmogorov inequality they proved that for $b \geq 1$,

$$P(v \geq b) \leq \frac{1}{x^{s+1}y^{b-1}} \left(\frac{ax + d}{a + d} \right)^L \quad (1.35)$$

In [15] Duffield generalized the work of Buffet and Duffield [8]. He considered an ATM multiplexer in which the input is a heterogeneous superposition of Markovian on-off sources and under weaker conditions, by applying martingale techniques he obtained exponential bounds of the form $P(\text{queue size} \geq b) \leq cy^{-b}$ for suitable c and y .

Monte Carlo (MC) simulation is another method to obtain the tail probabilities of waiting times or queue lengths and is particularly useful for the analysis of more complex systems where mathematical techniques may be difficult to apply. However if the probabilities are

very small, say 10^{-9} , then the conventional MC simulation is impractical. A powerful technique called importance sampling can be used to solve this problem in many cases. If one needs to calculate a probability of form $P(W > x) = \int_{-\infty}^{\infty} I(w > x) f_W(w) dw$ importance sampling involves choosing an appropriate probability density $g(\cdot)$ and observing that

$$P(W > x) = E_g \left(I(W > x) \frac{f_W(W)}{g(W)} \right)$$

where the expectation is with respect to density $g(\cdot)$. One then samples repeatedly from the density $g(\cdot)$, and estimates $P(W > x)$ by the sample mean of the observations of $I(W > x) f_W(W)/g(W)$. The term importance sampling is derived from the fact that one can choose $g(\cdot)$ to be large in the regions that are most important, namely where $|f|$ is largest. By properly choosing the density $g(\cdot)$ we can greatly increase the efficiency of MC simulation. Importance sampling is generally used to reduce the variance of estimation. For the details on importance sampling, see [22], [21] and [23]. On fast simulation using importance sampling for estimating the tail probabilities of waiting times or queue lengths, some results can be found in [4], [6], [40], [2], [9], [10], [11], [12], [17], [18] and [37]. In [11] Devetsikiotis and Townsend presented a dynamic importance sampling methodology for fast simulations of regenerative processes.

1.3 Thesis Overview

The TES (Transform-Expand-Sample) technique has emerged as a general methodology for modeling autocorrelated stationary random processes with arbitrary marginal distributions. Although TES models can be easily implemented for simulation studies the need still exists for analytical results on the performance of queueing systems under TES traffic. Of particular interest are the tails of the waiting time distribution in queues driven by TES-modeled bursty traffic. Such tail probabilities, when they become exceedingly small, may be difficult to obtain via conventional simulation.

The main purpose of this thesis is to develop exponential upper bounds for the tail probabilities of the waiting times in a TES/GI/1 queue system, in which the interarrival times are modeled by a TES sequence, while the service times are arbitrarily distributed independent random variables. Our work is motivated by the ability of TES to capture accurately the behavior of bursty traffic in ATM networks. Therefore, bounds derived here can be of great use in predicting the performance of ATM systems (e.g., buffers, multiplexers etc.) under bursty traffic. We solve this problem in two steps.

In the first step we consider a general $G/GI/1$ queue. Let $\{A_k\}_{k=-\infty}^{\infty}$ denote the interarrival times which form a stationary Markov chain and $\{B_k\}_{k=-\infty}^{\infty}$ denote the service times. Let W_n denote the waiting time of the n -th arrival and assume that $W_0 = 0$. Then for $n \geq 1$,

$$W_n = (W_{n-1} + B_{n-1} - A_n)^+ = \max \left\{ 0, \max_{1 \leq k \leq n} \left\{ \sum_{i=k}^n (B_{i-1} - A_i) \right\} \right\} \quad (1.36)$$

For $n \geq 1$, define

$$W_n^* = \max_{1 \leq k \leq n} \left\{ \sum_{i=-1}^{-k} (B_{i-1} - A_i) \right\} \quad (1.37)$$

Obviously $W_n^* \leq W_{n+1}^*$. Let

$$W^* = \lim_{n \rightarrow \infty} W_n^* \quad (1.38)$$

By the stationarity of $\{B_{n-1} - A_n\}$, we have that for $x > 0$,

$$\lim_{n \rightarrow \infty} P(W_n > x) = \lim_{n \rightarrow \infty} P(W_n^* > x) = P(W^* > x)$$

i.e. W_n converges to W^* in distribution. Therefore we are interested in deriving exponential bound for $P(W^* > x)$. Under some weak conditions we obtain that

$$P(W^* > x) \leq Ce^{-\theta^* x} \quad (1.39)$$

where C and θ^* are appropriate constants (see Theorem 2.1.1).

We use martingale techniques similar to those mentioned in Section 1.2 to prove (1.39). First, we use the one-step transition density of time-reversed Markovian chain of $\{A_k\}$, i.e.

$$\tilde{f}(y|x) = \frac{\partial}{\partial y} P(A_n \leq y | A_{n+1} = x)$$

to construct a suitable mapping in a function space and find out its fixed point. Second, we use this fixed point to construct a supermartingale relating to the waiting times, and finally by applying a martingale inequality we obtain our exponential upper bounds. We subsequently apply the general results to TES/GI/1 queues. Furthermore, in order to verify our results we run simulations employing importance sampling techniques and compare our analytical upper bounds with the simulation results.

Finally we consider the stability of queueing systems with TES traffic, an issue of great practical interest. We established a Strong Law of Large Numbers for TES sequence and used it to derive easily verifiable conditions on stability. In particular, we showed that if the squares of the interarrival times and service times are integrable, then the TES/GI/1 or TES/TES/1 queueing system is stable if the mean interarrival time is larger than the mean service time and it is unstable if the mean interarrival time is less than the mean service time.

The rest of our work is organized in two chapters. In Section 1 of Chapter 2, we describe our results concerning bounds on the tail probability of waiting times in the class of $G/GI/1$ queues with Markovian arrivals. In Section 2 of Chapter 2, we apply the general results to TES processes and give numerical examples. To obtain low probabilities a fast simulation technique based on importance sampling is used. We also compare our bounds with the

results from the simulations. In Section 3 of Chapter 2 we provide proofs for the general results. Chapter 3 contains results on the stability of queueing systems with TES traffic. In Section 1 of Chapter 3 we give and prove a Strong Law of Large Numbers for TES sequences and in Section 2 of Chapter 3 we prove theorems on stability of TES/GI/1 and TES/TES/1 queueing systems.

Chapter 2

Exponential Bounds of Waiting Times

This chapter is devoted to methods for exponential bounds for the waiting time in queues. Section 2.1 will present results on the exponential bounds for waiting time distribution in $G/GI/1$. Section 2.2 focuses on application of these results to $TES/GI/1$ queueing systems in which the interarrival times are modeled by a TES sequence. Furthermore we compare these exponential bounds with simulation results from numerical examples. Section 2.3 will provide the proof for the results of Section 2.1.

2.1 Exponential Bounds for the Waiting Time Distribution in $G/GI/1$ Queues

Let A_k and B_k denote the interarrival time and the service time for the k -th arrival respectively in a $G/GI/1$ queue. In this section we assume that $\{A_k\}_{k=-\infty}^{\infty}$ is a stationary Markov chain with continuous state space in $[0, \infty)$. We also assume that $\{A_k\}_{k=-\infty}^{\infty}$ is independent of $\{B_k\}_{k=-\infty}^{\infty}$. Suppose that A_k has density function $f(t)$ with $f(t) = 0$ if $t < 0$, and there exists the one-step transition density function $\tilde{f}(y|x)$ of the time-reversed Markov chain of $\{A_k\}$, i.e.

$$\tilde{f}(y|x) = \frac{\partial}{\partial y} P(A_k \leq y | A_{k+1} = x)$$

Furthermore, we assume that $\{B_k\}_{k=-\infty}^{\infty}$ are i.i.d. random variables with $B_k \geq 0$ and that $M_B(\theta) = E(e^{\theta B_k})$ exists for all θ in $[0, \theta_1]$ ($\theta_1 > 0$). Let W_n denote the waiting time for the n -th arrival. From Section 1.3, W_n converges to W^* in distribution, where

$$W^* = \lim_{n \rightarrow \infty} W_n^*$$

with W_n^* defined as follows:

$$W_n^* = \max_{1 \leq k \leq n} \left\{ \sum_{i=-1}^{-k} (B_{i-1} - A_i) \right\}.$$

We are interested in the tail probability $P(W^* > x)$. An exponential upper bound for $P(W^* > x)$ is given in the following theorem.

Theorem 2.1.1 *If for some positive number θ^* and some positive integer n_0 , the following inequality is true*

$$\eta_0 = \sup_{t \in \mathcal{R}(A_0)} [M_B(\theta^*)]^{n_0} E \left(e^{-\theta^* \sum_{i=-1}^{-n_0} A_i} | A_0 = t \right) < 1$$

where $\mathcal{R}(A_0)$ is the range of A_0 , then for $x > 0$,

$$P(W^* > x) \leq C e^{-\theta^* x}$$

where

$$C = \frac{\sum_{k=1}^{\infty} [M_B(\theta^*)]^k E \left(e^{-\theta^* \sum_{i=1}^k A_i} \right)}{1 + \inf_{t \in \mathcal{R}(A_0)} \left[\sum_{k=1}^{\infty} [M_B(\theta^*)]^k E \left(e^{-\theta^* \sum_{i=-1}^{-k} A_i} | A_0 = t \right) \right]}$$

From this theorem we can easily obtain the following corollaries:

Corollary 2.1.1 *Suppose that $\exists c_1 > 0$ such that $\tilde{f}(y|x) \leq c_1 f(y)$. If for some θ^* and some positive integer n_0 ,*

$$c_1 (M_B(\theta^*))^{n_0} E \left(e^{-\theta^* \sum_{i=1}^{n_0} A_i} \right) < 1$$

then the conclusion of Theorem 2.1.1 holds.

