



Centrum voor Wiskunde en Informatica

REPORTRAPPORT

PNA

Probability, Networks and Algorithms



Probability, Networks and Algorithms

Large buffer asymptotics for generalized processor sharing queues with Gaussian inputs

K. Dębicki, M.J.G. van Uitert

REPORT PNA-E0507 SEPTEMBER 2005

CWI is the National Research Institute for Mathematics and Computer Science. It is sponsored by the Netherlands Organization for Scientific Research (NWO).

CWI is a founding member of ERCIM, the European Research Consortium for Informatics and Mathematics.

CWI's research has a theme-oriented structure and is grouped into four clusters. Listed below are the names of the clusters and in parentheses their acronyms.

Probability, Networks and Algorithms (PNA)

Software Engineering (SEN)

Modelling, Analysis and Simulation (MAS)

Information Systems (INS)

Copyright © 2005, Stichting Centrum voor Wiskunde en Informatica

P.O. Box 94079, 1090 GB Amsterdam (NL)

Kruislaan 413, 1098 SJ Amsterdam (NL)

Telephone +31 20 592 9333

Telefax +31 20 592 4199

ISSN 1386-3711

Large buffer asymptotics for generalized processor sharing queues with Gaussian inputs

ABSTRACT

In this paper we derive large-buffer asymptotics for a two-class Generalized Processor Sharing (GPS) model. We assume both classes to have Gaussian characteristics. We distinguish three cases depending on whether the GPS weights are above or below the average rate at which traffic is sent. First, we calculate exact asymptotic upper and lower bounds, then we calculate the logarithmic asymptotics, and finally we show that the decay rates of the upper and lower bound match. We apply our results to two special Gaussian models: the integrated Gaussian process and the fractional Brownian motion. Finally we derive the logarithmic large-buffer asymptotics for the case where a Gaussian flow interacts with an on-off flow.

2000 Mathematics Subject Classification: 60K25

Keywords and Phrases: large-buffer asymptotics; Gaussian traffic; Generalized Processor Sharing; communication networks; differentiated services

Note: The second author is grateful to CWI for letting her finish this work.

Large buffer asymptotics for Generalized Processor Sharing queues with Gaussian inputs

Krzysztof Dębicki and Miranda van Uitert *

Abstract

In this paper we derive large-buffer asymptotics for a two-class Generalized Processor Sharing (GPS) model. We assume both classes to have Gaussian characteristics. We distinguish three cases depending on whether the GPS weights are above or below the average rate at which traffic is sent. First, we calculate exact asymptotic upper and lower bounds, then we calculate the logarithmic asymptotics, and finally we show that the decay rates of the upper and lower bound match. We apply our results to two special Gaussian models: the integrated Gaussian process and the fractional Brownian motion. Finally we derive the logarithmic large-buffer asymptotics for the case where a Gaussian flow interacts with an on-off flow.

Key words: large-buffer asymptotics – Gaussian traffic – Generalized Processor Sharing – communication networks – differentiated services

1 Introduction

In this paper we analyze the well-known Generalized Processor Sharing (GPS) mechanism. GPS has received a lot of attention over the last decade, mainly due to its isolation property. When the link capacity in a network is shared between two flows using GPS, then both classes are offered a fixed amount of the capacity. On top of this guaranteed capacity they can enjoy the capacity that is not needed by the other flow.

We choose to work with Gaussian processes in this paper. As both long-range dependent and short-range dependent characteristics can be modeled through Gaussian processes, they

*The first author is with the Mathematical Institute, University of Wrocław, pl. Grunwaldzki 2/4, 50-384 Wrocław, Poland. The second author is grateful to CWI, P.O. Box 94079, 1090 GB Amsterdam, the Netherlands, for letting her finish this work. Part of the work was done while the first author was also at CWI. Email: debicki@math.uni.wroc.pl.

provide a versatile framework for modeling a broad variety of traffic types. Moreover, the Gaussian model arises as a limiting process when considering the superposition of many on-off sources, see e.g. Taquu *et al.* [25] and Dębicki & Palmowski [8]. We also refer to Fraleigh *et al.* [11] and Kilpi & Norros [12] who give arguments in favor of Gaussian models when aggregating flows, possibly using TCP, in networks.

Until so far many authors have worked on large-buffer asymptotics for GPS. We mention the work by Borst *et al.* [1, 3, 4, 27] and Kotopoulos *et al.* [13] who derived results for heavy-tailed input processes, and Zhang [28, 29] who established results for light-tailed inputs. More recently, Mannersalo & Norros [20] and Mandjes & Van Uitert [18] have considered many-sources asymptotics for GPS with Gaussian inputs. We refer to Van Uitert [26] for a complete overview of the work that has been done on GPS. Most recent work on GPS involves calculating the optimal guaranteed rate [17] and many-sources delay asymptotics [16] with Gaussian inputs. However, results for large-buffer asymptotics with Gaussian inputs have remained unknown until this paper.

In this paper we first consider the asymptotic behavior of two Gaussian flows that share the network capacity using GPS. We distinguish three scenarios, depending on whether one of the flows sends on average above or below its guaranteed capacity. First calculating the exact asymptotic upper and lower bounds, and then taking the logarithm, we obtain the decay rate of the probability that one of the flows has a large backlog. We apply these results to two special cases of Gaussian processes: integrated Gaussian process, see e.g. [8, 9] and fractional Brownian motion, see for instance [22]. In the last part of this paper we derive the logarithmic large-buffer asymptotics for the situation where a Gaussian flow interacts with an on-off flow under the GPS discipline.

