Helsinki University of Technology, Institute of Mathematics, Research Reports
Teknillisen korkeakoulun matematiikan laitoksen tutkimusraporttisarja
Espoo 2005 A491

COMPARISON AND SCALING METHODS FOR PERFORMANCE ANALYSIS OF STOCHASTIC NETWORKS

Lasse Leskelä

TEKNILLINEN KORKEAKOULU
TEKNISKA HÖGSKOLAN
HELSINKI UNIVERSITY OF TECHNOLOGY
TECHNISCHE UNIVERSITÄT HELSINKI
UNIVERSITE DE TECHNOLOGIE D'HELSINKI

Helsinki University of Technology, Institute of Mathematics, Research Reports
Teknillisen korkeakoulun matematiikan laitoksen tutkimusraporttisarja
Espoo 2005 A491

COMPARISON AND SCALING METHODS FOR PERFORMANCE ANALYSIS OF STOCHASTIC NETWORKS

Lasse Leskelä

Dissertation for the degree of Doctor of Science in Technology to be presented with due permission of the Department of Engineering Physics and Mathematics for public examination and debate in Auditorium E at Helsinki University of Technology (Espoo, Finland) on 2 December 2005 at 12 noon.

Helsinki University of Technology
Department of Engineering Physics and Mathematics
Institute of Mathematics

Lasse Leskelä: *Comparison and scaling methods for performance analysis of stochastic networks*; Helsinki University of Technology, Institute of Mathematics, Research Reports A491 (2005).

Abstract: Stochastic networks are mathematical models for traffic flows in networks with uncertainty. The goal of this thesis is to develop new methods for analyzing performance and stability of stochastic networks, helping to better understand and control uncertainty in complex distributed systems.

The thesis considers three instances of stochastic networks, each representing a specific challenge for analytical modeling. The first case studies the impact of incomplete information to a queueing network with distributed admission control. Stability conditions for various admission policies are derived, together with a numerical algorithm for performance evaluation. In the second case, stochastic comparison is used to derive performance bounds for multiclass loss networks with overflow routing. The third model is a spatial random field generated by a large number of noninteracting sources, for which scaling and renormalization are used to show how the level of randomness of the individual sources may critically affect the macroscopic statistical properties of the field.

The results of the thesis illustrate the feasibility of stochastic comparison and stochastic analysis in deriving approximations and performance bounds for complex physical networks with uncertainty. Approximations and performance bounds based on exact mathematical methods have the advantage that they explicitly state the type of circumstances required for the accuracy of the estimates. The resulting analytical formulas can sometimes reveal interesting properties that are not easily detected using numerical simulation.

AMS subject classifications: 60K25, 68M20, 90B15, 90B22, 60E50, 60F17, 60G60, 60G18

Keywords: stochastic network, queueing, admission control, overflow routing, stochastic comparison, scaling, renormalization, spatial random field

Correspondence
lasse.leskela@iki.fi

ISBN 951-22-7917-7
ISSN 0784-3143

Helsinki University of Technology
Department of Engineering Physics and Mathematics
Institute of Mathematics
P.O. Box 1100, FI-02015 TKK, Finland
email:math@tkk.fi http://math.tkk.fi/

To Hannele

Acknowledgments

I am grateful above all to Professor Ilkka Norros, my instructor, who has guided me into studying many interesting problems in the field of applied probability. My sincere thanks also go to Professor Esko Valkeila for supporting me in taking the last big steps on the way towards becoming a doctor, and Professor Olavi Nevanlinna for providing me an excellent research environment at the TKK Institute of Mathematics. I thank Dr Takis Konstantopoulos for agreeing to be my opponent in the defense, and Professor Göran Högnäs and Professor Michel Mandjes for reviewing the thesis manuscript.

Large parts of the work presented in this dissertation were developed during the fall 2004 while I was visiting the Mittag-Leffler Institute in Djursholm, Sweden. I am grateful to all visitors at the institute with whom I had the privilege to share and discuss new mathematical ideas. Especially, I thank Professor Ingemar Kaj, Professor Jacques Resing, Professor Volker Schmidt, and Mr Matthieu Jonckheere, who appear as coauthors on the articles in this dissertation. Also big thanks to the organizers of the fall 2004 research program and the staff of the institute for giving me the opportunity to spend a semester in a wonderful research atmosphere.

I thank all my colleagues at the TKK Institute of Mathematics, TKK Networking Laboratory, VTT Research Centre, and the Finnish Graduate School in Stochastics for all the things they have taught me in various discussions, lectures, seminars, and summer schools. In particular, I thank Dr Jarmo Malinen for several conversations that have greatly developed my understanding of mathematics in general, and Professor Juha Kinnunen for teaching me many valuable things about the fundamentals of mathematical analysis.

For everything else that I have learned, I give my props to my fiancée, family, and friends.

Helsinki, November 2005 Lasse Leskelä

This thesis has been financed by the Finnish Graduate School in Stochastics, the Academy of Finland Teletronics II research program, Emil Aaltosen Säätiö, the Mittag-Leffler Institute, Tekniikan edistämissäätiö, Yrjö Uiton Säätiö, the Finnish Foundation for Share Promotion, and Ida Montinin Säätiö. I gratefully acknowledge the support of all these organizations.

List of publications

The dissertation consists of this overview and the following publications:

[I] Leskelä, L. (2006). Stabilization of an overloaded queueing network using measurement-based admission control. *Journal of Applied Probability* 43(1), to appear, 14 pages.

[II] Leskelä, L. and Resing, J. (2004). A tandem queueing network with feedback admission control. *Institut Mittag-Leffler Report No. 09, 2004/2005, fall*, 9 pages.

[III] George, L., Jonckheere, M. and Leskelä, L. (2005). Does repacking improve performance of multiclass loss networks with overflow routing? In X. Liang, Z. Xin, V. B. Iversen, G. S. Kuo (Eds.), *Proceedings of the 19th International Teletraffic Congress* (pp. 1365–1373). Beijing University of Posts and Telecommunications Press.

[IV] Kaj, I., Leskelä, L., Norros, I. and Schmidt, V. (2005). Scaling limits for random fields with long-range dependence. *Institut Mittag-Leffler Report No. 24, 2004/2005, fall*, 25 pages.

Contribution

[I] This article reports my independent research.

[II] The idea of adaptive sampling and the queue length analysis in Sections 2.1–2.2 are mainly due to J. Resing. I am mostly responsible for the performance analysis and the numerical computations parts, reported in Sections 3.1–3.3.