Corollary 2.1.2 *If the conditions of Theorem 2.1.1 (or Corollary 2.1.1) are satisfied and if there is a positive number c_0 such that $c_0 f(y) \leq \tilde{f}(y|x)$, then*

$$P(W^* > x) \leq c_0^{-1} e^{-\theta^* x}$$

The proof of Theorem 2.1.1 and subsequent Corollaries are based on a fixed point theorem for contraction mappings in metric spaces, and a martingale inequality. The proofs are given in Section 2.3.

We will show that our result in Theorem 2.1.1 is a “continuous” generalization of that for GI/GI/1 queueing systems. For this let us first prove the following lemma.

Lemma 2.1.1 *If $E(X) < 0$ for a random variable X and if there exists $\theta_1 > 0$ such that $1 < E(e^{\theta_1 X}) < \infty$, then $h(t) = E(e^{tX})$ is continuous and convex on $[0, \theta_1]$. If $t = \theta_0$ minimizes $h(t)$, then $E(e^{\theta_0 X}) < 1$ and $h(t)$ increases on $[\theta_0, \theta_1]$.*

Proof. Obviously $h'(t) = E(Xe^{tX})$ is continuous and $h'(0) = E(X) < 0$. So there is $t_1 > 0$ such that $h'(t) < 0$ for all $t \in [0, t_1]$, i.e. $h(t)$ decreases on $[0, t_1]$. Hence for $t \in [0, t_1]$, $h(t) < h(0) = 1$. Since $h''(t) = E(X^2 e^{tX}) > 0$, $h(t)$ is convex and $h'(t)$ is strictly increasing. It is easy to see that $h'(t) \rightarrow \infty$ as $t \rightarrow \infty$, and there exist $\theta_0 > 0$ such that $h'(\theta_0) = 0$, and $t = \theta_0$ minimizes $h(t)$, and $h(t)$ increases on $[\theta_0, \theta_1]$ because $h'(t) > 0$ for $t \in [\theta_0, \theta_1]$. We complete the proof.

Now assume that $\{A_k\}_{k=-\infty}^{\infty}$ are i.i.d. and there exists a θ_1 such that $1 < E(e^{\theta_1(B_0 - A_1)}) < \infty$. Then η_0 in Theorem 2.1.1 is equal to $[E(e^{\theta^*(B_0 - A_1)})]^{\eta_0} < 1$ which implies that $E(e^{\theta^*(B_0 - A_1)}) < 1$. By Lemma 2.1.1, there exists a $\theta' > 0$ such that $E(e^{\theta'(B_0 - A_1)}) = 1$ and there exists θ_0 with $\theta' > \theta_0 > 0$ such that for any $\theta \in [\theta_0, \theta']$, $E(e^{\theta(B_0 - A_1)}) < 1$. By Theorem 2.1.1,

$$P(W^* > x) \leq e^{-\theta x}$$

Letting $\theta \rightarrow \theta'$, we obtain

$$P(W^* > x) \leq e^{-\theta' x}$$

which is the optimal upper bound of $P(W^* > x)$ for GI/GI/1 queueing systems (see, for example, [28]). Hence our result is a “continuous” generalization of that for GI/GI/1 queues.

2.2 Application to TES/GI/1 Queues

In this section we apply our bounds to a TES/GI/1 queueing system with service times $\{B_k\}_{k=0}^{\infty}$ and interarrival times $\{A_k\}_{k=0}^{\infty}$. We assume that A_k , $k = 0, 1, \dots$ have common density function $f(t)$ and common distribution function $F(t)$. A_k is defined as follows

$$A_k = F^{-1}(U_k^+) \quad k = 0, 1, 2, \dots \quad (2.1)$$

where $\{U_k^+\}$ is the background TES sequence in which the innovation density $f_V(t)$ is given by

$$f_V(t) = \sum_{j=1}^K \frac{P_j}{\alpha_j} I_{[L_j, R_j)}(t) \quad (2.2)$$

with $-0.5 \leq L_1 < R_1 \leq \dots \leq L_K < R_K \leq 0.5$, $\alpha_j = R_j - L_j$, $P_j > 0$ and $\sum_{j=1}^K P_j = 1$.

Furthermore, we extend $\{U_n^+\}_{n=0}^{\infty}$ to $\{U_n^+\}_{n=-\infty}^{\infty}$, and extend $\{A_k\}_{k=0}^{\infty}$ to $\{A_k\}_{k=-\infty}^{\infty}$, and $\{B_k\}_{k=0}^{\infty}$ to $\{B_k\}_{k=-\infty}^{\infty}$ as follows. Let $\{V_k\}_{k=-\infty}^{\infty}$ be the innovations, i.e, they are i.i.d. random variables with common density $f_V(t)$. Define

$$U_0^+ = U_0 \quad (2.3)$$

and for $k \geq 1$,

$$U_{-k}^+ = \langle U_{-(k-1)} - V_{-(k-1)} \rangle \quad (2.4)$$

where $\langle \cdot \rangle$ is the modulo-1 operator defined in (1.1). The definition of U_k^+ for $k \geq 1$ is the same as before. Then by the properties of the modulo-1 operation, for

$$U_0^+ = \langle U_{-1} + V_0 \rangle \quad (2.5)$$

$$U_{-1}^+ = \langle U_{-2} + V_{-1} \rangle \quad (2.6)$$

$\vdots$

$$U_{-k}^+ = \langle U_{-k-1} + V_{-k} \rangle \quad (2.7)$$

$\vdots$

So for any $n = \dots, -2, -1, 0, 1, 2, \dots$,

$$U_k^+ = \langle U_{k-1} + V_k \rangle \quad (2.8)$$

Lemma 2.2.1 $\{U_n^+\}_{n=-\infty}^\infty$ is a stationary Markov chain. The one-step transition density of $\{U_n^+\}_{n=-\infty}^\infty$ is

$$g^+(y|x) = \sum_{j=1}^K \frac{P_j}{\alpha_j} I[y \in C(x, L_j, R_j)] \quad (2.9)$$

and the one-step transition density of time-reversed Markov chain of $\{U_n^+\}_{n=0}^\infty$ is

$$\tilde{g}^+(y|x) = \sum_{j=1}^K \frac{P_j}{\alpha_j} I[x \in C(y, L_j, R_j)] \quad (2.10)$$

where $I[y \in C(x, L_j, R_j)]$ is the indicator function of $C(x, L_j, R_j)$ and $C(x, L_j, R_j)$ is defined in (1.9).

Proof. It is obvious that $\{U_n^+\}_{n=-\infty}^\infty$ is a stationary Markov chain. (2.9) is (1.10) and (2.10) is also true because that $\tilde{g}^+(y|x) = g^+(x|y)$ (the reader should refer to page 7).

Now for $k = \dots, -2, -1, 0, 1, 2, \dots$, define $A_k = F^{-1}(U_k^+)$. It is easy to extend $\{B_k\}_{k=0}^\infty$ to $\{B_k\}_{k=-\infty}^\infty$ just by adding $\{B_k\}_{k=-\infty}^{-1}$ to $\{B_k\}_{k=0}^\infty$ such that $\{B_k\}_{k=-\infty}^\infty$ are i.i.d.

The following corollary provides an upper bound for $P(W^* > x)$ in the case of TES/GI/1 queues:

Corollary 2.2.1 Let $f_V(t)$ be defined by (2.2) with $-0.5 = L_1 < R_1 = L_2 < R_2 = \dots = L_K < R_K = 0.5$. If for some positive number θ^* and a positive integer n_0 the following inequality is true:

$$\beta(M_B(\theta^*))^{n_0} E\left(e^{-\theta^* \sum_{i=1}^{n_0} A_i}\right) < 1 \quad (2.11)$$

then

$$P(W^* > x) \leq \frac{1}{\alpha} e^{-\theta^* x} \quad (2.12)$$

where $\alpha = \min_{1 \leq j \leq K} \{P_j/\alpha_j\}$ and $\beta = \max_{1 \leq j \leq K} \{P_j/\alpha_j\}$.

Proof. Let $\tilde{f}(y|x)$ be the one-step transition density of $\{A_k\}$. By Lemma 2.2.1,

$$\tilde{f}(y|x) = \frac{\partial}{\partial y} P(A_k \leq y | A_{k+1} = x) \quad (2.13)$$

$$= \frac{\partial}{\partial y} P(U_k \leq F(y) | U_{k+1} = F(x)) \quad (2.14)$$

$$= \tilde{g}^+(F(y)|F(x)) \frac{d}{dy} F(y) \quad (2.15)$$

$$= f(y)g^+(F(x)|F(y)) \quad (2.16)$$

where $f(y)$ is the density function of A_k . Therefore, by (2.10), $\alpha f(y) \leq \tilde{f}(y|x) \leq \beta f(y)$. Corollary 2.2.1 follows directly from Corollary 2.1.2.

To apply Theorem 2.1.1 and the relevant corollaries we need to evaluate parameter θ^* . Let

$$\eta_n \triangleq \sup_{t \in \mathcal{R}(A_0)} [M_B(\theta^*)]^n E \left(e^{-\theta^* \sum_{i=-1}^{-n} A_i} | A_0 = t \right)$$

Assume that $\eta_{n_0} < 1$, then

$$\begin{aligned} \eta_{2n_0} &= \sup_{t \in \mathcal{R}(A_0)} [M_B(\theta^*)]^{2n_0} E \left(E(e^{-\theta^* \sum_{i=-1}^{-2n_0} A_i} | A_{-n_0}) e^{-\theta^* \sum_{i=-1}^{-n_0} A_i} | A_0 = t \right) \\ &\leq \eta_{n_0} < 1 \end{aligned}$$

So there exists $\theta^{**} \geq \theta^*$ such that

$$\sup_{t \in \mathcal{R}(A_0)} [M_B(\theta^{**})]^{2n_0} E \left(e^{-\theta^{**} \sum_{i=-1}^{-2n_0} A_i} | A_0 = t \right) < 1$$

In most practical cases, assuming that θ^* is bounded, by progressively increase n_0 we can get better estimates of θ^* . However we do not have a systematic way to evaluate θ^* and for the purpose of this thesis θ^* was estimated through Monte Carlo simulations

We now provide two examples in which we will verify the validity of our claim.