This paper is organized as follows. In Section 2 we describe the GPS mechanism in more detail, give the necessary notation, and discuss the specific details of the input models. We present the main results, without proof, for two Gaussian flows sharing a GPS queue in Section 3. Important applications are discussed in Section 4. Then in Section 5 we give the proofs of the results in Section 3. In Section 6 we derive the asymptotic behavior for a Gaussian flow sharing a GPS queue with an on-off process.

2 Model and preliminaries

2.1 Generalized Processor Sharing

We consider a GPS server with service rate c , which is fed by two classes that both have their own queue. Both classes are assigned a positive weight ϕ_i , $i = 1, 2$, and we assume without loss of generality that $\phi_1 + \phi_2 = 1$. If both classes are backlogged, then class i is served at

rate $\phi_i c$, which we call the *guaranteed rate* of class i . If one of the classes has no backlog then the other class gets the excess service rate (note that a class can send traffic at a rate smaller than its guaranteed rate and is not backlogged). We denote by $Q_i(t), t \in \mathbb{R}$ the stationary workload process of class i at time t , and let $Q_i \equiv Q_i(0)$. The probability of interest is

$$\mathbb{P}(Q_1 > u), \text{ with } u \rightarrow \infty. \quad (1)$$

Note that the analysis for Q_2 is completely symmetric. We denote by $A_i(s, t)$ the amount of traffic that is put into class i during the interval $(s, t]$, with $s < t$ and $s, t \in \mathbb{R}$. The amount of service that is obtained by class i in the interval $(s, t]$ is defined by $B_i(s, t)$. The following identity is straightforward,

$$Q_i(t) = Q_i(s) + A_i(s, t) - B_i(s, t), \quad \forall s < t, \quad s, t \in \mathbb{R}. \quad (2)$$

Following Reich's representation [24] we can also write

$$Q_i(t) = \sup_{s \leq t} \{A_i(s, t) - B_i(s, t)\},$$

where the optimizing s denotes the beginning of the busy period that contains time t . Similar to [18, 19], we use this result to obtain

$$\begin{aligned} Q_1 \equiv Q_1(0) &= Q_1(0) + Q_2(0) - Q_2(0) \\ &= \sup_{s \geq 0} \left\{ A_1(-s, 0) + A_2(-s, 0) - cs - \sup_{t \in [0, s)} \{A_2(-t, 0) - B_2(-t, 0)\} \right\}, \end{aligned}$$

with the negative of the optimizing s (respectively t) denoting the beginning of the busy period in the total queue (respectively queue 2) containing time epoch 0. Obviously,

$$Q_1(0) \geq \sup_{s \geq 0} \left\{ A_1(-s, 0) + A_2(-s, 0) - cs - \sup_{t \in [0, s)} \{A_2(-t, 0) - \phi_2 ct\} \right\}. \quad (3)$$

We use the following additional notation. We denote by Q_i^c the stationary workload of flow i served in isolation at rate c . For any two real functions $g(\cdot)$ and $h(\cdot)$, we use $g(u) \sim h(u)$ to denote

$$\lim_{u \rightarrow \infty} \frac{g(u)}{h(u)} = 1, \text{ or equivalently, } g(u) = h(u)(1 + o(1)) \text{ as } u \rightarrow \infty.$$

Similarly we use \gtrsim and \lesssim to denote an asymptotic lower respectively upper bound for $u \rightarrow \infty$. We use $\stackrel{d}{=}$ to denote equality in distribution.

2.2 Input processes

We assume the inputs of both classes to be Gaussian processes with stationary increments. When we use the notation i in the remainder of this section, we mean $i = 1, 2$. We define $\mu_i(t-s) := \mathbb{E}[A_i(s, t)]$ to be the mean amount of traffic that arrived in the interval $(s, t]$. The variance over an interval $(s, t]$ is defined by $v_{A_i}(t-s) := \text{Var}[A_i(s, t)]$. For convenience we define $A_i(t) := A_i(t, 0)$ for $t < 0$ and analogously $A_i(t) := A_i(0, t)$ for $t \geq 0$. We use the additional notation $\bar{A}_i(t) := A_i(t) - \mu_i t$ for the centered version, and define the total input process by $A_i := \{A_i(t), t \in \mathbb{R}\}$. We make the following assumptions on the variance function $v_{A_i}(\cdot)$:

- C1** $v_{A_i}(t) \in C([0, \infty))$ is ultimately strictly increasing;
- C2** $v_{A_i}(t)$ is regularly varying at 0 with index $\beta_i \in (0, 2]$ and $v_{A_i}(t)$ is regularly varying at ∞ with index $\alpha_i \in (0, 2)$;
- C3** $v_{A_i}(t)$ is ultimately convex.

When referring to either **C1**, **C2** or **C3** for a Gaussian process in the remainder of this paper, we mean that the variance function of this process satisfies the corresponding regularity condition as defined above. In particular, given the applications in Gaussian fluid models, we focus on the following special cases of the input processes.

- **Integrated Gaussian (IG)**: $A_i(t) = \int_0^t (Z_i(s) + \mu_i) ds$, where $Z_i(t)$ is a stationary centered Gaussian process with continuous covariance function $R_i(t) = \text{Cov}(Z_i(s+t), Z_i(s))$ such that $R_i(t) > 0$ for each $t \geq 0$ and $\mu_i > 0$.
- **Fractional Brownian motion (FBM)**: $A_i(t) = B_{H_i}(t) + \mu_i t$, where $B_{H_i}(t)$ is the fractional Brownian motion with Hurst parameter $H_i \geq 1/2$. In particular, for $H_i = 1/2$ we have that $B_{1/2}(t) = W(t)$ is a standard Brownian motion.

The relevance of **FBM** input is discussed in e.g. [22]. The importance of **IG** input in the theory of fluid models is shown in e.g. [8] and [9].