[III] The analytical part of the paper is joint work of me and M. Jonckheere, while L. George is to be credited for the numerical simulations. The key mathematical result given in Section 3.2, required for proving the main theorems, is my independent work.

[IV] The main theorem, given in Section 3.2, is joint work of I. Kaj, me, and I. Norros. The theorems and proofs in Sections 2.2, 3.1, 5.1–5.3, and 5.5 are my independent work; so is most of the mathematics reported in Sections 3.3, 5.4, and 5.8, as well as the discussion of long-range dependence presented in Sections 2.1 and 4.1–4.2. The initiative for the study was suggested by V. Schmidt.

For all articles included in the thesis, I have done most of the writing.

Contents

Acknowledgments **v**

List of publications **vi**

Contribution **vi**

1 Introduction **1**

2 Analytical methods **3**
2.1 Stability of Markov processes 3
2.2 Markov reward approach 4
2.3 Coupling 5
2.4 Regular variation 5
2.5 Maximal functions 6

3 Queueing network with feedback admission control **7**
3.1 Model description 7
3.2 Stability analysis 7
3.3 Performance analysis 8
3.4 Discussion 9

4 Loss network with overflow routing **11**
4.1 Model description 11
4.2 Performance analysis 11
4.3 Discussion 12

5 Spatial random field generated by noninteracting sources **13**
5.1 Model description 13
5.2 Scaling approximations 14
5.3 Discussion 15

6 Conclusion **17**

References **19**

Publications [I–IV]

1 Introduction

Stochastic networks are mathematical models for traffic flows in networks with uncertainty, where uncertainty is modeled using random variables. The advantage of using random variables is that they allow one to quantify the amount of uncertainty, and to analyze how uncertainty cumulates in systems composed of independent random components. Some practical problems that can be studied in terms of stochastic networks include:

- How does adaptive packet dropping at the boundary routers of a packet-switched communications network affect performance?
- In a telephone call center, what is the tradeoff between employing well-trained call agents, capable of serving many customer types, versus untrained agents?
- In a mobile radio network, is there a parsimonious way of describing the spatial interference generated by the transmitting terminals?

Rather than developing a general analytical framework, this thesis finds answers to the above type of questions by considering three concrete instances of stochastic networks, each representing a specific challenge for analytical modeling. The first is a queueing network with distributed admission control for which stability conditions and a numerical algorithm for performance evaluation are presented in Section 3. The second is a multiclass loss network with overflow routing, for which analytical performance bounds are proved in Section 4. The third case, treated in Section 5, considers approximations for a spatial random field generated by a large number of noninteracting sources.

The above type of networks are hard to analyze because of the nonlinear blocking and feedback effects in the network interactions; exact mathematical model descriptions are typically not available, and rarely useful in helping to understand the model dynamics. To overcome these difficulties, two types of methodologies are applied and developed in this thesis: stochastic comparison and stochastic scaling. The idea of stochastic comparison is to find an analytically or computationally simpler reference model and prove that it performs better or worse than the original model. Upper and lower bounds for the performance of the original model can then be calculated using the reference model. The same idea is also used for deriving stability conditions, by proving that the stability of a model implies the stability of another, analytically tractable model. Stochastic scaling here means looking at renormalized limit distributions of models composed of independent random objects, as the number of objects grows to infinity. This type of limits are used to smooth out unnecessarily complex model details and derive simple approximations for macroscopic statistical properties of the system.

The rest of the thesis is organized as follows. Section 2 outlines various mathematical techniques of probability theory and functional analysis that have been developed and applied in the study of the models. Sections 3–5 summarize the results that have been obtained in the original publications [I–IV] with a short discussion on extensions, relevance, and originality of the results appended at the end of each section. Section 6 concludes the thesis, and the original research articles are attached to the end.

2 Analytical methods

This section gives a walk-through of the key results in probability theory and functional analysis that are used and developed in the thesis. Some of the results, with a Roman numeral referring to the publication that contains the proof, have been independently formulated and proved by me.

2.1 Stability of Markov processes

The following continuous-time analogue of Foster's classical theorem [5] provides a sufficient condition for a discrete-space Markov process to be ergodic, that is, irreducible and positive recurrent. If $q(x,y)$ is a set of transition rates of a Markov process, denote the net drift out of state x by $q(x) = \sum_{y \neq x} q(x,y)$.

Theorem (Tweedie [29]). *Let X be an irreducible Markov process on a countable state space S generated by transition rates $q(x,y)$ so that $q(x) < \infty$ for all x. Assume that there exists a function $V : S \to \mathbb{R}$ such that*

- $\sum_{y \neq x} |V(y) - V(x)|\, q(x,y) < \infty$ *for all x,*
- $\lim_{x \to \infty} V(x) = \infty$, *and*
- $\sup_{x \in S \setminus S_0} \sum_{y \neq x} (V(y) - V(x))\, q(x,y) < 0$ *for some finite set S_0.*

Then X is ergodic.

The above theorem is used to derive sufficient stability conditions for the queueing network studied in Section 3. One way to study the necessity of the conditions is to show that the stability of the original system implies the stability of a simpler system. In proving this type of implication, the following lemma is used. If $q(x,y)$ and $q'(x,y)$ are generators of Markov processes on a countable state space S, denote

$$D(q,q') = \{x : q(x,y) \neq q'(x,y) \text{ for some } y\}$$

and

$$\overline{D}(q,q') = D(q,q') \cup \{y : q(x,y) > 0 \text{ or } q'(x,y) > 0 \text{ for some } x \in D(q,q')\}.$$

Lemma [I]. *Let X and X' be irreducible Markov processes on a countable state space S generated by $q(x,y)$ and $q'(x,y)$, respectively, with $q(x)$, $q'(x) < \infty$ for all x. Assume that $\overline{D}(q,q')$ is finite. Then X is ergodic if and only if X' is ergodic.*

2.2 Markov reward approach

Let X be an irreducible and positive recurrent continuous-time Markov process on a countable state space S with generator matrix Q. Assume that the entries $q(x,y)$ of Q satisfy $\sum_{y \neq x} q(x,y) \leq \Lambda$ for all x, where Λ is a positive constant independent of x. Let Y_n be a Markov chain with transition matrix $Q_\Lambda = I + \Lambda^{-1}Q$, and let N be a Poisson process with rate Λ, independent of the chain Y_n. Then the processes $X(t)$ and $Y_{N(t)}$ have the same distribution (Kallenberg [13] Proposition 12.20). Moreover, it follows from the definition of Q_Λ that the processes X and Y share the same invariant probability distribution. Let r be a positive function on S, and let

$$V^n(x) = \sum_{j=0}^{n-1} Q_\Lambda^j r(x),$$

where $Q_\Lambda^j r(x) = \mathrm{E}_x\, r(Y_j)$ is the expectation of r evaluated at the j-th step of the chain Y started at x. Then the ergodicity of X implies that for all initial states x,

$$\lim_{n \to \infty} \frac{V^n(x)}{n} = \mathrm{E}\, r(X).$$

Assume that X' is another continuous-time Markov process on S with generator Q' satisfying the same regularity assumptions as X. Let r' be a positive function on S. The following theorem, a consequence of the above limit property, is the key result of the Markov reward approach.