Example 2.2.1. Let us consider a TES-M/M/1 queue, where TES-M means that A_k is defined in (2.1) and $F(\cdot)$, the interarrival time distribution, is an exponential cumulative

Figure 2.1: Comparison between our upper bound and simulation results for Example 2.2.1.

distribution function with mean $1/\lambda = 1$, i.e., $F(t) = 1 - \exp(-t)$. In this example we will try to derive an exponential bound for the delay of a TES/GI/1 queue in which the interarrival times are TES random variables that are weakly correlated. We know that a TES sequence with innovation density $f_V(t) = I_{[-0.5, 0.5]}(t)$ will result in an uncorrelated random sequence. In order to create a weakly correlated random sequence we “slightly” alter the the above innovation density and we consider a TES sequence with innovation density given by

$$f_V(t) = \frac{0.4}{0.5} I_{[-0.5, 0.0)}(t) + \frac{0.6}{0.5} I_{[0.0, 0.5)}(t).$$

Furthermore, let B_k be exponentially distributed with a density $f_B(t) = \mu \exp(-\mu t)$ for $t \geq 0$ with $\mu = 2$. Since the support of the innovation density covers the whole interval $[-0.5, 0.5]$ Corollary 2.2.1 can be applied. We calculated η_0 for $n_0 = 100$ using Monte Carlo simulation. In this case, using 16,000,000 repetitions, we found $\theta^* \approx 1.02$. The 99% upper confidence limit for η_0 was found to be 0.9697, thus we can say with statistical significance

that η_0 is less than unity for this particular θ^* . The quantity $1/\alpha$ is equal to 1.25. Hence from Corollary 2.2.1 we conclude that $P(W^* > x) \leq 1.25e^{-1.02x}$. A comparison between our upper bound and simulation results for $P(W^* > x)$ is shown in Fig. 2.1.

For the first (highest) simulation point, we used conventional regenerative Monte Carlo (MC) simulation, with regeneration cycles defined as in [3]. We observed 100,000 regeneration cycles to ensure convergence of the estimates. The remaining (lower) five points would have required an impractically large amount of run time using conventional MC simulation. Instead, in these cases we accelerated estimation by applying a technique based on *importance sampling*, in a way similar to [10]. Here, for simplicity, we modified only the average arrival rate to be $\lambda^* = 1.403$ and the average service rate to be $\mu^* = 1.099$. Likelihood ratios were taken as in [3]. Approximately 5,000,000 cycles ($\approx 325,000,000$ arrivals) were observed. The 99% confidence intervals for all simulation results are shown in Fig. 2.1.

Example 2.2.2. We consider a TES-M/M/1 queue with $\lambda = 1$, $\mu = 2$ as in the previous example. However, we are now going to produce a TES arrival stream with stronger correlation than the one in the previous example. In order to do this we consider a “narrow” pulse as the innovation density, namely,

$$f_V(t) = \frac{1}{0.3} I_{[-0.1, 0.2]}(t).$$

We have to apply Theorem 2.1.1 to obtain the bound since the support of the innovation density does not cover the whole interval $[-0.5, 0.5]$. Choosing $n_0 = 100$ and applying the MC simulation technique we obtain $\theta^* \approx 0.68$. Using 16,000,000 repetitions, the 99% upper confidence limit for η_0 was found to be 0.97912.

Obviously, for large N ,

$$C \doteq \frac{\sum_{k=1}^N [M_B(\theta^*)]^k E \left(e^{-\theta^* \sum_{i=1}^k A_i} \right)}{1 + \inf_{t \in \mathcal{R}(A_0)} \left[\sum_{k=1}^N [M_B(\theta^*)]^k E \left(e^{-\theta^* \sum_{i=1}^k A_i} | A_0 = t \right) \right]}$$

For $N = 400$, applying again the Monte Carlo technique (1,000,000 repetitions) we obtain $C \approx 8.0436$. Hence by Theorem 2.1.1, $P(W^* > x) \leq 8.0436e^{-0.68x}$. A comparison between our upper bound and simulation results for $P(W^* > x)$ is shown in Fig. 2.2.

As in Example 2.2.1, for the first three (higher) simulation points, we used conventional regenerative MC simulation. As in the previous example, 500,000 regeneration cycles were enough to ensure convergence of the estimates. Again, the remaining (lower) three points would have required an impractically large amount of run time using conventional MC simulation. We accelerated estimation by applying the importance sampling technique. Likelihood ratios were taken as in [3]. For the cases of $x = 20$ and $x = 25$ we modified only the average service rate to be $\mu^* = 1.48$, and we observed 500,000 regeneration cycles.

For the case of $x = 30$ we modified the average service rate to be $\mu^* = 1.333$ and the innovation density to be:

$$\begin{aligned} f_V^*(t) = & \frac{0.2208}{0.05} I_{[-0.1, -0.05]}(t) + \frac{0.2007}{0.05} I_{[-0.05, 0.0]}(t) \\ & + \frac{0.1956}{0.05} I_{[0.0, 0.05]}(t) + \frac{0.1388}{0.05} I_{[0.05, 0.1]}(t) \\ & + \frac{0.1422}{0.05} I_{[0.1, 0.15]}(t) + \frac{0.1016}{0.05} I_{[0.15, 0.2]}(t) \end{aligned}$$

Approximately 1,000,000 cycles ($\approx 492,000,000$ arrivals) were observed. The 99% confidence intervals for all simulation results are shown in Fig. 2.2.

Figure 2.2: Comparison between our upper bound and simulation results for Example 2.2.2.

2.3 Proofs of Theorem 2.1.1 and its Corollaries

This section is divided into three subsections. In the first subsection we introduce concepts and basic results from martingale theory and fixed point theory needed in our proofs. In the second subsection we prove some preliminary lemmas. In the third subsection we will give the proof of Theorem 2.1.1 and its corollaries.

We apply the idea in [8,15] to prove our results. In [8], the authors consider a FIFO (first-in-first-out) queue in which the arrival process comprises a sum of independent Markov processes on the state space $\{0,1\}$. Furthermore, in [15] they consider an ATM multiplexer in which the input is a superposition of heterogeneous Markovian on-off sources. As we also mentioned in Section 1.2, in [8, 15] the authors appropriately define a martingale which is used in conjunction with a martingale inequality to obtain the exponential upper bounds. For G/GI/1 queuing systems in which the interarrival times form a Markov chain with *continuous* state space like TES, it may be difficult to follow a similar approach. In this case

we appropriately construct a function space and on this space we suitably define a contraction mapping relating to the transition density function of the Markov chain describing the TES arrival process. We then apply a fixed point theorem to get a function which is the fixed point of the contraction mapping. Finally we utilize the fixed point to construct a supermartingale and then apply a martingale inequality to derive an exponential upper bound of the waiting time in the queue.

2.3.1 Preliminaries

1) Conditional Expectation and Martingales. Let $(\Omega, \mathcal{F}, P)$ be a probability space in which Ω is the sample space, $\mathcal{F}$ a σ -algebra of events and P a probability measure.

Let X be a random variable. Then X is $\mathcal{F}$ -measurable, i.e. for any real α the set $\Omega_\alpha \triangleq \{\omega \in \Omega; X(\omega) < \alpha\} \in \mathcal{F}$. We use $\sigma(X)$ to denote the minimum σ -algebra containing all Ω_α . For random variables $X_1, X_2, \dots, X_n$, $\sigma(X_1, X_2, \dots, X_n)$ denote the minimum σ -algebra containing $\sigma(X_1), \sigma(X_2), \dots, \sigma(X_n)$. We often say that the σ -algebra $\sigma(X)$ is generated by X and $\sigma(X_1, X_2, \dots, X_n)$ is generated by $X_1, X_2, \dots, X_n$.

Before we introduce the definition of conditional expectation let us first give the following fundamental theorem (see Page 84 in [43]).

Proposition 2.3.1 *Let X be a random variable with $E(|X|) < \infty$. Let $\mathcal{G}$ be a sub- σ -algebra of $\mathcal{F}$. Then there exists a random variable Y such that*

- (a) Y is $\mathcal{G}$ -measurable,
- (b) $E(|Y|) < \infty$,
- (c) for every set G in $\mathcal{G}$,

$$\int_G Y dP = \int_G X dP, \quad \forall G \in \mathcal{G}$$

Moreover, if $\tilde{Y}$ is another random variable with these properties then $Y = \tilde{Y}$ a.s., that is, $P(Y = \tilde{Y}) = 1$.

A random variable Y with the properties (a), (b) and (c) in Proposition 2.3.1 is called a conditional expectation of X given $\mathcal{G}$, and is denoted by $E(X|\mathcal{G})$, i.e., $Y \triangleq E(X|\mathcal{G})$. Note that we often write $E(X|Z)$ for $E(X|\sigma(Z))$ and $E(X|Z_1, Z_2, \dots)$ for $E(X|\sigma(Z_1, Z_2, \dots))$.

The conditional expectation has the following properties ([43], page 88).