3 Main results

In this section we present the main results of the paper, that is, the logarithmic asymptotics of the probability that queue 1 reaches a large workload (1). Since most of the proofs is rather technical, we first present the results without proof in this section. The proofs will be given in Section 5, after the next section where we apply our results to two important special cases: **IG** and **FBM**. Although the asymptotic behavior in Theorem 3.1 is equal to that in Theorem 3.3, we present it separately as it requires a different approach in the proof. Throughout the

paper we assume that flow 1 sends on average at a rate smaller than the guaranteed rate, i.e., $\mu_1 < \phi_1 c$, and we say that flow 1 is in *underload*. For flow 2 we distinguish between underload, i.e., $\mu_2 < \phi_2 c$, and *overload*, i.e., $\mu_2 > \phi_2 c$. The following scenarios, corresponding with the theorems are considered:

S1 (flow 2 in overload): $\mu_2 > \phi_2 c$;

S2 (flow 2 in underload and flow 1 dominant): $\mu_2 < \phi_2 c$ and $\lim_{t \rightarrow \infty} \frac{v_2(t)}{v_1(t)} = 0$;

S3 (flow 2 in underload and flow 2 dominant): $\mu_2 < \phi_2 c$ and $\lim_{t \rightarrow \infty} \frac{v_1(t)}{v_2(t)} = 0$.

We remark that we do not consider the scenario with flow 1 in overload and flow 2 in underload. For this scenario a completely different approach is needed, which is not within the scope of this paper. To ensure the stability of the system in the remainder of the paper we tacitly assume that

$$\mu_1 + \mu_2 < c; \quad \mu_1 \geq 0, \mu_2 \geq 0.$$

We start the analysis with scenario **S1**. In this regime, flow 2 is not stable if it would only obtain service rate $\phi_2 c$; it needs the remaining service capacity of flow 1 for stability. The logarithmic asymptotics for this regime are as follows.

Theorem 3.1 (S1) *If A_1 and A_2 satisfy C1-C2 then*

$$\begin{aligned} \log \mathbb{P}(Q_1 > u) &\sim \log \mathbb{P}(Q_1^{\phi_1 c} > u) \\ &\sim -\frac{1}{2}(\phi_1 c - \mu_1)^{\alpha_1} \left(\frac{\alpha_1}{2 - \alpha_1}\right)^{-\alpha_1} \left(\frac{2}{2 - \alpha_1}\right)^2 \frac{u^2}{v_{A_1}(u)}. \end{aligned}$$

Intuitively, the result states that in order for flow 1 to have a large workload, it has to send at a rate higher than its guaranteed rate $\phi_1 c$. As $\phi_1 c$ is larger than the rate at which flow 1 sends on average, we say that flow 1 has to show abnormal behavior. That is, on average flow 1 is likely to send at a smaller rate. Note that the behavior of flow 2 does not show up in the asymptotics. That is because flow 2 already uses its guaranteed capacity $\phi_2 c$ just by sending on its average rate μ_2 .

The results for scenario **S2** are different. In this scenario flow 2 remains stable, even if it only receives service at its guaranteed rate $\phi_2 c$. Recall that in this regime, the variance function of flow 1 dominates that of flow 2. The logarithmic asymptotics for this regime are as follows.

Theorem 3.2 (S2) *If A_1 and A_2 satisfy C1-C2 then*

$$\begin{aligned} \log \mathbb{P}(Q_1 > u) &\sim \log \mathbb{P}(Q_1^{c - \mu_2} > u) \\ &\sim -\frac{1}{2}(c - \mu)^{\alpha_1} \left(\frac{\alpha_1}{2 - \alpha_1}\right)^{-\alpha_1} \left(\frac{2}{2 - \alpha_1}\right)^2 \frac{u^2}{v_{A_1}(u)}. \end{aligned}$$

The result is a so-called reduced-load equivalent, as the workload asymptotics of flow 1 are similar to the workload asymptotics of this flow if it were served in isolation at service rate $c - \mu_2$. Intuitively two scenarios may lead to a large workload of flow 1 in this regime. One is the reduced-load equivalence, where flow 2 behaves normal in the sense that it sends on average at rate μ_2 , and flow 1 takes all capacity it can get. In the other scenario, flow 2 behaves ‘abnormal’ as well. That is, flow 2 takes its reserved capacity $\phi_2 c$ during the epoch that flow 1 needs to build up a large workload with service rate $\phi_1 c$. Because of the assumption on the variance functions, the reduced-load equivalence result turns out to be the most likely scenario in this regime. Not surprisingly, the other scenario is most likely in the regime **S3**, where we assume the variance function of flow 2 to dominate that of flow 1.

For scenario **S3** the logarithmic asymptotics are as follows.

Theorem 3.3 (S3) *If A_1 and A_2 satisfy **C1-C3** then*

$$\begin{aligned} \log \mathbb{P}(Q_1 > u) &\sim \log \mathbb{P}(Q_1^{\phi_1 c} > u) \\ &\sim -\frac{1}{2}(\phi_1 c - \mu_1)^{\alpha_1} \left(\frac{\alpha_1}{2 - \alpha_1}\right)^{-\alpha_1} \left(\frac{2}{2 - \alpha_1}\right)^2 \frac{u^2}{v_{A_1}(u)}. \end{aligned}$$

Qualitatively the result is the same as that of Theorem 3.1, covering regime **S1** where flow 2 sends on average at a rate higher than $\phi_2 c$. Again the workload of flow 1 reaches a large level only because it does something extraordinary itself.

4 Examples

In this section we analyze the asymptotic behavior of the probability that Q_1 gets large for two important applications: **IG** and **FBM** processes.