Theorem (van Dijk [3]). *Suppose that for all $x \in S$ and all n,*

$$(r' - r)(x) + \sum_{y \neq x} \big(q'(x,y) - q(x,y)\big)\big(V^n(y) - V^n(x)\big) \geq 0.$$

Then

$$\mathrm{E}\, r'(X') \geq \mathrm{E}\, r(X).$$

In large complex systems, the stationary distributions of X and X' are often impossible to evaluate numerically. The above theorem is important, because it allows one to compare $\mathrm{E}\, r(X)$ and $\mathrm{E}\, r'(X')$ without explicitly knowing the stationary distributions. This theorem provides a method for proving the results of Section 4, where a numerically computable upper bound for $\mathrm{E}\, r(X)$ is derived by finding a structurally simpler Markov process X' whose generator satisfies the condition of the theorem with $r' = r$.

Analogous results for comparing Markov processes with respect to stochastic integral orderings are given by Massey [24] and Whitt [30]. However, these stronger ordering results are not directly applicable to the model studied in Section 4, due to multidimensional blocking.

2.3 Coupling

Let $X = (X_n)_{n\geq 0}$ and $Y = (Y_n)_{n\geq 0}$ be two discrete-time stochastic processes on a state space S. Assume that there exist two S-valued stochastic processes $\tilde{X}$ and $\tilde{Y}$ defined on the same probability space $(\Omega, \mathcal{F}, \mathrm{P})$ such that $\tilde{X}$ has the same distribution as X, and $\tilde{Y}$ has the same distribution as Y. Then the pair $(\tilde{X}, \tilde{Y})$ is called a coupling of the processes X and Y.

Finding a coupling of two processes X and Y with special path properties sometimes allows one to make interesting conclusions on how the distributions of X and Y are related. For example, assume that there exists a coupling $(\tilde{X}, \tilde{Y})$ such that $r(\tilde{Y}_n) \geq r(\tilde{X}_n)$ almost surely for all n, where r is positive function on S. Then it immediately follows that $\mathrm{E}\, r(Y_n) \geq \mathrm{E}\, r(X_n)$ for all n.

While making distributional conclusions using a coupling is trivial, finding a coupling with desired path properties is often not, requiring a careful analysis of the dynamics of the two processes. The key observation required for proving the results of Section 4 is to couple two copies of the same Markov process started at different initial states in such a way that the trajectories of the processes remain close to each other [III, Lemma 1]. This coupling is later combined with the Markov reward approach introduced in Section 2.3 to compare the stationary performance of two stochastic networks [III, Lemma 2].

2.4 Regular variation

A probability distribution F on $\mathbb{R}_+$ is said to have a regularly varying tail of exponent $\gamma > 0$, if

$$\lim_{v\to\infty} \frac{\bar{F}(av)}{\bar{F}(v)} = a^{-\gamma} \quad \text{for all} \quad a > 0,$$

where $\bar{F}(v) = 1 - F(v)$. The following results are consequences of Karamata's theorem [14] (see Bingham, Goldie, and Teugels [2] for a textbook exposition).

Lemma [IV]. *Let F be a probability distribution on $\mathbb{R}_+$ with a regularly varying tail of exponent $\gamma > 0$, and define the scaled distribution F_ρ for $\rho > 0$ by $F_\rho(v) = F(v/\rho)$. Assume that $f(v)$ is a continuous function on $\mathbb{R}_+$ such that for some $0 < p < \gamma < q$,*

$$\limsup_{v\to\infty} v^{-p}|f(v)| < \infty \quad \text{and} \quad \limsup_{v\to 0} v^{-q}|f(v)| < \infty.$$

Then

$$\int_{\mathbb{R}_+} f(v) F_\rho(dv) \sim \bar{F}_\rho(1) \int_0^\infty f(v)\, \gamma v^{-\gamma-1} dv \quad \text{as } \rho \to 0.$$

Lemma [IV]. *Let F and F_ρ be defined as above and let $f_\rho(v)$ be a family of measurable functions on $\mathbb{R}_+$. Assume that for some $0 < p < \gamma < q$, either*

$$\lim_{\rho\to 0} \sup_{v>a} v^{-p} \bar{F}_\rho(1)|f_\rho(v)| = 0 \quad \text{for all } a > 0,$$
$$\text{and} \quad \bar{F}_\rho(1)|f_\rho(v)| \leq cv^q \quad \text{for all } \rho, v,$$

or

$$\lim_{\rho\to 0} \sup_{a\le v} v^{-q} \bar{F}_\rho(1) |f_\rho(v)| = 0 \quad \textit{for all } a > 0,$$

$$\textit{and} \quad \bar{F}_\rho(1)|f_\rho(v)| \le cv^p \quad \textit{for all } \rho, v,$$

Then

$$\lim_{\rho\to 0} \int_{\mathbb{R}_+} f_\rho(v)\, F_\rho(dv) = 0.$$

2.5 Maximal functions

Let C be a bounded measurable set in $\mathbb{R}^d$ with unit volume. If ϕ is a locally integrable function, define the averages $m_\phi(x,v)$ by

$$m_\phi(x,v) = v^{-1} \int_{x+v^{1/d}C} \phi(y)\, dy,$$

and let ϕ_* be the maximal function of ϕ given by

$$\phi_*(x) = \sup_{v>0}\; v^{-1} \int_{x+v^{1/d}C} |\phi(y)|\, dy.$$

The following result, a consequence of Hardy–Littlewood maximal function theory [10] (see Rudin [27] for a textbook exposition), is used to find integrable upper bounds for the characteristic functional of the random field studied in Section 5.

Lemma [IV]. *Let C be a bounded measurable set in $\mathbb{R}^d$ with unit volume.*

(i) If $\phi \in L^1$, then $\lim_{v\to 0} m_\phi(x,v) = \phi(x)$ for almost all x.

(ii) If $\phi \in L^1$, then $\phi_*(x) < \infty$ for almost all x.