Proposition 2.3.2 *Let X, X_1, X_2 and Z be integrable random variables, a_1 and a_2 constants, and $\mathcal{G}$ a sub- σ -algebra of $\mathcal{F}$. We have*

- (a) $E(E(X|\mathcal{G})) = E(X)$,
- (b) $E(a_1X_1 + a_2X_2|\mathcal{G}) = a_1E(X_1|\mathcal{G}) + a_2E(X_2|\mathcal{G})$,
- (c) if $\varphi : R \rightarrow R$ be a convex function and $E(|\varphi(X)|) < \infty$, then

$$\varphi(E(X|\mathcal{G})) \leq E(\varphi(X)|\mathcal{G}) \quad \text{a.s.}$$

- (d) if Z is $\mathcal{G}$ -measurable and $E(|ZX|) < \infty$, then

$$E(ZX|\mathcal{G}) = ZE(X|\mathcal{G}) \quad \text{a.s.}$$

- (e) if $\{Y_k\}_{k=1}^n$ are random variables independent of X and random variable: $\{Z_k\}_{k=1}^m$, then

$$E(X|Y_1, \dots, Y_n, Z_1, \dots, Z_m) = E(X|Z_1, \dots, Z_m).$$

Let $\{\mathcal{F}_n\}_{n=0}^\infty$ be a sequence of sub- σ -algebra of $\mathcal{F}$. $\{\mathcal{F}_n\}_{n=0}^\infty$ is called a filtration if

$$\mathcal{F}_0 \subseteq \mathcal{F}_1 \subseteq \dots \subseteq \mathcal{F}_n \subseteq \dots$$

A process $\{X_n\}$ is called a **martingale** relative to $\{\mathcal{F}_n\}$ if

- (i) for each n , X_n is $\mathcal{F}_n$ -measurable,
- (ii) $E(|X_n|) < \infty, \forall n$,
- (iii) $E(X_n|\mathcal{F}_{n-1}) = X_{n-1}$, a.s. ($n \geq 1$).

A **supermartingale** (relative to $\{\mathcal{F}_n\}$) is defined similarly, except that (iii) is replaced by

$$E(X_n|\mathcal{F}_{n-1}) \leq X_{n-1} \quad \text{a.s.} \quad (n \geq 1)$$

and a submartingale is defined with (iii) replaced by

$$E(X_n | \mathcal{F}_{n-1}) \geq X_{n-1} \quad \text{a.s.} \quad (n \geq 1).$$

We now recall the following well-known martingale inequality (see Theorem 2.6.10 in [29]):

Proposition 2.3.3 *If $\{X_n\}_{n=1}^{\infty}$ is a positive supermartingale, and $\lambda > 0$, then*

$$\lambda P(\sup_{n \geq 1} X_n \geq \lambda) \leq E(X_1) \quad (2.17)$$

2) Metric Space and Fixed Point Theorem. A metric space is a pair (M, d) , where M is a set and d is a metric on M (or a distance function on M), that is, a function defined on $M \times M = \{(a, b); a, b \in M\}$ such that for all $x, y, z \in M$ we have

(M1) d is real-valued, finite and nonnegative.

(M2) $d(x, y) = 0$ if and only if $x = y$.

(M3) $d(x, y) = d(y, x)$.

(M4) $d(x, y) \leq d(x, z) + d(z, y)$.

It is obvious that if M_1 is a subset of M then (M_1, d) is also a metric space. For the sake of presentation that will follow the following facts may be useful.

Fact 2.3.1 *Let R denote the set of real numbers and $B(R)$ be the set of real-valued and bounded functions defined on R . For $f, g \in B(R)$, define*

$$d(f, g) = \|f - g\| = \sup_{t \in R} |f(t) - g(t)|.$$

Then $(B(R), d)$ is a metric space. In fact, we can see directly that the conditions (M1), (M2) and (M3) are true for this d . Since if $h \in B(R)$

$$|f(t) - g(t)| \leq |f(t) - h(t)| + |h(t) - g(t)| \leq \sup_{t \in R} |f(t) - h(t)| + \sup_{t \in R} |h(t) - g(t)|$$

Hence $d(f, g) \leq d(f, h) + d(h, g)$.

Let (M, d) be the metric space. A sequence $\{x_n\}$ in (M, d) is said to converge or to be convergent if there is an $x \in M$ such that

$$\lim_{n \rightarrow \infty} d(x_n, x) = 0$$

x is called the *limit* of $\{x_n\}$ and we write

$$\lim_{n \rightarrow \infty} x_n = x$$

or, simply

$$x_n \longrightarrow x$$

We say that $\{x_n\}$ converges to x or has limit x .

A sequence $\{x_n\}$ in (M, d) is said to be Cauchy or to be a Cauchy sequence if

$$\lim_{\substack{n, m \geq N \\ N \rightarrow \infty}} d(x_m, x_n) = 0.$$

(M, d) is said to be a *complete metric space* if for each Cauchy sequence $\{x_n\}$ of (M, d) there exists $x \in M$ such that $x_n \longrightarrow x$.

Fact 2.3.2 $(B(R), d)$ in Fact 2.3.1 is a complete metric space. In fact, if $\{f_n\}$ is a Cauchy sequence of $(B(R), d)$, then for any small $\varepsilon > 0$, there exists a subsequence $\{f_{n_k}\}$ of $\{f_n\}$ with $n_1 < n_2 < \dots$ such that

$$d(f_{n_k}, f_{n_{k+1}}) < \varepsilon^k.$$

Since

$$|f_{n_1}| + \sum_{k=1}^{\infty} |f_{n_{k+1}} - f_{n_k}| \leq B + \sum_{k=1}^{\infty} \varepsilon^k = B + \frac{\varepsilon}{1 - \varepsilon}$$

where B is the upper bound of $|f_{n_1}|$, we conclude that

$$f_{n_1} + \sum_{k=1}^{\infty} (f_{n_{k+1}} - f_{n_k}) \triangleq f \in B(R)$$

and we can easily check that $f_n \longrightarrow f$.

Fact 2.3.3 *Define*

$$M = \{f \in B(R); f \text{ is measurable}\}$$

and d as in Fact 2.3.1. Then (M, d) is a complete metric space. Actually if $\{f_n\}$ is a Cauchy sequence of $(M, d) \subseteq (B(R), d)$, then the limit f in Fact 2.3.2 is measurable.

Obviously if

$$M_1 = \{f \in M; 0 \leq f(t) \leq 1\}$$

and d is as in Example 2.3.2, then (M_1, d) is also a complete metric space.

We now introduce certain facts on "fixed points".

A fixed point of a mapping $T : M \rightarrow M$ of a set M into itself is an $x \in M$ such that

$$Tx = x.$$

Let (M, d) be a metric space. A mapping $T : M \rightarrow M$ is called a contraction on M if there exist a positive real number $\alpha < 1$ such that for all $x, y \in M$

$$d(Tx, Ty) \leq \alpha d(x, y) \quad (2.18)$$

Proposition 2.3.4 (Lemma 5.4-3 in [30], page 329) Let $T : \mathcal{X} \rightarrow \mathcal{X}$ be a mapping on a complete metric space $\mathcal{X} = (\mathcal{X}, d)$ and suppose that T^m is a contraction on $\mathcal{X}$ for some positive integer m , i.e., for all $x, y \in \mathcal{X}$

$$d(T^m x, T^m y) \leq \alpha d(x, y)$$

where $0 < \alpha < 1$. Then T has a unique fixed point.

2.3.2 Introductory Lemmas

Let (M, d) and (M_1, d) be as in Fact 2.3.3 of Section 2.3.1. We now introduce the following mappings in M : Let $\tilde{f}(y|x)$ be the one-step transition density of time-reversed Markov chain of $\{A_k\}$ defined in Section 2.1. For $g(x) \in M$, define

$$L_{\theta^*}(g)(x) = M_B(\theta^*) \int_0^\infty e^{-\theta^* y} g(y) \tilde{f}(y|x) dy \quad (2.19)$$

where θ^* is given in Theorem 2.1.1, Let n_0 and η_0 be given by Theorem 2.1.1, and let

$$\tilde{L}_{\theta^*}(g)(x) = \varepsilon_0 + L_{\theta^*}(g)(x) \quad (2.20)$$

where

$$\varepsilon_0 = \frac{1 - \eta_0}{1 + \sum_{k=1}^{n_0-1} [M_B(\theta^*)]^k} \quad (2.21)$$

It is easy to see that L_{θ^*} and $\tilde{L}_{\theta^*}$ are mappings from M to M . Note that the purpose to define ε_0 as in (2.21) is to guarantee that $\tilde{L}_{\theta^*}$ will be a contraction (see the proof of Lemma 2.3.3).

Lemma 2.3.1 For $g \in M$ and any positive integer n ,

$$L_{\theta^*}^n(g)(x) = [M_B(\theta^*)]^n E \left(e^{-\theta^* \sum_{i=1}^n A_i} g(A_n) | A_0 = x \right) \quad (2.22)$$

or

$$L_{\theta^*}^n(g)(x) = [M_B(\theta^*)]^n \underbrace{\int_0^\infty \cdots \int_0^\infty}_n e^{-\theta^* \sum_{i=1}^n t_i} g(t_n) \left[\prod_{i=2}^n \tilde{f}(t_i | t_{i-1}) \right] \tilde{f}(t_1 | x) dt_1 \cdots dt_n \quad (2.23)$$

Proof. We use induction technique to prove (2.23). By (2.19), relation (2.23) holds for $n = 1$. Assume that (2.23) holds for $n = k$. Then we have

$$\begin{aligned} L_{\theta^*}^{k+1}(g)(x) &= L_{\theta^*}^k (L_{\theta^*}(g)(\cdot))(x) \\ &= [M_B(\theta^*)]^k \underbrace{\int_0^\infty \cdots \int_0^\infty}_k e^{-\theta^* \sum_{i=1}^k t_i} L_{\theta^*}(g)(t_k) \left[\prod_{i=2}^k \tilde{f}(t_i | t_{i-1}) \right] \tilde{f}(t_1 | x) dt_1 \cdots dt_k \\ &= [M_B(\theta^*)]^{k+1} \underbrace{\int_0^\infty \cdots \int_0^\infty}_{k+1} e^{-\theta^* \sum_{i=1}^{k+1} t_i} g(t_{k+1}) \left[\prod_{i=2}^{k+1} \tilde{f}(t_i | t_{i-1}) \right] \tilde{f}(t_1 | x) dt_1 \cdots dt_{k+1}. \end{aligned}$$