4.1 Integrated Gaussian input

When a large number of sources shares a queue then the integrated Gaussian process can be used to model the accumulated input. See for instance Dębicki & Palmowski [8] and Kulkarni & Rolski [14], who show that **IG** appears as the limit approximation, in a heavy traffic environment, of a superposition of a large number of integrated on-off processes. In the case of single fluid queueing models, this class of processes was intensively investigated by many authors; see e.g. [7, 14] and references therein.

In Proposition 4.1 we present the results for scenarios **S1**, **S2** and **S3** when the **IG** process exhibits a short-range dependent structure. Results for long-range dependent **IG** processes are then given in Proposition 4.2.

Proposition 4.1 *Let A_2 satisfy **C1-C3** and $A_1(t) = \int_0^t Z(s) + \mu_1 ds$ with $Z(t)$ a centered stationary Gaussian process with continuous covariance function $R(t)$ such that $R(t) > 0$ for all $t \geq 0$ and $\int_0^\infty R(t)dt < \infty$, then*

$$[\mathbf{S1}] \quad \log \mathbb{P}(Q_1 > u) \sim -\frac{\phi_1 c - \mu_1}{\int_0^\infty R(t)dt} u;$$

$$[\mathbf{S2}] \quad \log \mathbb{P}(Q_1 > u) \sim -\frac{c - \mu}{\int_0^\infty R(t)dt} u;$$

$$[\mathbf{S3}] \quad \log \mathbb{P}(Q_1 > u) \sim -\frac{\phi_1 c - \mu_1}{\int_0^\infty R(t)dt} u.$$

Proof Observe that

$$v_{A_1}(t) = 2 \int_0^t ds \int_0^s R(v)dv,$$

meaning that $v_{A_1}(t)$ satisfies **C1-C2** with $\alpha_1 = 1$ and $\beta_1 = 2$. Moreover **C3** is fulfilled due to the nonnegativity of $R(\cdot)$. Applying Theorems 3.2, 3.3 and 3.1 respectively completes the proof. \square

Proposition 4.2 *Let A_2 satisfy **C1-C3** and $A_1(t) = \int_0^t Z(s)ds + \mu_1 t$ with $Z(t)$ a centered stationary Gaussian process with continuous covariance function $R(t)$ such that $R(t) > 0$ for all $t \geq 0$ and $R(t) = l(t)t^{\alpha-2}$ for $\alpha \in (1, 2)$ with $l(\cdot)$ slowly varying at ∞ , then*

$$[\mathbf{S1}] \quad \log \mathbb{P}(Q_1 > u) \sim -\frac{1}{2}(\phi_1 c - \mu_1)^\alpha \left(\frac{\alpha}{2-\alpha}\right)^{-\alpha} \left(\frac{2}{2-\alpha}\right)^2 \frac{\alpha(\alpha-1)}{2} \frac{u^{2-\alpha}}{l(u)};$$

$$[\mathbf{S2}] \quad \log \mathbb{P}(Q_1 > u) \sim -\frac{1}{2}(c - \mu)^\alpha \left(\frac{\alpha}{2-\alpha}\right)^{-\alpha} \left(\frac{2}{2-\alpha}\right)^2 \frac{\alpha(\alpha-1)}{2} \frac{u^{2-\alpha}}{l(u)};$$

$$[\mathbf{S3}] \quad \log \mathbb{P}(Q_1 > u) \sim -\frac{1}{2}(\phi_1 c - \mu_1)^\alpha \left(\frac{\alpha}{2-\alpha}\right)^{-\alpha} \left(\frac{2}{2-\alpha}\right)^2 \frac{\alpha(\alpha-1)}{2} \frac{u^{2-\alpha}}{l(u)}.$$

Proof The proof is analogous to that of Proposition 4.1. \square

4.2 Fractional Brownian motion input

In this section we focus on the case, where the input process of the first queue is given by a fractional Brownian motion with Hurst parameter $H \geq 1/2$. The idea of using **FBM** as input process was motivated by some measurements of internet traffic, see e.g. [15], where the presence of *long-range dependency* and *self similarity* was statistically verified. The theoretical justification of **FBM** as an appropriate approximation for input traffic can be found in Taqqu *et al.* [25] and Mikosch *et al.* [21].

In the following proposition we give the results for the scenarios **S1**, **S2** and **S3** when flow 1 is an **FBM** process.

Proposition 4.3 *Let A_2 satisfy **C1-C3** and $A_1(t) = B_H(t) + \mu_1 t$, an FBM with $H \geq 1/2$, then*

$$[\mathbf{S1}] \quad \log \mathbb{P}(Q_1 > u) \sim -\frac{1}{2}(\phi_1 c - \mu_1)^{2H} \left(\frac{H}{1-H}\right)^{-2H} \left(\frac{1}{1-H}\right)^2 u^{2-2H};$$

$$[\mathbf{S2}] \quad \log \mathbb{P}(Q_1 > u) \sim -\frac{1}{2}(c - \mu)^{2H} \left(\frac{H}{1-H}\right)^{-2H} \left(\frac{1}{1-H}\right)^2 u^{2-2H};$$

$$[\mathbf{S3}] \quad \log \mathbb{P}(Q_1 > u) \sim -\frac{1}{2}(\phi_1 c - \mu_1)^{2H} \left(\frac{H}{1-H}\right)^{-2H} \left(\frac{1}{1-H}\right)^2 u^{2-2H}.$$

Proof The proof is analogous to that of Proposition 4.1. □

5 Proofs

In this section we give the proofs of Theorems 3.1, 3.2 and 3.3. Before presenting the proofs we need some technical lemmas.

We use the notation $X := \{X(t), t \in \mathbb{R}\}$ to denote a centered Gaussian process. The following lemma, which was proved in [2], gives an upper bound for the supremum of a centered Gaussian process over a finite interval. The lemma will be used in the proofs of Theorems 3.1 and 3.2.