(iii) If $\phi \in L^p$ for some $p > 1$, then $\phi_* \in L^p$.

3 Queueing network with feedback admission control

3.1 Model description

The input to the system is modeled as a Poisson process with rate λ, and the service times of jobs in nodes 1 and 2 are assumed to be independent exponentially distributed random variables with parameters μ_1 and μ_2, respectively. The state of the network is denoted by $X(t) = (X_1(t), X_2(t))$, where $X_i(t)$ is the amount of jobs in node i at time t. The admission control scheme is based on X_2, the number of jobs present in node 2. As long as X_2 is smaller than or equal to a certain threshold level K, new jobs arriving to the network are accepted; otherwise they are rejected (Figure 1).

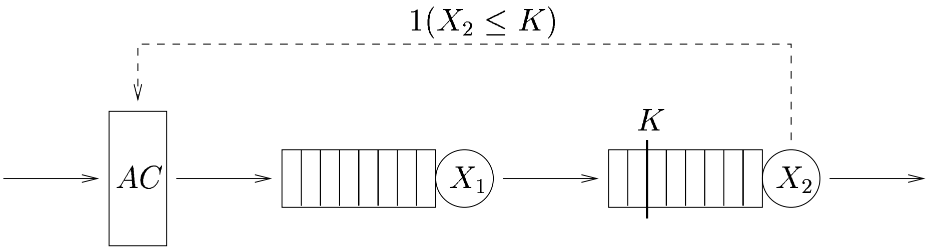

Figure 1: The admission control mechanism.

3.2 Stability analysis

The stability region of the queueing network of Figure 1 is given by the following theorem.

Theorem [I]. *The Markov process X is positive recurrent if and only if*

$$\lambda \left(1 - (\mu_1/\mu_2)^{K+1}\right) < \mu_1.$$

Especially, when $\mu_1 \geq \mu_2$, the system is stable for all λ and K.

Observe that $\mu_1 < \mu_2$ for unstable networks, so that the rate at which work is fed into the second server is strictly less than its service capacity. Thus, intuition suggests that only the first queue will grow to infinity. The next theorem makes these heuristics rigorous.

Theorem [II]. *In an unstable system, the process X started from an arbitrary initial state satisfies as $t \to \infty$,*

$$\begin{aligned} X_1(t) &\to \infty \quad \textit{almost surely}, \\ X_2(t) &\to Z \quad \textit{in distribution}, \end{aligned}$$

where Z is a geometrically distributed random variable with parameter μ_1/μ_2.

3.3 Performance analysis

Assume from now on that the process $X = (X_1, X_2)$ describing the amount of customers in the system is ergodic, and consider the *censored process* $Y = (Y_1, Y_2)$ constructed by sampling X at periods of time during which $X_2 \leq K$. It follows that also Y is a Markov process. Denote by I the identity matrix, and let T_L and T_R be the left and right shift matrices, and let U_0 and U_K be the projection matrices onto the 0-th and the K-th coordinate in $\mathbb{R}^{K+1}$, respectively. The generator of Y can then be written in the form

$$Q = \begin{pmatrix} B_0 & A_0 & 0 & 0 & \cdots \\ B_1 & A_1 & A_0 & 0 & \cdots \\ B_2 & A_2 & A_1 & A_0 & \cdots \\ B_3 & A_3 & A_2 & A_1 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \end{pmatrix},$$

where the matrices A_n and B_n are given by

$$\begin{aligned} A_0 &= \lambda I, \\ A_1 &= \mu_2 T_L - (\lambda + \mu_1 + \mu_2) I + \mu_2 U_0, \\ A_2 &= \mu_1 (T_R + q_1 U_K), \\ A_{n+1} &= \mu_1 q_n U_K, \quad n \geq 2, \end{aligned}$$

and

$$\begin{aligned} B_0 &= \mu_2 T_L - (\lambda + \mu_2) I + \mu_2 U_0, \\ B_1 &= \mu_1 (T_R + U_K), \\ B_{n+1} &= \mu_1 (1 - q_1 - \cdots - q_n) U_K, \quad n \geq 1, \end{aligned}$$

and where the constants q_n are given by

$$q_{n+1} = \frac{1}{n+1} \binom{2n}{n} \left(\frac{\mu_1}{\mu_1 + \mu_2} \right)^n \left(\frac{\mu_2}{\mu_1 + \mu_2} \right)^{n+1}.$$

Denote by e_k the k-th unit (row) vector of $\mathbb{R}^{K+1}$, and let $e = \sum_{k=0}^{K} e_k$. Then according to Neuts [26], the steady-state probabilities of the censored process can be expressed in the matrix-geometric form

$$\mathrm{P}(Y = (n,k)) = x_0 R^n e_k^T, \quad (n,k) \in S^-, \tag{1}$$

where the matrix R is the unique minimal non-negative solution of

$$\sum_{n=0}^{\infty} R^n A_n = 0, \tag{2}$$

and x_0 is the unique positive row vector satisfying

$$x_0 \sum_{n=0}^{\infty} R^n B_n = 0 \quad \text{and} \quad x_0 (I - R)^{-1} e^T = 1. \tag{3}$$

The matrix R can be numerically solved from equation (2) using the method of successive substitutions, as described by Neuts [26]. When R is calculated, the vector x_0 will be obtained from (3). The following theorem shows how the throughput θ and the mean sojourn time $\mathrm{E}(D)$ in the original system can be calculated using the steady-state distribution (1) of the censored process Y.

Theorem [II]. *The steady-state throughput θ and the steady-state mean sojourn time $\mathrm{E}(D)$ of jobs accepted to the network are given in terms of the equilibrium distribution of Y by*

$$\frac{1}{\theta} = \frac{1}{\lambda}\,\mathrm{P}(Y_2 = 0) + \frac{1}{\mu_2},$$

and

$$\mathrm{E}(D) = \frac{1}{\lambda}\,\mathrm{E}(Y_1; Y_2 = 0) + \frac{1}{\mu_2}\,\mathrm{E}(Y_1 + Y_2 + 1).$$

3.4 Discussion

The stability analysis presented in this section can be extended to more general admission control policies, where the input rate to the first node is allowed to be an arbitrary function of the number of customers in the second node [I, Theorems 2–4]. This type of generalization allows modeling of networks where during congestion, the input traffic is gradually thinned by randomly rejecting a certain proportion of the incoming traffic. Assuming that the generalized admission policy strictly blocks all incoming traffic when the amount of jobs in the second node exceeds a certain maximum threshold, then the stability region of the system can be characterized exactly [I, Theorem 5]. The performance analysis presented in Section 3.3 also extends without difficulties to this type of generalized admission policies [II, Remark 2.1].