So the equality (2.23) is valid. Define $\tilde{A}_n = (A_{-1}, A_{-2}, \dots, A_{-n})$ and let $f_{\tilde{A}_n}(t_1, \dots, t_n | x)$ be the conditional density function defined as

$$f_{\tilde{A}_n}(t_1, \dots, t_n | x) = \frac{\partial}{\partial t_1} \cdots \frac{\partial}{\partial t_n} P(A_{-1} \leq t_1, \dots, A_{-n} \leq t_n | A_0 = x) \quad (2.24)$$

By the properties of $\{A_k\}$ we have

$$f_{\tilde{A}_n}(t_1, \dots, t_n | x) = \tilde{f}(t_n | t_{n-1}) \tilde{f}(t_{n-1} | t_{n-2}) \cdots \tilde{f}(t_2 | t_1) \tilde{f}(t_1 | x) \quad (2.25)$$

Furthermore,

$$E \left(e^{-\theta^* \sum_{i=1}^n A_i} g(A_{-n}) | A_0 = x \right) \quad (2.26)$$

$$= \underbrace{\int_0^\infty \cdots \int_0^\infty}_n e^{-\theta^* \sum_{i=1}^n t_i} g(t_n) f_{\tilde{A}_n}(t_1, \dots, t_n | x) dt_1 \cdots dt_n \quad (2.27)$$

$$= \underbrace{\int_0^\infty \cdots \int_0^\infty}_n e^{-\theta^* \sum_{i=1}^n t_i} g(t_n) \left[\prod_{i=2}^n \tilde{f}(t_i | t_{i-1}) \right] \tilde{f}(t_1 | x) dt_1 \cdots dt_n \quad (2.28)$$

Therefore checking with (2.23) we see that (2.22) is also valid.

Lemma 2.3.2 *For any function g in M and a positive integer n*

$$\begin{aligned} \tilde{L}_{\theta^*}^n(g)(x) &= \varepsilon_0 \left[1 + \sum_{k=1}^{n-1} (M_B(\theta^*))^k E \left(e^{-\theta^* \sum_{i=1}^k A_i} | A_0 = x \right) \right] \\ &\quad (M_B(\theta^*))^n E \left(e^{-\theta^* \sum_{i=1}^n A_i} g(A_{-n}) | A_0 = x \right) \end{aligned} \quad (2.29)$$

Proof: By (2.20), we have

$$\tilde{L}_{\theta^*}^n(g)(x) = \varepsilon_0 + L_{\theta^*}(\varepsilon_0)(x) + \cdots + L_{\theta^*}(\varepsilon_0)(x) + L_{\theta^*}^n(g)(x)$$

By Lemma 2.3.1, we obtain (2.29).

Lemma 2.3.3 *Let the conditions of Theorem 2.1.1 be satisfied. Then $\tilde{L}_{\theta^*}^{n_0}$ is a contraction from M_1 to M_1 .*

Proof: For every function g in M_1 , we will show that $\tilde{L}_{\theta^*}^{n_0}(g)$ is also in M_1 . Obviously $\tilde{L}_{\theta^*}^{n_0}(g)(x) \geq 0$. Since $0 \leq g(x) \leq 1$ and by (2.21) and Lemma 2.3.2,

$$\tilde{L}_{\theta^*}^{n_0}(g)(x) \leq \varepsilon_0 \left(1 + \sum_{k=1}^{n_0-1} [M_B(\theta^*)]^k \right) + \eta_0 = 1$$

Therefore, $\tilde{L}_{\theta^*}^{n_0}$ is a mapping from M_1 to M_1 . We now show that $\tilde{L}_{\theta^*}^{n_0}$ is a contraction on M_1 . For $g_1, g_2 \in M_1$, by the assumption of Theorem 2.1.1 and Lemma 2.3.2,

$$\begin{aligned} & d(\tilde{L}_{\theta^*}^{n_0}(g_1), \tilde{L}_{\theta^*}^{n_0}(g_2)) \\ &= \|\tilde{L}_{\theta^*}^{n_0}(g_1) - \tilde{L}_{\theta^*}^{n_0}(g_2)\| \\ &\leq [M_B(\theta^*)]^{n_0} E\left(e^{-\theta^* \sum_{i=-1}^{-n_0} A_i} | A_0 = x\right) \|g_1 - g_2\| \\ &\leq \eta_0 d(g_1, g_2) \end{aligned}$$

where $d(\cdot, \cdot)$ and $\|\cdot\|$ are defined in Example 2.3.1. By the assumption of Theorem 2.1.1, $\eta_0 < 1$, and Lemma 2.3.3 follows.

Lemma 2.3.4 *If the conditions of Theorem 2.1.1 are satisfied, then $\tilde{L}_{\theta^*}$ has the following fixed point*

$$g_0(x) = \varepsilon_0 \left[1 + \sum_{k=1}^{\infty} [M_B(\theta^*)]^k E\left(e^{-\theta^* \sum_{i=-1}^{-k} A_i} | A_0 = x\right) \right] \leq 1$$

and hence

$$K_0 \triangleq \sum_{k=1}^{\infty} [M_B(\theta^*)]^k E\left(e^{-\theta^* \sum_{i=1}^k A_i}\right) \leq \frac{1}{\varepsilon_0} - 1$$

Proof: Let $L = \tilde{L}_{\theta^*}^{n_0}$. By Lemma 2.3.3, L is a contraction on M_1 . By Proposition 2.3.4, $\tilde{L}_{\theta^*}$ have a unique fixed point $\hat{x}$. Since $\varepsilon_0 \in M_1$, we have

$$\hat{x} = \lim_{n \rightarrow \infty} L_{\theta^*}^n(\varepsilon_0).$$

By Lemma 2.3.2,

$$\tilde{L}_{\theta^*}^n(\varepsilon) = \varepsilon_0 \left[1 + \sum_{k=1}^n [M_B(\theta^*)]^k E\left(e^{-\theta^* \sum_{i=-1}^{-k} A_i} | A_0 = x\right) \right]$$

Hence by letting $n \rightarrow \infty$, we obtain $\hat{x} = g_0(x)$. Since

$$E\left(e^{-\theta^* \sum_{i=1}^k A_i}\right) = E\left(e^{-\theta^* \sum_{i=-1}^{-k} A_i}\right) = E\left[E\left(e^{-\theta^* \sum_{i=-1}^{-k} A_i} | A_0\right)\right],$$

we have

$$1 + \sum_{k=1}^{\infty} [M_B(\theta^*)]^k E \left(e^{-\theta^* \sum_{i=1}^k A_i} \right) \quad (2.30)$$

$$= E \left[1 + \sum_{k=1}^{\infty} [M_B(\theta^*)]^k E \left(e^{-\theta^* \sum_{i=1}^k A_i} \mid A_0 \right) \right] \quad (2.31)$$

$$= \varepsilon_0^{-1} E(g_0(A_0)) \leq \frac{1}{\varepsilon_0} \quad (2.32)$$

For $k = 1, 2, \dots$, we define

$$\xi_k = g_0(A_{-k}) e^{\theta^* \sum_{i=-1}^{-k} (B_{i-1} - A_i)}$$

where θ^* is given by Theorem 2.1.1 and $g_0(x)$ is given by Lemma 2.3.4. For $k \geq 1$ define the σ -field $\mathcal{F}_k$ as follows

$$\mathcal{F}_k \triangleq \sigma(B_0, B_{-1}, \dots, B_{-k+1}, A_{-1}, \dots, A_{-k})$$

Lemma 2.3.5 $\{\xi_k, \mathcal{F}_k\}$ is a positive supermartingale.

Proof: By the definition of supermartingale we need to check

$$(A) \quad \mathcal{F}_k \subseteq \mathcal{F}_{k+1} \text{ and } \xi_k \text{ is } \mathcal{F}_k\text{-measurable}$$

$$(B) \quad E|\xi_k| < \infty$$

$$(C) \quad E(\xi_{k+1} | \mathcal{F}_k) \leq \xi_k$$

Conditions (A) and (B) are obvious. We now proceed to prove (C). By Proposition 2.3.2,

$$\begin{aligned} E(\xi_{k+1} | \mathcal{F}_k) &= E \left(g_0(A_{-k-1}) e^{\theta^* \sum_{i=-1}^{-k-1} (B_{i-1} - A_i)} \mid B_0, B_{-1}, \dots, B_{-k+1}, A_{-1}, \dots, A_{-k} \right) \\ &= e^{\theta^* \sum_{i=-1}^{-k} (B_{i-1} - A_i)} E \left(g_0(A_{-k-1}) e^{\theta^* (B_{-k} - A_{-k-1})} \mid A_{-k} \right) \end{aligned}$$

By (2.20) and since g_0 is a fixed point for $\tilde{L}_{\theta^*}$

$$\begin{aligned} &E \left(g_0(A_{-k-1}) e^{\theta^* (B_{-k} - A_{-k-1})} \mid A_{-k} = x \right) \\ &= M_B(\theta^*) \int_0^\infty e^{-\theta^* y} g_0(y) \tilde{f}(y|x) dy \\ &= g_0(x) - \varepsilon_0 \leq g_0(x) \end{aligned}$$

Hence $E(\xi_{k+1} | \mathcal{F}_k) \leq g_0(A_{-k}) e^{\theta^* \sum_{i=-1}^{-k} (B_{i-1} - A_i)} = \xi_k$, and Lemma 2.3.5 follows.