Lemma 5.1 *Let X be a centered Gaussian stochastic process with stationary increments. If X satisfies conditions **C1-C2**, then there exist constants $K > 0$ and $q > 0$ such that for t large enough and every $u > q\sqrt{v_X(t)}$,*

$$\mathbb{P}\left(\sup_{s \in [0, t]} X(s) > u\right) \leq K e^{-\frac{1}{2}\left(\frac{u}{\sqrt{v_X(t)}} - q\right)^2}. \quad (4)$$

In the next lemma the logarithmic asymptotics are given for the supremum of a centered Gaussian process over an infinite interval, i.e., for the stationary version of a queue with input process X and service rate c . We use this lemma in the proofs of Theorem 3.1, 3.2 and 3.3.

Lemma 5.2 *Let X be a centered Gaussian stochastic process with stationary increments. If X satisfies conditions **C1-C2** with regularity indexes α, β respectively, then, for $c > 0$,*

$$(i) \quad \log \mathbb{P}\left(\sup_{t \geq 0} \{X(t) - ct\} > u\right) \sim -\frac{1}{2}c^\alpha \left(\frac{\alpha}{2-\alpha}\right)^{-\alpha} \left(\frac{2}{2-\alpha}\right)^2 \frac{u^2}{v_X(u)};$$

$$(ii) \quad \log \mathbb{P}\left(\sup_{t \geq 0} \{X(t) - ct\} > u\right) \sim \log \mathbb{P}\left(\mathcal{N} > \frac{u + ct_u}{\sqrt{v_X(t_u)}}\right),$$

for $t_u = \arg \inf_{t \geq 0} \frac{(u+ct)^2}{2v_X(t)}$. Moreover t_u is regularly varying at ∞ with index 1, as $u \rightarrow \infty$.

Proof The proof of (i) straightforwardly follows from Proposition 1 in [10].

In order to prove (ii) it is enough to combine Theorem 2.1 in [6] with the observation that

$$\log \mathbb{P} \left(\mathcal{N} > \frac{u + ct_u}{\sqrt{v_X(t_u)}} \right) \sim \inf_{t \geq 0} \frac{(u + ct)^2}{2v_X(t)}.$$

Moreover the combination of Theorem 1 in [10] with (5.14) and (5.15) in [31] gives that $t_u \sim \frac{1-\alpha}{c} u$ as $u \rightarrow \infty$ which shows that t_u is regularly varying at ∞ with index 1. \square

The following lemma is used in the lower bound of Theorem 3.3. It provides an asymptotic lower bound for the infimum of a centered Gaussian process over a finite interval. In terms of queues we give an asymptotic lower bound for the probability that the queue is backlogged during a finite interval.

Lemma 5.3 *If X satisfies conditions C1-C3, then for each $\epsilon \in (0, 1)$,*

$$\log \mathbb{P} \left(\inf_{t \in [\epsilon u, u]} \{X(t) - ct\} > 0 \right) \gtrsim -\frac{c^2 u^2}{2v_X(\epsilon u)}.$$

Proof Note that for sufficiently large u

$$\begin{aligned} \mathbb{P} \left(\inf_{t \in [\epsilon u, u]} \{X(t) - ct\} > 0 \right) &\geq \mathbb{P} \left(\inf_{t \in [\epsilon u, u]} X(t) > cu \right) \\ &= \mathbb{P} \left(\sup_{t \in [\epsilon u, u]} X(t) < -cu \right) \\ &\geq \mathbb{P} \left(\sup_{t \in [\epsilon u, u]} W(v_X(t)) < -cu \right) \\ &= \mathbb{P} \left(\sup_{t \in [v_X(\epsilon u), v_X(u)]} W(t) < -cu \right), \end{aligned} \tag{5}$$

where W is a standard Brownian motion, and (5) follows from Slepian's inequality (see e.g. Theorem C.1 in [23]) combined with the fact that $v_X(t)$ is convex. Because W has independent increments, we obtain

$$\begin{aligned} &\mathbb{P} \left(\sup_{t \in [v_X(\epsilon u), v_X(u)]} W(t) < -cu \right) \\ &= \mathbb{P} \left(W(v_X(\epsilon u)) + \sup_{t \in [v_X(\epsilon u), v_X(u)]} \{W(t) - W(v_X(\epsilon u))\} < -cu \right) \\ &= \mathbb{P} \left(W_1(v_X(\epsilon u)) + \sup_{t \in [0, v_X(u) - v_X(\epsilon u)]} W_2(t) < -cu \right), \end{aligned}$$

where both W_1 and W_2 are independent copies of W . For $\gamma > 0$, such that $1 > \gamma > \alpha/2$,

$$\begin{aligned}
& \mathbb{P} \left(W_1(v_X(\epsilon u)) + \sup_{t \in [0, v_X(u) - v_X(\epsilon u)]} W_2(t) < -cu \right) \\
& \geq \mathbb{P} \left(W_1(v_X(\epsilon u)) + \sup_{t \in [0, v_X(u) - v_X(\epsilon u)]} W_2(t) < -cu; \sup_{t \in [0, v_X(u) - v_X(\epsilon u)]} W_2(t) < u^\gamma \right) \\
& \geq \mathbb{P}(W_1(v_X(\epsilon u)) + u^\gamma < -cu) \mathbb{P} \left(\sup_{t \in [0, v_X(u) - v_X(\epsilon u)]} W_2(t) < u^\gamma \right) \\
& = \mathbb{P} \left(\mathcal{N} > \frac{cu + u^\gamma}{\sqrt{v_X(\epsilon u)}} \right) \left(1 - 2\mathbb{P} \left(\mathcal{N} > \frac{u^\gamma}{\sqrt{v_X(u) - v_X(Au)}} \right) \right), \tag{6}
\end{aligned}$$