Other types of queueing networks with feedback signaling have been studied earlier. In the papers of van Foreest, Mandjes, van Ommeren, and Scheinhardt [4], Grassman and Drekic [9], Konheim and Reiser [17], and Latouche and Neuts [21], the first server stops processing when the number of jobs at the second station becomes too high. Kroese, Scheinhardt, and Taylor [19] have studied a network where new jobs are rejected whenever the number of jobs at the first station reaches a certain threshold. The queueing network studied in this section differs from the above models in that both queues can grow arbitrarily large, which makes stability analysis more difficult. For networks with unlimited buffers and state-dependent service times, Bambos and Walrand [1] have derived stability results extending to non-Markovian systems, however ruling out the type of feedback signaling loops presented here.

4 Loss network with overflow routing

4.1 Model description

Consider a loss network serving K classes of jobs. The network consists of M_k monoskill servers assigned to serving jobs of class k, and N multiskill servers capable of serving all classes of jobs. Customers of class k arrive according to a Poisson process of intensity λ_k and require exponential service times with mean $1/\mu_k$. Assume that the arrival processes and the service times are independent. Arriving service requests are routed according to an overflow policy, where a class-k job is always routed to a corresponding monoskill server, if there is one available. If all monoskill servers for class k are busy, the job is routed to a multiskill server. If all multiskill servers are also busy, then the service request is rejected (Figure 2). The rejected requests are assumed to leave the system without retrials.

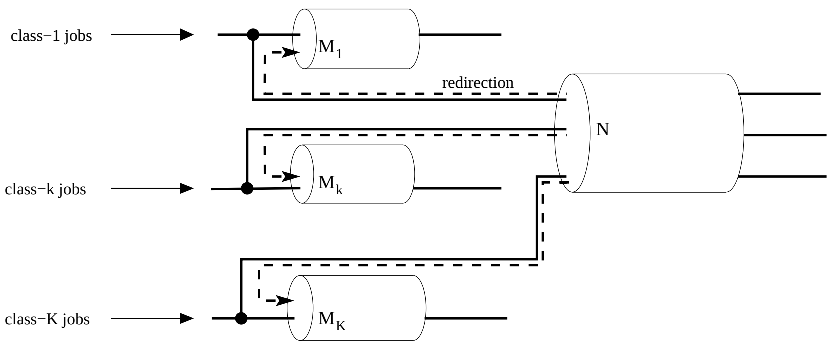

Figure 2: Overflow system and repacking.

Let $X_{1,k}$, $X_{2,k}$ be the number of class-k jobs being served by the monoskill and multiskill servers, respectively. Then the network state can be described by the random vector $X = (X_{i,k})$ indexed by $i = 1, 2$ and $k = 1, \ldots, K$. By construction, X is a Markov process with a finite state space.

Consider the same network with slightly modified routing, where jobs are redirected from multiskill to monoskill servers as soon as places become vacant, and denote the corresponding Markov process by $X' = (X'_{i,k})$. The repacking policy guarantees that all states x with $x_{1,k} < M_k$ and $x_{2,k} > 0$ are transient for X' and thus have zero stationary probability. As a consequence of this special feature, the stationary distribution of the process $\tilde{X}' = (X'_{1,k} + X'_{2,k})_{k=1}^K$ describing the net amount of class-k jobs can be expressed in product form (see for example Kaufman [15])

$$\frac{\mathrm{P}(\tilde{X}' = x)}{\mathrm{P}(\tilde{X}' = 0)} = \prod_k \frac{(\lambda_k/\mu_k)^{x_k}}{x_k!}. \tag{4}$$

4.2 Performance analysis

Performance is measured in terms of *carried load*, defined as the stationary mean rate of work served by the network, and *overall blocking prob-*

ability, which is the probability that an arbitrary arriving job is rejected. Because neither X nor its aggregated version are reversible, there are no simple closed form expressions for the performance quantities of the original system. Moreover, the size of the state space makes brute force numerical solution of the stationary distribution of X unfeasible (for example with $K = 2$, $M_1 = M_2 = N = 9$, the generator of X has over 30 million entries). On the other hand, the above performance quantities for the system with repacking are readily computed using (4). The following theorem together with its corollaries shows how the system with repacking provides upper bounds for the performance of the original system.

Theorem [III]. *In the stationary regime, the mean net amount of jobs in the system with repacking is greater than or equal to the corresponding quantity in the original system.*

Corollary [III]. *The carried load in the system with repacking is greater than or equal to the carried load in the original system.*

Corollary [III]. *When $\mu_k = \mu$ for all k, then the overall blocking probability for the system with repacking is less than or equal to the overall blocking probability in the original system.*

4.3 Discussion

Increasing the system utilization by repacking may lead to worse performance in terms of overall blocking probability, if the mean service times of job classes differ significantly [III, Example 1]. However, there are numerical simulations illustrating that the performance of the system with repacking is close to the original system for a wide range of parameters [III, Sections 4.1–4.2]. Practical relevance of the performance bounds can be criticized, because they tell nothing about the worst-case behavior of the original system. The derivation of computable lower bounds for the system performance remains an interesting open problem.

Performance of loss networks is a well-studied topic in applied probability (Kelly [16]). The computational complexity of loss networks has been discussed by Louth, Mitzenmacher, and Kelly [22]. Some numerical methods for calculating the blocking probability in multiclass loss networks include the Hayward–Fredericks method [7] and the hyperexponential decomposition introduced by Franx, Koole, and Pot [6]; the latter provides accurate estimates but is computationally demanding. The repacking system discussed in this thesis is computationally efficient; its accuracy as an approximation for the multiclass loss network deserves to be studied in more detail.