2.3.3 Proof of Theorem 2.1.1

We now proceed with the proof of the main results in Section 2.1, namely Theorem 2.1.1, and Corollaries 2.1.1 and 2.1.2.

Proof of Theorem 2.1.1: For positive integer n , set $W_n^* = \max_{1 \leq k \leq n} \{\sum_{i=-1}^{-k} (B_{i-1} - A_i)\}$, then for $x \geq 0$ and $m \leq n$, $\{W_m^* > x\} \subset \{W_n^* > x\}$. Thus

$$P(W^* > x) = P\left(\bigcup_{n \geq 1} \{W_n^* > x\}\right) = \lim_{n \rightarrow \infty} P(W_n^* > x)$$

Since

$$\begin{aligned} P(W_n^* > x) &= P\left(\max\left\{\sum_{i=-1}^{-k} (B_{i-1} - A_i), k = 1, 2, \dots, n\right\} > x\right) \\ &= P\left(\bigcup_{k=1}^n \left[\sum_{i=-1}^{-k} (B_{i-1} - A_i) > x\right]\right) \end{aligned}$$

and

$$\begin{aligned} &\left[\sum_{i=-1}^{-k} (B_{i-1} - A_i) > x\right] \\ &= \left[g_0(A_{-k})e^{\theta^* \sum_{i=-1}^{-k} (B_{i-1} - A_i)} > g_0(A_{-k})e^{\theta^* x}\right] \\ &\subset \left[g_0(A_{-k})e^{\theta^* \sum_{i=-1}^{-k} (B_{i-1} - A_i)} > (\inf_t g_0(t))e^{\theta^* x}\right] \\ &= \left[\xi_k > (\inf_t g_0(t))e^{\theta^* x}\right] \end{aligned}$$

we have

$$\begin{aligned} P(W_n^* > x) &\leq P\left(\bigcup_{k=1}^n \left[\xi_k > (\inf_t g_0(t))e^{\theta^* x}\right]\right) \\ &= P\left(\max_{1 \leq k \leq n} \xi_k > (\inf_t g_0(t))e^{\theta^* x}\right) \end{aligned}$$

By Lemma 2.3.5 and Proposition 2.3.3,

$$P(W^* > x) = \lim_{n \rightarrow \infty} P(W_n^* > x) \leq \frac{E(\xi_1)}{\inf_t g_0(t)} e^{-\theta^* x} \quad (2.33)$$

Further,

$$\begin{aligned} E(\xi_1) &= E(g_0(A_{-1})e^{\theta^*(B_{-2}-A_{-1})}) \\ &= E[M_B(\theta^*)E(g_0(A_{-1})e^{-\theta^*A_{-1}}|A_0)] \end{aligned}$$

Since $g_0(t)$ is a fixed point of $\tilde{L}_{\theta^*}$ we have

$$\begin{aligned} &M_B(\theta^*)E(g_0(A_{-1})e^{-\theta^*A_{-1}}|A_0 = t) \\ &= M_B(\theta^*) \int_0^\infty e^{-\theta^*y} g_0(y) \tilde{f}(y|t) dy \\ &= L_{\theta^*}(g)(t) \\ &= \tilde{L}_{\theta^*}(g)(t) - \varepsilon_0 \\ &= g_0(t) - \varepsilon_0 \end{aligned}$$

Hence

$$E(\xi_1) = E(g_0(A_0) - \varepsilon_0)$$

Therefore, by the definition of K_0 (see Lemma 2.3.4)

$$E(\xi_1) = \varepsilon_0 K_0 \tag{2.34}$$

Combining (2.33) and (2.34), we immediately obtain the bound of Theorem 2.1.1, namely:

$$P(W^* > x) \leq C e^{-\theta^* x}$$

Proof of Corollary 2.1.1: We only need to show the conditions of Theorem 2.1.1 are satisfied. Assume that the conditions of Corollary 2.1.1 are satisfied. By Lemma 2.3.1, for $t \in \mathcal{R}(A_0)$ we have

$$\begin{aligned}
& [M_B(\theta^*)]^{n_0} E \left(e^{-\theta^* \sum_{i=-1}^{-n_0} A_i} \mid A_0 = t \right) \\
&= [M_B(\theta^*)]^{n_0} \underbrace{\int_0^\infty \cdots \int_0^\infty}_{n_0} e^{-\theta^* \sum_{i=1}^{n_0} t_i} \left[\prod_{i=2}^{n_0} \tilde{f}(t_i \mid t_{i-1}) \right] \tilde{f}(t_1 \mid t) dt_1 \cdots dt_{n_0} \\
&\leq c_1 [M_B(\theta^*)]^{n_0} \underbrace{\int_0^\infty \cdots \int_0^\infty}_{n_0} e^{-\theta^* \sum_{i=1}^{n_0} t_i} \left[\prod_{i=2}^{n_0} \tilde{f}(t_i \mid t_{i-1}) \right] \tilde{f}(t_1) dt_1 \cdots dt_{n_0} \\
&= c_1 [M_B(\theta^*)]^{n_0} E \left(e^{-\theta^* \sum_{i=-1}^{-n_0} A_i} \right) \\
&= c_1 [M_B(\theta^*)]^{n_0} E \left(e^{-\theta^* \sum_{i=1}^{n_0} A_i} \right) < 1
\end{aligned}$$

Hence η_0 in Theorem 2.1.1 is less than 1, and Corollary 2.1.1 follows.

Proof of Corollary 2.1.2: To prove Corollary 2.1.2 we only need to show that C in Theorem 2.1.1 is less than c_0^{-1} . Using the assumptions of Corollary 2.1.2, similar to the argument in the proof of Corollary 2.1.1, we have that for $t \in \mathcal{R}(A_0)$

$$\begin{aligned}
& [M_B(\theta^*)]^{n_0} E \left(e^{-\theta^* \sum_{i=-1}^{-n_0} A_i} \mid A_0 = t \right) \\
&\geq c_0 [M_B(\theta^*)]^{n_0} E \left(e^{-\theta^* \sum_{i=1}^{n_0} A_i} \right)
\end{aligned}$$

and Corollary 2.1.2 follows immediately.

Chapter 3

Strong Law of Large Numbers of TES and Stability of Queueing Systems with TES Traffic

Our main purpose in this chapter is to derive easily checked conditions for stability of queueing systems with TES traffic. This chapter consists of two sections. In the first section we will prove a Strong Law of Large Numbers (SLLN) for TES sequence. In the second section we use this SLLN to prove theorems on stability of TES/GI/1 and TES/TES/1 queueing systems.

3.1 Strong Law of Large Numbers for TES Processes

In this section we prove theorems on the Strong Law of Large Numbers for TES. For general stationary processes we have the following theorem (see [13] pp. 465 and 469).

Theorem 3.1.1 *Let $\{X_n, n \geq 0\}$ be a stationary stochastic process in the strict sense, with $E(|X_0|) < \infty$. Then with probability 1,*

$$\lim_{n \rightarrow \infty} \frac{X_0 + X_1 + \cdots + X_n}{n+1} = X \quad (3.1)$$

and

$$\lim_{n \rightarrow \infty} E \left| \frac{X_0 + X_1 + \cdots + X_n}{n+1} - X \right| = 0 \quad (3.2)$$

where X is a random variable and $E(X) = E(X_0)$.

The Strong Law of Large Numbers expressed by (3.1) may not be suitable in practice because the limit is a random variable. For example, to discuss the stability of the TES/GI/1 queueing systems one hopes that the limit is a constant. In order to obtain this constant limit we need to add another condition which may not be easy to check, that is, that the process is “metrically transitive” or “ergodic” (see [13] and [42] for the definition).

Let $\{A_k\}$ be a TES sequence with common distribution function $F(x)$, i.e.

$$A_k = F^{-1}(U_k^+) \quad k \geq 0 \quad (3.3)$$

where $\{U_k^+\}$ is a background TES sequence with innovation density f_V (see Section 1.1). Let $D_F(x) = F^{-1}(x)$ and $R(q) = \text{Cov}(A_0, A_q)$ be the autocovariance of $\{A_k\}$ and let $\tilde{D}_F(s)$ and $\tilde{f}_V(s)$ be the Laplace transforms of D_F and f_V respectively, i.e.

$$\tilde{D}_F(s) = \int_0^1 e^{-sx} D_F(x) dx \quad (3.4)$$

and

$$\tilde{f}_V(s) = \int_{-\infty}^{\infty} e^{-sx} f_V(x) dx. \quad (3.5)$$

We have the following theorem.

Theorem 3.1.2 *Let $\{A_k\}$ be the TES sequence defined above. Suppose that $E(A_0^2) < \infty$. Then*

$$\lim_{n \rightarrow \infty} \frac{A_1 + A_2 + \cdots + A_n}{n} = E(A_1) \quad (3.6)$$

We can see that this result is quite strong since for the TES sequence $\{A_k\}$ the sample average converges to a constant, and the only condition needed is that $E(A_0^2) < \infty$. We will assume this condition throughout this section.

Before proving this Theorem 3.1.2 we introduce the following lemmas.

Lemma 3.1.1 *Let f be a complex-valued function defined on R . Suppose that f is integrable on R and*

$$\left| \int_R f(x) dx \right| = \int_R |f(x)| dx.$$

Then there exists a constant α such that for all $x \in R$, $\alpha f(x) = |f(x)|$.

Proof. See 1.39 Theorem in [38].