where (6) follows from the fact that $\mathbb{P}(\sup_{t \in [0, T]} W(t) > x) = 2\mathbb{P}(\mathcal{N} > x/\sqrt{T})$. Finally, combining (6) with the fact that

$$\log \mathbb{P} \left(\mathcal{N} > \frac{cu + u^\gamma}{\sqrt{v_X(\epsilon u)}} \right) \sim -\frac{c^2 u^2}{2v_X(\epsilon u)},$$

and

$$\mathbb{P} \left(\mathcal{N} > \frac{u^\gamma}{\sqrt{v_X(u) - v_X(\epsilon u)}} \right) \rightarrow 0, \text{ as } u \rightarrow \infty,$$

the proof is completed. \square

5.1 Proof of Theorem 3.1

Upper bound. This follows straightforwardly from the fact that

$$\mathbb{P}(Q_1 > u) \leq \mathbb{P}(Q_1^{\phi_1 c} > u) = \mathbb{P} \left(\sup_{t \geq 0} (A_1(t) - \phi_1 ct) > u \right)$$

and Lemma 5.2 (i).

Lower bound. Defining

$$t_u := \arg \inf_{t \geq 0} \frac{(u + (\phi_1 c - \mu_1)t)^2}{2v_1(t)},$$

we use (3) to obtain

$$\begin{aligned}
& \mathbb{P}(Q_1 > u) \\
& \geq \mathbb{P}\left(A_1(-t_u, 0) + A_2(-t_u, 0) - ct_u - \sup_{t \in [0, t_u]} \{A_2(-t, 0) - \phi_2 ct\} > u\right) \\
& = \mathbb{P}\left(\bar{A}_1(t_u) + \bar{A}_2(t_u) - (c - \mu)t_u - \sup_{t \in [0, t_u]} \{\bar{A}_2(t) + (\mu_2 - \phi_2 c)t\} > u\right) \\
& \geq \mathbb{P}\left(\bar{A}_1(t_u) - (c - \mu)t_u - (\mu_2 - \phi_2 c)t_u - \sup_{t \in [0, t_u]} \{\bar{A}_2(t) - \bar{A}_2(t_u)\} > u\right) \\
& = \mathbb{P}\left(\bar{A}_1(t_u) - (\phi_1 c - \mu_1)t_u - \sup_{t \in [0, t_u]} \bar{A}_2(t) > u\right), \tag{7}
\end{aligned}$$

where (7) follows from the fact that $\sup_{t \in [0, t_u]} \{\bar{A}_2(t) - \bar{A}_2(t_u)\} \stackrel{d}{=} \sup_{t \in [0, t_u]} \bar{A}_2(t)$. Adding an additional constraint, the probability in line (7) is bounded below by

$$\begin{aligned}
& \mathbb{P}\left(\bar{A}_1(t_u) + (\mu_1 - \phi_1 c)t_u - \sup_{t \in [0, t_u]} \bar{A}_2(t) > u, \sup_{t \in [0, t_u]} \bar{A}_2(t) < (v_{A_2}(t_u))^{\frac{1}{2} + \epsilon}\right) \\
& \geq \mathbb{P}\left(\bar{A}_1(t_u) + (\mu_1 - \phi_1 c)t_u - (v_{A_2}(t_u))^{\frac{1}{2} + \epsilon} > u\right) \mathbb{P}\left(\sup_{t \in [0, t_u]} \bar{A}_2(t) < (v_{A_2}(t_u))^{\frac{1}{2} + \epsilon}\right),
\end{aligned}$$

with $\epsilon > 0$ such that $(\frac{1}{2} + \epsilon)\alpha_2 < 1$. Applying Lemma 5.1 we obtain

$$\mathbb{P}(Q_1 > u) \geq \mathbb{P}\left(\mathcal{N} > \frac{u + (\phi_1 c - \mu_1)t_u + (v_{A_2}(t_u))^{\frac{1}{2} + \epsilon}}{\sqrt{v_{A_1}(t_u)}}\right) \left(1 - Ke^{-\kappa((v_{A_2}(t_u))^\epsilon - q)^2}\right),$$

with \mathcal{N} a standard Gaussian random variable, and K, κ and q some positive constants. Because $t_u \rightarrow \infty$ as $u \rightarrow \infty$ and $(\frac{1}{2} + \epsilon)\alpha_2 < 1$, we obtain consecutively, as $u \rightarrow \infty$,

$$\log \mathbb{P}(Q_1 > u) \gtrsim -\frac{\left(u + (\phi_1 c - \mu_1)t_u + (v_{A_2}(t_u))^{\frac{1}{2} + \epsilon}\right)^2}{2v_{A_1}(t_u)} \sim -\frac{(u + (\phi_1 c - \mu_1)t_u)^2}{2v_{A_1}(t_u)}.$$

Hence combining the fact that

$$-\frac{(u + (\phi_1 c - \mu_1)t_u)^2}{2v_{A_1}(t_u)} \sim \log \mathbb{P}\left(\mathcal{N} > \frac{u + (\phi_1 c - \mu_1)t_u}{\sqrt{v_X(t_u)}}\right)$$

with Lemma 5.2 we obtain

$$\log \mathbb{P}(Q_1 > u) \gtrsim -\frac{1}{2}(\phi_1 c - \mu_1)^{\alpha_1} \left(\frac{\alpha_1}{2 - \alpha_1}\right)^{-\alpha_1} \left(\frac{2}{2 - \alpha_1}\right)^2 \frac{u^2}{v_{A_1}(u)},$$

which completes the proof. \square

5.2 Proof of Theorem 3.2

Upper bound. Note that

$$\mathbb{P}(Q_1 > u) \leq \mathbb{P}(Q_1 + Q_2 > u) = \mathbb{P}\left(\sup_{t \geq 0} A_1(t) + A_2(t) - ct > u\right).$$