5 Spatial random field generated by noninteracting sources

5.1 Model description

Let C be a bounded measurable set in $\mathbb{R}^d$ such that $|C| = 1$ and $|\partial C| = 0$, where ∂C denotes the boundary of C and $|\cdot|$ is the Lebesgue measure. Let $X_j + (\rho V_j)^{1/d} C$ be a family of random sets, called grains, with random locations X_j and random volumes ρV_j. Assume that X_j are uniformly distributed in the space according to a Poisson random measure with mean density $\lambda > 0$, and that V_j are independent copies of a positive random variable V with $\mathrm{E}\, V = 1$, also independent of the locations X_j. Hence the scalar $\rho > 0$ equals the mean grain volume. Our goal is to study the cumulative mass field $J_{\lambda,\rho}$ generated by the grains, defined by

$$J_{\lambda,\rho}(A) = \sum_j \left| A \cap \left(X_j + (\rho V_j)^{1/d} C \right) \right|,$$

where A ranges over bounded measurable sets in $\mathbb{R}^d$. Figure 3 illustrates a pseudorandom sample of the grain field $J_{\lambda,\rho}$ simulated over a rectangular grid on the unit rectangle with 512x512 resolution, where the volume has Pareto distribution with parameter $\gamma = 1.5$. On the left, the grain shape C is a circle, and on the right, ellipse. When $\mathrm{E}\, V^2 = \infty$, the functional $J_{\lambda,\rho}$ has

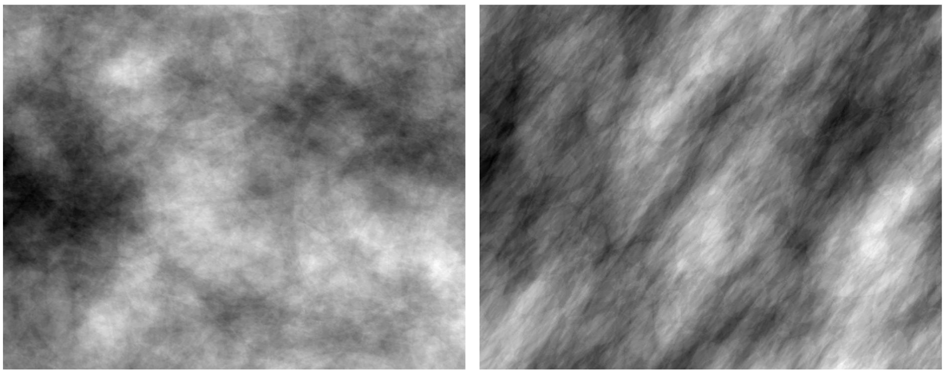

Figure 3: Simulated grain field generated by symmetric (left) and asymmetric (right) grains, with $\lambda = 1000$, $\rho = 0.1$.

long-range spatial dependence [IV] in the sense that

$$\lim_{r\to\infty} \left| \mathrm{Cov}\left(J_{\lambda,\rho}(B_1),\, J_{\lambda,\rho}(B_r \setminus B_1) \right) \right| = \infty,$$

where B_r denotes the open ball centered at the origin with radius r.

For analytical convenience, extend $J_{\lambda,\rho}$ to a random linear functional, by defining for an arbitrary test function $\phi \in L^1$,

$$J_{\lambda,\rho}(\phi) = \sum_j \int_{X_j + (\rho V_j)^{1/d} C} \phi(y)\, dy.$$

5.2 Scaling approximations

Let us now consider the limiting behavior of $J_{\lambda,\rho}$ as the mean grain density λ grows to infinity and the mean grain volume ρ shrinks to zero. When the grain volume distribution has finite variance, the following theorem shows that the centered and renormalized version of $J_{\lambda,\rho}$ converges to white Gaussian noise.

All limits in this section are to be understood in the sense of finite-dimensional distributions of random linear functionals on $L^1 \cap L^2$.

Theorem [IV]. *Let C be a bounded set with $|C| = 1$ and $|\partial C| = 0$, and assume $\mathrm{E}\, V^2 < \infty$. Then as $\lambda \to \infty$ and $\rho \to 0$,*

$$\frac{J_{\lambda,\rho}(\phi) - \mathrm{E}\, J_{\lambda,\rho}(\phi)}{\rho(\lambda \,\mathrm{E}\, V^2)^{1/2}} \longrightarrow W(\phi),$$

where W is the centered Gaussian random linear functional on L^2 with $\mathrm{E}\, W(\phi)W(\psi) = \iint \phi(x)\psi(y)\, dx\, dy$.

The situation is different if the grain volumes are heavy-tailed with regularly varying tail of index $\gamma \in (1,2)$, which implies $\mathrm{E}\, V^2 = \infty$. As $\rho \to 0$, the expected number of grains with volume larger than one that cover the origin is asymptotically equivalent to a constant times $\lambda \bar{F}_\rho(1)$. Consequently, let us distinguish the following three scaling regimes (the constant γ^{-1} has been chosen for notational convenience):

Large-grain scaling	$\lambda \bar{F}_\rho(1) \to \infty$
Intermediate scaling	$\lambda \bar{F}_\rho(1) \to \gamma^{-1}$
Small-grain scaling	$\lambda \bar{F}_\rho(1) \to 0$

The following theorem describes the nontrivial stochastic limits that can be derived for different scaling regimes using suitable renormalization.

Theorem [IV]. *Let C be a bounded set with $|C| = 1$ and $|\partial C| = 0$, and assume that V has a regularly varying tail with exponent $\gamma \in (1,2)$. Then the following three limits hold as $\lambda \to \infty$ and $\rho \to 0$:*

(i) (Large-grain scaling) *If If $\lambda \bar{F}_\rho(1) \to \infty$, then*

$$\frac{J_{\lambda,\rho}(\phi) - \mathrm{E}\, J_{\lambda,\rho}(\phi)}{\left(\gamma \lambda \bar{F}_\rho(1)\right)^{1/2}} \longrightarrow W_{\gamma,C}(\phi),$$

where $W_{\gamma,C}$ is the centered Gaussian random linear functional on $L^1 \cap L^2$ with $\mathrm{E}\, W_{\gamma,C}(\phi) W_{\gamma,C}(\psi) = \iint \phi(x) K_{\gamma,C}(x-y)\psi(y)\, dx\, dy$, and

$$K_{\gamma,C}(x) = \int_0^\infty \left| \left(v^{-1/d}x + C\right) \cap C \right| v^{-\gamma} dv.$$

(ii) (Intermediate scaling) *If $\lambda \bar{F}_\rho(1) \to \gamma^{-1}$, then*

$$J_{\lambda,\rho}(\phi) - \mathrm{E}\, J_{\lambda,\rho}(\phi) \longrightarrow J^*_{\gamma,C}(\phi),$$

where $J^*_{\gamma,C}$ is a random linear functional with the same second order statistics as $W_{\gamma,C}$, and characteristic functional

$$\mathrm{E}\, e^{iJ^*_{\gamma,C}(\phi)} = \exp \int_{\mathbb{R}^d} \int_0^\infty \Psi \left(\int_{x+v^{1/d}C} \phi(y)\, dy \right) dx\, v^{-\gamma-1} dv,$$

with $\Psi(u) = e^{iu} - 1 - iu$.