Lemma 3.1.2 *For any integer v ,*

$$|\tilde{f}_v(i2\pi v)| < 1.$$

Proof. We have

$$|\tilde{f}_v(i2\pi v)| = \left| \int_{-\infty}^{\infty} e^{-i2\pi vt} f_v(t) dt \right| \leq \int_{-\infty}^{\infty} |e^{-i2\pi vt} f_v(t)| dt = \int_{-\infty}^{\infty} f_v(t) dt = 1.$$

If $|\tilde{f}_v(i2\pi v)| = 1$, then,

$$\left| \int_{-\infty}^{\infty} e^{-i2\pi vt} f_v(t) dt \right| = \int_{-\infty}^{\infty} |e^{-i2\pi vt} f_v(t)| dt.$$

By Lemma 3.1.1 there exists a constant α such that

$$\alpha e^{-i2\pi vt} f_v(t) = |e^{-i2\pi vt} f_v(t)| = f_v(t).$$

Now we obtain that

$$\alpha = e^{i2\pi vt} \quad \text{for all } t \text{ with } f_v(t) \neq 0$$

which contradicts the fact that α is a constant.

Jagerman and Melamed proved the following result (Theorem 3 in Part I of [25]).

Lemma 3.1.3 *Under the conditions of Theorem 3.1.2 and for $q = 1, 2, \dots$*

$$R(q) = \sum_{\substack{v = -\infty \\ v \neq 0}}^{\infty} \left(\tilde{f}_v(i2\pi v) \right)^q |\tilde{D}_F(i2\pi v)|^2 \quad (3.7)$$

Using Lemma 3.1.3 we can prove

Lemma 3.1.4 *Let $R(q)$ be defined above. Under the conditions of Theorem 3.1.2,*

$$\lim_{q \rightarrow \infty} R(q) = 0 \quad (3.8)$$

Proof. Obviously

$$|\tilde{D}_F(i2\pi v)| \leq \int_0^1 |e^{-i2\pi vx} D_F(x)| dx \leq \int_0^1 |D_F(x)| dx = \int |t| dF(t) = E|A_0| < \infty. \quad (3.9)$$

By (3.7), for any $\varepsilon > 0$, there exist positive integer N such that

$$\left| \sum_{|v| > N} (\tilde{f}_V(i2\pi v))^q |\tilde{D}_F(i2\pi v)|^2 \right| < \varepsilon \quad (3.10)$$

By Lemma 3.1.3 and (3.10),

$$\limsup_{q \rightarrow \infty} |R(q)| \leq \limsup_{q \rightarrow \infty} \left| \sum_{|v| \leq N} (\tilde{f}_V(i2\pi v))^q |\tilde{D}_F(i2\pi v)|^2 \right| + \varepsilon \quad (3.11)$$

Since by Lemma 3.1.2 for any $v \neq 0$, $|\tilde{f}_V(i2\pi v)| < 1$,

$$\limsup_{q \rightarrow \infty} |R(q)| \leq \varepsilon \quad (3.12)$$

Hence $\limsup_{q \rightarrow \infty} |R(q)| = 0$.

Lemma 3.1.5 *Under the conditions of Theorem 3.1.2,*

$$\lim_{n \rightarrow \infty} E \left(\left| \frac{A_1 + A_2 + \cdots + A_n}{n} - E(A_0) \right|^2 \right) = 0 \quad (3.13)$$

Proof. It is easy to see that

$$E \left(\left| \frac{A_1 + A_2 + \cdots + A_n}{n} - E(A_0) \right|^2 \right) \quad (3.14)$$

$$= n^{-2} \left\{ n \text{Var}(A_0) + 2 \sum_{k=1}^{n-1} \sum_{j=k+1}^n \text{Cov}(A_k, A_j) \right\} \quad (3.15)$$

$$= n^{-2} \left\{ n \text{Var}(A_0) + 2 \sum_{k=1}^{n-1} \sum_{j=1}^{n-k} R(j) \right\} \quad (3.16)$$

$$= n^{-2} \left\{ n \text{Var}(A_0) + 2 \sum_{k=1}^{n-1} (n-k) R(k) \right\} \quad (3.17)$$

Furthermore by assumption $\text{Var}(A_0) < \infty$. Therefore,

$$\lim_{n \rightarrow \infty} n^{-2}(n \text{Var}(A_0)) = 0 \quad (3.18)$$

and by Lemma 3.1.4

$$\left| n^{-2} \sum_{k=1}^{n-1} (n-k) R(k) \right| = \left| n^{-1} \sum_{k=1}^{n-1} \frac{n-k}{n} R(k) \right| \leq n^{-1} \sum_{k=1}^{n-1} |R(k)| \rightarrow 0. \quad (3.19)$$

Proof of Theorem 3.1.2. Since TES processes are strictly stationary, by Theorem 3.1.1 there exist a random variable A such that

$$\lim_{n \rightarrow \infty} \frac{A_1 + A_2 + \cdots + A_n}{n} = A \quad (3.20)$$

and

$$\lim_{n \rightarrow \infty} E \left| \frac{A_1 + A_2 + \cdots + A_n}{n} - A \right| = 0. \quad (3.21)$$

On the other hand, by Lemma 3.1.5

$$\lim_{n \rightarrow \infty} E \left| \frac{A_1 + A_2 + \cdots + A_n}{n} - E(A_0) \right| = 0. \quad (3.22)$$

Hence $A = E(A_0)$ with probability 1 because by (3.21) and (3.22) we have

$$E|A - E(A_0)| \leq E \left| \frac{A_1 + A_2 + \cdots + A_n}{n} - A \right| + E \left| \frac{A_1 + A_2 + \cdots + A_n}{n} - E(A_0) \right| \rightarrow 0 \quad (3.23)$$

So in (3.20) A can be replaced by $E(A_0)$.

3.2 Stability Results of Queueing Systems with TES Traffic

The stability of queueing systems is an important issue since if a queueing system is not stable then this system is not very useful in practice. In this section we only consider queueing systems with FIFO (first in - first out) discipline. It is well known that for GI/GI/1 queueing systems if the mean service time is less than the mean interarrival time then this queueing system is stable. For G/G/1 queueing systems let A_n and B_n be the interarrival time and the service time for the n -th arrival respectively. Put

$$Y_n = B_{n-1} - A_n \quad n = \dots, -2, -1, 0, 1, 2, \dots \quad (3.24)$$

Loynes [33] considered the case that $\{Y_n\}$ is strictly stationary and metrically transitive. Then the queueing system is stable if $E(Y_1) < 0$, is unstable if $E(Y_1) > 0$ and if $E(Y_1) = 0$ then it may be stable or unstable.

To prove these results Loynes [33] first proved the following theorem

Theorem 3.2.1 *A queue is stable if $\limsup_{n \rightarrow \infty} \sum_1^n Y_{-k} < \infty$, with probability one, and unstable otherwise.*

Note that the transitivity condition is not necessary for this theorem. Using the Strong Law of Large Numbers for strictly stationary and metrically transitive processes (see [13], pp.465) we have

$$\lim_{n \rightarrow \infty} n^{-1} \sum_1^n Y_{-k} = E(Y_1) \quad (3.25)$$

Then $\limsup_{n \rightarrow \infty} \sum_1^n Y_{-k} < \infty$ if $E(Y_1) < 0$ and $\limsup_{n \rightarrow \infty} \sum_1^n Y_{-k} = \infty$ if $E(Y_1) > 0$, and therefore by Theorem 3.2.1 the queueing system is stable if $E(Y_1) < 0$ and it is unstable if $E(Y_1) > 0$.

Applying the Loynes' results we need to check a condition that $\{Y_n\}$ is metrically transitive, which is usually difficult to verify. Loynes [33] also stated that the metrically transitivity

may not be necessary. Instead we can drop this condition and replace $E(Y_1) < 0$ or $E(Y_1) > 0$ by $E(Y_1|\mathcal{F}) < 0$ a.s. or $E(Y_1|\mathcal{F}) > 0$ a.s., where $\mathcal{F}$ is an appropriate " σ -field of invariant events". Unfortunately these conditions may not be easy to verify in practical cases.

In [42], Theorem 7.4.5 states that if $\{Y_n\}$ is stationary and ergodic and if $E(Y_1) < 0$, then the queueing system is stable. In order to use this theorem verification of ergodicity condition may be difficult.

In this section we are going to discuss the stability of queueing systems with TES traffic and will provide some easily verifiable stability conditions.

According to [33], a sequence of almost everywhere (a.e) finite random variables X_n with (honest) distribution function F_n will said to be

- (1) stable, if F_n tends to some (honest) distribution function F at all its points of continuity;
- (2) substable, if each subsequence contains a stable sub-sub-sequence;
- (3) unstable, if it is not substable.

A random variable and its distribution function are said to be honest if the random variable is a.s. finite. A queueing system is said to be stable if the waiting times process $\{W_n\}$ is stable, where W_n denotes the waiting time of the n -th arrival.

Let $\{A_k\}$, $\{B_k\}$ and $\tilde{f}_V(s)$ be defined as in Section 3.1. We have

Theorem 3.2.2 *Suppose that $E(A_1^2) < \infty$ and $E(|B_0|) < \infty$. Then a TES/GI/1 queue is stable if $E(Y_1) < 0$ and is unstable if $E(Y_1) > 0$.*

Proof. By the assumption and the Law of Large Numbers for i.i.d. random variables,

$$\lim_{n \rightarrow \infty} n^{-1} \sum_{k=1}^n B_k = E(B_1) \quad (3.26)$$

By (3.26) and Theorem 3.1.2 we have

$$\lim_{n \rightarrow \infty} n^{-1} \sum_{k=1}^n Y_k = E(Y_1).$$

Then applying Theorem 3.2.1 we reach our conclusion.

From Theorem 3.2.2 we can very easily determine whether a TES/GI/1 queueing system is stable or not if $E(Y_1) \neq 0$. The conditions we need to check for the stability are $E(|B_0|) < \infty$ and $E(A_1^2) < \infty$, and obviously checking of these conditions are straightforward.

Now consider TES/TES/1 queueing system, where the interarrival and the service times are TES sequences. From the proof of Theorem 3.2.2 we can see a further generalization for the TES/TES/1 queueing system.