Then applying Lemma 5.2 (i) and using that $v_{A_1+A_2}(t) = v_{A_1}(t) + v_{A_2}(t)$ is regularly varying at ∞ with index α_1 , we obtain

$$\begin{aligned} \log \mathbb{P}(Q_1 > u) &\lesssim -\frac{1}{2}(c - \mu)^{\alpha_1} \left(\frac{\alpha_1}{2 - \alpha_1}\right)^{-\alpha_1} \left(\frac{2}{2 - \alpha_1}\right)^2 \frac{u^2}{v_{A_1}(u) + v_{A_2}(u)} \\ &\sim -\frac{1}{2}(c - \mu)^{\alpha_1} \left(\frac{\alpha_1}{2 - \alpha_1}\right)^{-\alpha_1} \left(\frac{2}{2 - \alpha_1}\right)^2 \frac{u^2}{v_{A_1}(u)}, \end{aligned}$$

with \sim due to S2.

Lower bound. Defining (note that this t_u differs in a subtle way from the t_u defined in the previous section)

$$t_u := \arg \inf_{t \geq 0} \frac{(u + (c - \mu)t)^2}{2v_{A_1}(t)},$$

we obtain, using (3),

$$\begin{aligned} \mathbb{P}(Q_1 > u) &= \mathbb{P}(Q_1(0) > u) \\ &\geq \mathbb{P}\left(A_1(-t_u, 0) + A_2(-t_u, 0) - ct_u - \sup_{t \in [0, t_u]} \{A_2(-t, 0) - \phi_2 ct\} > u\right) \\ &= \mathbb{P}\left(\bar{A}_1(t_u) - (\phi_1 c - \mu_1)t_u - \sup_{t \in [0, t_u]} \{\bar{A}_2(t) - \bar{A}_2(t_u) + (\phi_2 c - \mu_2)(t_u - t)\} > u\right) \\ &= \mathbb{P}\left(\bar{A}_1(t_u) - (\phi_1 c - \mu_1)t_u - \sup_{t \in [0, t_u]} \{\bar{A}_2(t) + (\phi_2 c - \mu_2)t\} > u\right) \\ &\geq \mathbb{P}\left(\bar{A}_1(t_u) - (c - \mu)t_u - \sup_{t \in [0, t_u]} \bar{A}_2(t) > u\right) \\ &\geq \mathbb{P}\left(\bar{A}_1(t_u) - (c - \mu)t_u - (v_2(t_u))^{\frac{1}{2} + \epsilon} > u\right) \mathbb{P}\left(\sup_{t \in [0, t_u]} \bar{A}_2(t) < (v_{A_2}(t_u))^{\frac{1}{2} + \epsilon}\right) \end{aligned}$$

for some $\epsilon > 0$ such that $(\frac{1}{2} + \epsilon)\alpha_2 < 1$. The remainder of the proof is analogous to that of the lower bound of Theorem 3.1. \square

5.3 Proof of Theorem 3.3

Upper bound. Goes line by line the same as the corresponding part of the proof of Theorem 3.1.

Lower bound. Let

$$t_u := \arg \inf_{t \geq 0} \frac{(u + (\phi_1 c - \mu_1)t)^2}{2v_1(t)},$$

which is the same definition as in the proof of Theorem 3.1. Following (2) we have

$$\begin{aligned} \mathbb{P}(Q_1 > u) &= \mathbb{P}(Q_1(t_u) > u) \\ &= \mathbb{P}(Q_1(0) + A_1(0, t_u) - B_1(0, t_u) > u) \\ &\geq \mathbb{P}(A_1(0, t_u) - B_1(0, t_u) > u). \end{aligned} \tag{8}$$

Hence for $\varepsilon \in (0, 1)$

$$\begin{aligned} \mathbb{P}(Q_1 > u) &\geq \\ &\geq \mathbb{P}\left(A_1(0, t_u) - B_1(0, t_u) > u; \inf_{t \in [\varepsilon t_u, t_u]} Q_2(t) > 0\right) \\ &= \mathbb{P}\left(A_1(0, t_u) - B_1(0, t_u) > u; \inf_{t \in [\varepsilon t_u, t_u]} Q_2(t) > 0; B_2(0, \varepsilon t_u) \geq \phi_2 c(1 - \varepsilon)t_u\right) \\ &= \mathbb{P}\left(A_1(0, t_u) - B_1(0, \varepsilon t_u) - B_1(\varepsilon t_u, t_u) > u; \inf_{t \in [\varepsilon t_u, t_u]} Q_2(t) > 0; B_2(0, \varepsilon t_u) \geq \phi_2 c(1 - \varepsilon)t_u\right) \\ &\geq \mathbb{P}\left(A_1(0, t_u) - c\varepsilon t_u - \phi_1 c(1 - \varepsilon)t_u > u; \inf_{t \in [\varepsilon t_u, t_u]} Q_2(t) > 0\right) \end{aligned} \tag{9}$$

$$\geq \mathbb{P}\left(A_1(0, t_u) - c\varepsilon t_u - \phi_1 c(1 - \varepsilon)t_u > u; \inf_{t \in [\varepsilon t_u, t_u]} A_2(t) - ct > 0\right) \tag{10}$$

$$= \mathbb{P}(A_1(0, t_u) - c\varepsilon t_u - \phi_1 c(1 - \varepsilon)t_u > u) \mathbb{P}\left(\inf_{t \in [\varepsilon t_u, t_u]} A_2(t) - ct > 0\right) \tag{11}$$