(iii) (Small-grain scaling) *If $\lambda \bar{F}_\rho(1) \to 0$, then*

$$\frac{J_{\lambda,\rho}(\phi) - \mathrm{E}\, J_{\lambda,\rho}(\phi)}{c_\gamma (1/\bar{F}_\rho)^{\leftarrow}(\gamma\lambda)} \longrightarrow \Lambda_\gamma(\phi),$$

where Λ_γ is the independently scattered γ-stable random measure on $\mathbb{R}^d$ with Lebesgue control measure and unit skewness, $(1/\bar{F}_\rho))^{\leftarrow}(u) = \inf\{v : 1/\bar{F}_\rho(v) \geq u\}$ is the quantile function of F_ρ, and

$$c_\gamma = \left(-\frac{\Gamma(2-\gamma)}{\gamma(\gamma-1)} \cos\left(\frac{\pi\gamma}{2}\right) \right)^{-1/\gamma}.$$

To study the role of symmetry, let us define a slightly modified model $\widetilde{J}_{\lambda,\rho}$ where the grains are randomly oriented. Let θ_j be independent random rotations in $\mathbb{R}^d$, each uniformly distributed according to the Haar measure, and assume that θ_j are independent of the locations X_j and the volumes V_j. Then

$$\widetilde{J}_{\lambda,\rho}(\phi) = \sum_j \int_{X_j + (\rho V_j)^{1/d}\theta_j C} \phi(y)\, dy.$$

defines the analogue of $J_{\lambda,\rho}$ with randomly rotated grains $\theta_j C$. Renormalizing $\widetilde{J}_{\lambda,\rho}$ in the same way as $J_{\lambda,\rho}$, it is shown [IV] that the small-grain limit coincides with Λ_γ, and the intermediate limit has similar structure as $J^*_{\gamma,C}$, with the characteristic functional

$$\exp \int_{\mathbb{R}^d} \int_0^\infty \int_{SO(d)} \Psi \left(\int_{x+v^{1/d}\theta C} \phi(y)\, dy \right) dx\, v^{-\gamma-1} dv\, d\theta,$$

where $\Psi(u) = e^{iu} - 1 - iu$ and $d\theta$ is the Haar measure in the group of rotations $SO(d)$ in $\mathbb{R}^d$. Moreover, in the large-grain scaling, the renormalized limit of $\widetilde{J}_{\lambda,\rho}$ corresponds to a constant multiple of fractional Gaussian noise with Hurst parameter $H = (3-\gamma)/2$, as defined by Mandelbrot and van Ness [23].

5.3 Discussion

The results in this section illustrate the difficulty of deriving robust statistical approximations for cumulative random fields generated as superpositions of independent random objects, if the objects have long-range spatial dependence. Fractional Gaussian noise and stable random measures, both obtained as renormalized limits of the same model, represent opposite extremes in the range of commonly used stochastic models; and it is impossible to distinguish from a finite number of statistical tests which of the two (or in the intermediate case, neither) should be used for approximating a given model. Although the results presented here do not seem very useful from the modeling point of view, they might be valuable in explaining how macroscopic physical structures with long-range dependent statistical features are formed. Some of the limit results discussed in this section can also be extended to random functionals over spaces of signed measures [IV]. This type of generalized viewpoint might be helpful in revealing connections with other parts of mathematics, such as potential theory and geometric measure theory.

The type of dichotomy discussed above has earlier been observed by Taqqu and Levy [28], who showed that both fractional Brownian motion and stable Lévy motion can appear as renormalized limits of a renewal-reward process, depending on the order of taking a double limit. Analogous limit results have been derived by Konstantopoulos and Lin [18], Kurtz [20], and Mikosch, Resnick, Rootzén, and Stegeman [25], who studied simultaneuous double limits are for a data traffic model known as the infinite source Poisson model. The critical boundary between the two scaling regimes leading to fractional Brownian motion and stable Lévy motion has been studied by Gaigalas and Kaj [8], Kaj and Martin–Löf [11], and Kaj and Taqqu [12]. The results reported in this thesis extend the analysis of the above type of double limits to spatial random fields, and show how results analogous to those of Kaj and Taqqu [12] and Mikosch et al. [25] can be obtained for spatial models with long-range dependence. Another new contribution is the use of test functions for characterizing the limits. In contrast with dimension one, where stationary random processes can be analyzed in terms of their integrated versions (for example fractional Brownian motion), in multidimensional spaces it is necessary to consider the stationary objects themselves as generalized random functions. The results characterizing test function spaces suitable for analyzing fractional Gaussian noise and related anisotropic self-similar random fields [IV] can be helpful in applications of spatial stochastic models.

6 Conclusion

In this thesis I have studied how stochastic comparison and scaling methods can be applied to derive simple approximative descriptions for performance and stability of complex network models involving feedback, blocking, and long-range dependence. Stochastic comparison methods can be useful in stability analysis, as illustrated by the results of Section 2.1 applied to the queueing network in Section 3, and in performance analysis, as discussed in Section 4. For systems composed of a large number of independently acting random sources, the results of Section 5 show how stochastic scaling methods can be used to approximate cumulative macroscopic effects of the sources.

The results of the thesis illustrate the feasibility of stochastic comparison and stochastic analysis in deriving approximations and performance bounds for complex physical networks with uncertainty. Approximations and performance bounds based on exact mathematical methods have the advantage that they explicitly state the type of circumstances required for the accuracy of the estimates. The resulting analytical formulas can reveal interesting properties that are not easily detected using numerical simulation, such as the stability region of the queueing network in Section 3, the monotonicity properties of the loss network in Section 4, or the critical scaling regime for the spatial field of Section 5.

Two probabilistic assumptions, independence and lack of memory (the exponential probability distribution), have been actively used throughout the thesis. These assumptions may be criticized to be unrealistic in many applications. For example, if the queueing network of Section 3 is used to model the transfer of random-sized files in a computer network with deterministic servers, then the service times in different nodes of the network are not independent. Extending the results of the thesis to models where the independence and lack of memory assumptions are relaxed remains an important direction for future research.

References

[1] Bambos, N., and Walrand, J. On stability of state-dependent queues and acyclic queueing networks. *Adv. Appl. Probab. 21* (1989), 618–701.

[2] Bingham, N. H., Goldie, C. M., and Teugels, J. L. *Regular Variation.* Cambridge University Press, 1987.

[3] van Dijk, N. M. Bounds and error bounds for queueing networks. *Ann. Oper. Res. 79* (1998), 295–319.