Theorem 3.2.3 *Let $\{A_k\}$ be the interarrival times and $\{B_k\}$ the service times in TES/TES/1 queueing system. Suppose that $\{A_k\}$ is independent of $\{B_k\}$ and $E(B_0^2) < \infty$ and $E(A_1^2) < \infty$. Then the queueing system is stable if $E(B_0) < E(A_1)$ and is unstable if $E(B_0) > E(A_1)$*

Conclusions

The TES (Transform-Expand-Sample) technique has emerged as a general methodology for modeling autocorrelated random processes with arbitrary marginal distributions. Because of their generality and practical applicability, TES models can be readily used to accurately characterize bursty traffic streams in ATM networks (e.g., variable-bit-rate compressed video).

This thesis studies queueing systems with TES traffic. We derived exponential upper bounds for the waiting times in TES/GI/1 queueing systems and found easily-checked conditions for stability of queueing systems with TES traffic. To derive these bounds, we applied techniques from martingale theory and functional analysis. The difficulty in using the martingale technique is the construction of a martingale from interarrival times and service times of a TES/GI/1 queueing system. We overcome this difficulty by successfully using a fixed point theorem from the theory of functional analysis. The comparison of the exponential upper bounds and the results of simulations shows that our upper bounds are tight. The exponential bounds may be useful in the analysis and design of communication networks for deciding on the buffer size of ATM multiplexers. Finally we considered stability of queueing systems with TES traffic. We derived a strong law of large numbers for TES sequences and by using this result we established conditions for the stability of TES/GI/1 and TES/TES/1 queueing systems that can be easily checked.

In terms of further research we are currently investigating techniques that can be used for fast computation of the constants C and θ^* in Theorem 2.1.1. One approach we may follow is based on numerical techniques and the theory of integral operators relating to TES processes.

Let $\{U_k^+\}_{k=0}^\infty$ be the *background* TES sequence and $g^+(y|x)$ be the one-step transition density function. In TES/GI/1 queueing system let the interarrival times $\{A_k\}_{k=0}^\infty$ be defined by (2.1) with distribution function $F(x)$ and let the service times $\{B_k\}_{k=0}^\infty$ be i.i.d. random variables. Let $D(x) = F^{-1}(x)$ and $M_B(\theta) = E(e^{\theta B_k})$. Jagerman [24] suggests that we might define an integral operator as follows:

$$T\varphi(u) = \int_0^1 e^{-\theta D(v)} g^+(v|u) dv$$

and if we could find the eigenvalue $\lambda(\theta)$ of above integral operator then θ^* in Corollary 2.2.1 could be determined by the following equality

$$M_B(\theta)\lambda(\theta) = 1.$$

Another approach that we may follow is based on quick computation of the quantity

$$E\left(e^{-\theta \sum_{i=1}^k A_i} \mid A_0 = x\right)$$

for $x > 0$, $\theta > 0$ and $k > 0$. More specifically we concentrate on the use of fast simulation based on importance sampling for the estimation of the above mentioned expectation.

Finally the exponential bounding techniques we established might be generalized to TES/TES/1 and TES/GI/m systems. The research is under progress in all such directions.

Bibliography

- [1] R.G. Addie and M. Zukerman, A Gaussian traffic model for a B-ISDN statistical multiplexer, 1513-1517, 1992 IEEE.
- [2] Al-Qaq, M. Devetsikiotis and K.R. Townsend, Importance sampling methodologies for simulation of communication systems with adaptive equalizers and time varying channels, IEEE Journal on Selected Areas in Communications 11: 317-327, 1993.
- [3] S. Asmussen & B. Melamed, Regeneration and likelihood ratios in TES processes, preprint.
- [4] S. Asmussen, Conjugate processes and simulation of ruin problems, Stoch. processes and their Appl., 20: 213-229, 1985.
- [5] D.D. Botvich and N. G. Duffield, Large deviations, the shape of the loss curve, and economics of scale in large multiplexers, May 13, 1994, preprint.
- [6] J.A. Bucklew, Large deviation techniques in decision, simulation and estimation, Wiley, New York, 1990.
- [7] J.A. Bucklew, P. Ney and J.S. Sadowsky, Monte Carlo simulation and large deviation theory for uniformly recurrent Markov chains, J Appl. Prob. 27: 44-59, 1990.
- [8] E. Buffet and N. G. Duffield, Exponential upper bounds via martingales for multiplexers with Markovian arrivals, J. Appl. Prob., 1994, to appear.

- [9] C.S. Chang, P. Heidelberger, S. Juneja, and P. Shahabuddin, The application of effective bandwidth to fast simulation of communication networks, IBM Research Report RC 18877, Yorktown Heights, New York, 1993.
- [10] M. Devetsikiotis and K.R. Townsend, On the efficient simulation of large communication networks using importance sampling, In proc. of IEEE Globecom '92, Computer Society Press, 1992.
- [11] M. Devetsikiotis and K.R. Townsend, A dynamic importance sampling methodology for the efficient estimation of rare event probabilities in regenerative simulation of queueing systems, In proc. of IEEE ICC '92, Conference, 1290-1297, Computer Society Press, 1992.
- [12] M. Devetsikiotis and K.R. Townsend, Statistical optimization of dynamic importance sampling parameters for efficient simulation of communication networks, IEEE/ACM Transactions on Networking, 1(3), June 1993.
- [13] J.L. Doob, Stochastic processes, John Wiley & Sons Inc., New York, 1953.
- [14] N.G. Duffield and N. O'Connell, Large deviations and overflow probabilities for the general single-server queue, with application, 1993, preprint.
- [15] N.G. Duffield, Rigorous bounds for queue lengths in heterogeneous ATM multiplexers, preprint.
- [16] W. Feller, An introduction to probability theory and its applications, Vol. II, John Wiley & Sons, New York, 1971.
- [17] M.R. Frater and B.D. Anderson, Fast estimation of the statistics of excessive backlogs in tandem networks of queue, Australian Telecommunications Research 23:49-55, 1989.

- [18] M.R. Frater, J. Walrand and B.D.O Anderson, Optimally efficient simulation of buffer overflows in queues with deterministic service times via importance sampling, Australian Telecommunications Research 24: 1-8, 1990.
- [19] D.Geist & B. Melamed, *TEStool: An environment for visual interactive modeling auto-correlated traffic*, Proceedings of ICC'92, Chicago, Illinois (to appear).
- [20] P.W. Glynn and W. Whitt, Logarithmic asymptotics for steady-state tail probabilities in a single-server queue, J. Appl. Prob., to appear (1994).
- [21] P.W. Glynn and D.L. Iglehart, Importance sampling for stochastic simulation, Management Science 35: 1367-1392, 1989.
- [22] J.M. Hammersley and D.C. Handscomb, Monte Carlo methods, London: Methuen and Co. Ltd, 1964.
- [23] P. Heidelberger, Fast simulation of rare events in queueing and reliability models, IBM research report, 1993.
- [24] D. Jagerman, TES processes, personal communication, 1994.
- [25] D. L. Jagerman & B. Melamed, *The transition and autocorrelation structure of TES processes I+II* Commun. Statist.- Stochastic Models, 8(2) 193-219(1992) and 8(3) 499-527(1992).
- [26] F.P. Kelly, Reversibility and Stochastic networks, Wiley, 1979.
- [27] J.F.C. Kingman, On the algebra of queues, J. Appl. Prob., 3: 285-326, 1966.
- [28] J.F.C. Kingman, A martingale inequality in the theory of queues, Proc. Camb. Phil. Soc. 59: 359-361, 1964.

- [29] P. E. Kopp, *Martingales and Stochastic integrals*, Cambridge University Press, Cambridge, 1984.
- [30] Erwin Kreyszig, *Introductory functional analysis with applications* John Wiley & Son, 1978.
- [31] A.J. Lawrance, Directionality in time series, *International Statistical Review* Vol. 59, 67-79, 1991.
- [32] P. A. W. Lewis and E. McKenize, Minification processes and their transformations, *J. Appl. Prob.*, Vol. 28, 45 - 57, 1991.
- [33] R.M. Loynes, The stability of a queue with non-independent inter-arrival and service times, *Proc. Camb. Phil. Soc.* Vol. 58, 497-520, 1962.
- [34] B. Melamed, *TES: A class of methods for generating autocorrelated uniform variates*, *ORSA J. on Computing*, Vol 3, 317-329, 1991.
- [35] B. Melamed & B. Sengupta, *TES modeling of video traffic*, *IEEE Trans. Commun.* Vol. E75-B, No. 12, 1292-1300, 1992.
- [36] M. Neuts and Y. Takahashi, Asymptotic behavior of the stationary distribution in a GI/PH/c queue with heterogeneous servers, *Z. Wahr.*, 57: 441-452, 1981.
- [37] S. Parekh and J. Walrand, A quick simulation method for excessive backlogs in networks of queues, *IEEE Transaction on Automatic Control* 34: 54-56, 1989.
- [38] W. Rudin, *Real and Complex Analysis*, New York, McGraw-Hill, 1974.
- [39] J.S. Sadowsky and W. Szpankowski, On the analysis of the tail queue length and waiting time distributions of a GI/GI/c queue, *Performance '90*, P.J.B. King, I. Mitrani and R.j. Pooky, Elsevier Science Publisher B.V. (North-Holland) 93-107, 1990.

- [40] J.S. Sadowsky, Large deviations theory and efficient simulation of excessive backlogs in a GI/GI/m queue, IEEE Transactions on Automatic Control 36: 1383-1394, 1991.
- [41] Y. Takahashi, Asymptotic exponentially of the tail of the waiting times distribution in a PH/PH/c queue, Adv. Appl. Prob., 13: 619-630, 1981.
- [42] J. Walrand, An introduction to queueing networks, Prentice Hall, 1988.
- [43] D. Williams, Probability with martingales, CUP, Cambridge, 1991.

END

1 7 - 1 / 1 - 9 5

FIN