[4] van Foreest, N. D., Mandjes, M. R. H., van Ommeren, J. C. W., and Scheinhardt, W. R. W. A tandem queue with server slow-down and blocking. *Stochastic Models 21*, 2–3 (2005), 695–724.

[5] Foster, F. G. On the stochastic matrices associated with certain queuing processes. *Ann. Math. Statist. 24* (1953), 355–360.

[6] Franx, G. J., Koole, G., and Pot, A. Approximating multi-skill blocking systems by hyperexponential decomposition. *Perform. Evaluation* (2005). In press.

[7] Fredericks, A. Congestion in blocking systems – a simple approximation technique. *Bell Syst. Tech. J. 59*, 6 (1980), 805–827.

[8] Gaigalas, R., and Kaj, I. Convergence of scaled renewal processes and a packet arrival model. *Bernoulli 9*, 4 (2003), 671–703.

[9] Grassmann, W. K., and Drekic, S. An analytical solution for a tandem queue with blocking. *Queueing Syst. 36*, 1 (2000), 221–235.

[10] Hardy, G. H., and Littlewood, J. E. A maximal theorem with function-theoretic applications. *Acta Math. 54* (1930), 81–116.

[11] Kaj, I., and Martin-Löf, A. Scaling limit results for the sum of many inverse Lévy subordinators. *Stoch. Proc. Appl.* (2005). In press.

[12] Kaj, I., and Taqqu, M. S. Convergence to fractional Brownian motion and to the Telecom process: The integral representation approach. Preprint 2004:16. Dept. of Math., Uppsala University, 2004.

[13] Kallenberg, O. *Foundations of Modern Probability.* Springer, 2001.

[14] Karamata, J. Sur un mode de croissance régulière des fonctions. *Mathematica (Cluj) 4* (1930), 38–53.

[15] Kaufman, J. S. Blocking in a shared resource environment. *IEEE Trans. Comm. 29*, 10 (1981), 1474–1481.

[16] Kelly, F. P. Loss networks. *Ann. Appl. Probab. 1* (1991), 319–378.

[17] KONHEIM, A. G., AND REISER, M. A queueing model with finite waiting room and blocking. *J. ACM 23*, 2 (1976), 328–341.

[18] KONSTANTOPOULOS, T., AND LIN, S.-J. Macroscopic models for long-range dependent network traffic. *Queueing Syst. 28*, 1–3 (1998), 215–243.

[19] KROESE, D. P., SCHEINHARDT, W. R. W., AND TAYLOR, P. G. Spectral properties of the tandem Jackson network, seen as a quasi-birth-and-death process. *Ann. Appl. Probab. 14*, 4 (2004), 2057–2089.

[20] KURTZ, T. G. Limit theorems for workload input models. In *Stochastic networks: Theory and applications*, F. P. Kelly, S. Zachary, and I. Ziedins, Eds. Oxford University Press, 1996, pp. 119–139.

[21] LATOUCHE, G., AND NEUTS, M. F. Efficient algorithmic solutions to exponential tandem queues with blocking. *SIAM J. Algebra. Discr. 1* (1980), 93–106.

[22] LOUTH, G., MITZENMACHER, M., AND KELLY, F. P. Computational complexity of loss networks. *Theor. Comp. Sc. 125* (1994), 45–59.

[23] MANDELBROT, B. B., AND VAN NESS, J. W. Fractional Brownian motions, fractional noises and applications. *SIAM Rev. 10*, 4 (1968), 422–437.

[24] MASSEY, W. A. Stochastic orderings for Markov processes on partially ordered spaces. *Math. Oper. Res. 12*, 2 (1987), 350–367.

[25] MIKOSCH, T., RESNICK, S. I., ROOTZÉN, H., AND STEGEMAN, A. Is network traffic approximated by stable Lévy motion or fractional Brownian motion? *Ann. Appl. Probab. 12*, 1 (2002), 23–68.

[26] NEUTS, M. F. *Matrix-Geometric Solutions in Stochastic Models.* John Hopkins University Press, 1981.

[27] RUDIN, W. *Real and Complex Analysis*, third ed. McGraw–Hill, 1987.

[28] TAQQU, M. S., AND LEVY, J. B. Using renewal processes to generate long-range dependence and high variability. In *Dependence in Probability and Statistics*, E. Eberlein and M. S. Taqqu, Eds. Birkhäuser, 1986, pp. 73–89.

[29] TWEEDIE, R. L. Sufficient conditions for regularity, recurrence and ergodicity of Markov processes. *Math. Proc. Cambridge 78* (1975), 125–136.

[30] WHITT, W. Stochastic comparisons for non-Markov processes. *Math. Oper. Res. 11*, 4 (1986), 608–618.

(continued from the back cover)

A484 Visa Latvala , Niko Marola , Mikko Pere
Harnack's inequality for a nonlinear eigenvalue problem on metric spaces
March 2005

A482 Mikko Lyly , Jarkko Niiranen , Rolf Stenberg
A refined error analysis of MITC plate elements
April 2005

A481 Dario Gasbarra , Tommi Sottinen , Esko Valkeila
Gaussian Bridges
December 2004

A480 Ville Havu , Jarmo Malinen
Approximation of the Laplace transform by the Cayley transform
December 2004

A479 Jarmo Malinen
Conservativity of Time-Flow Invertible and Boundary Control Systems
December 2004

A478 Niko Marola
Moser's Method for minimizers on metric measure spaces
October 2004

A477 Tuomo T. Kuusi
Moser's Method for a Nonlinear Parabolic Equation
October 2004

A476 Dario Gasbarra , Esko Valkeila , Lioudmila Vostrikova
Enlargement of filtration and additional information in pricing models: a Bayesian approach
October 2004

A475 Iivo Vehviläinen
Applyining mathematical finance tools to the competitive Nordic electricity market
October 2004

HELSINKI UNIVERSITY OF TECHNOLOGY INSTITUTE OF MATHEMATICS
RESEARCH REPORTS

The list of reports is continued inside. Electronical versions of the reports are available at *http://math.tkk.fi/reports/* .

A490 Anders Björn , Niko Marola
Moser iteration for (quasi)minimizers on metric spaces
September 2005

A489 Sampsa Pursiainen
A coarse-to-fine strategy for maximum a posteriori estimation in limited-angle computerized tomography
September 2005

A487 Ville Turunen
Differentiability in locally compact metric spaces
May 2005

A486 Hanna Pikkarainen
A Mathematical Model for Electrical Impedance Process Tomography
April 2005

A485 Sampsa Pursiainen
Bayesian approach to detection of anomalies in electrical impedance tomography
April 2005

ISBN 951-22-7917-7
ISSN 0784-3143