where (9) follows from the fact that $B_1(0, \varepsilon t_u) \leq c\varepsilon t_u$ and $B_1(\varepsilon t_u, t_u) + B_2(\varepsilon t_u, t_u) \leq ct_u(1 - \varepsilon)$, and (10) follows from the trivial inequality $Q_1(t) \geq Q_1^c(t) \forall t \in \mathbb{R}$. Applying Lemma 5.3 we infer that

$$\begin{aligned} \log \mathbb{P}(Q_1 > u) &\gtrsim -\frac{(u + (\phi_1 c - \mu_1 + \varepsilon \phi_2 c)t_u)^2}{2v_{A_1}(t_u)} - \frac{(c - \mu_2)^2 t_u^2}{2v_{A_2}(\varepsilon t_u)} \\ &\sim -\frac{(u + (\phi_1 c - \mu_1 + \varepsilon \phi_2 c)t_u)^2}{2v_{A_1}(t_u)}. \end{aligned}$$

Now putting $\varepsilon \downarrow 0$ we apply Lemma 5.2 to obtain

$$\begin{aligned} \log \mathbb{P}(Q_1 > u) &\gtrsim -\frac{(u + (\phi_1 c - \mu_1)t_u)^2}{2v_{A_1}(t_u)} \\ &\sim -\frac{1}{2}(\phi_1 c - \mu_1)^{\alpha_1} \left(\frac{\alpha_1}{2 - \alpha_1}\right)^{-\alpha_1} \left(\frac{2}{2 - \alpha_1}\right)^2 \frac{u^2}{v_{A_1}(u)}, \end{aligned}$$

which completes the proof. \square

6 Mixing a Gaussian flow with an on-off flow

In this section we consider the same model as in the previous sections, but we now assume $\{A_2(t), t \in \mathbb{R}\}$ to be an on-off process with regularly varying residual on times. Recall that we assume the system to be stable, i.e., $\mu_1 + \mu_2 < c$. We introduce some additional notation. The on periods of flow 2 are denoted by T_{On} , the fraction of time the process is off by p_{Off} , and the on rate by r_{On} . We define the indicator function $I_{\{E\}}$ returning 1 in case of event E and 0 otherwise, and we define $d_t A_2(t)$ to be the input rate of flow 2 at time t . We need two additional lemmas to derive the result in Theorem 6.1. The first result is well-known, it gives the asymptotic behavior of the residual busy period, denoted by P^r , in an ordinary queue with service rate c fed by an on-off process, say process Y . Let us denote the on periods of Y by Y_{On} , the on rate by r_Y , the average rate by μ_Y and the fraction of on time by p_Y . Then the lemma, which is due to [5, 30], is as follows.

Lemma 6.1 *If Y_{On} is intermediately regularly varying, and $\mu_Y < c < r_Y$, then*

$$\mathbb{P}(P^r > u) \sim (1 - p_Y) \frac{c}{c - \mu_Y} \mathbb{P}\left(Y_{\text{On}}^r > \frac{u(c - \mu_Y)}{r_Y - \mu_Y}\right).$$

The second result is proved in [31] (Theorem 5.1). It shows that an ordinary queue fed by a centered Gaussian process X and an on-off process Y as described above, but now with regularly varying residual on times, reaches a large value because of atypical behavior of both processes.

Lemma 6.2 *Let X satisfy **C1-C2** with regularity index α , and let Y_{On}^r be regularly varying at ∞ with $r_Y < c$, then*

$$\mathbb{P}(Q_{X+Y} > u) \sim p_Y \mathbb{P}(Q_X^{c-r_Y} > u) \mathbb{P}\left(Y_{\text{On}}^r > \frac{1}{c - r_Y} \frac{\alpha}{2 - \alpha} u\right).$$

Using these lemmas we obtain the following result.

Theorem 6.1 *If T_{On}^r is regularly varying, $\mu_1 < \phi_1 c$, and A_1 satisfies **C1-C2**, then*

$$\log \mathbb{P}(Q_1 > u) \sim \log \mathbb{P}(Q_1^{\phi_1 c} > u).$$

Qualitatively we have a result similar to that of Theorems 3.1 (**S1**) and 3.3 (**S3**). The workload of the Gaussian flow, flow 1, gets large if it shows atypical behavior itself. Flow 2 does not show up in the asymptotics, as it only takes its guaranteed rate in the interval that flow 1 needs to build up a large queue. This is due to the long-range dependent behavior of flow 2. As its residual on times are regularly varying, it easily takes its guaranteed rate $\phi_2 c$ for the necessary amount of time.

Proof Trivially,

$$\begin{aligned}\mathbb{P}(Q_1 > u) &= \mathbb{P}\left(\sup_{t \geq 0} \left\{ A_1(t) - \phi_1 ct - \int_0^t I_{\{Q_2(s)=0\}}(\phi_2 c - d_s A_2(s)) ds \right\} > u\right) \\ &= \mathbb{P}\left(\sup_{t \geq 0} \left\{ A_1(t) - \phi_1 ct - \phi_2 c \int_0^t I_{\{Q_2(s)=0\}} ds + \int_0^t d_s A_2(s) I_{\{Q_2(s)=0\}} ds \right\} > u\right)\end{aligned}$$

Observe that the event $\{Q_2(s) = 0\}$ implies that flow 2 is off at time s . Hence,

$$\mathbb{P}(Q_1 > u) = \mathbb{P}\left(\sup_{t \geq 0} \left\{ A_1(t) - ct + \phi_2 c \int_0^t I_{\{Q_2(s) > 0\}} ds \right\} > u\right).$$

We now proceed to derive an upper and a lower bound for $\mathbb{P}(Q_1 > u)$, which will appear to logarithmically coincide.

Upper bound. Observing that $\int_0^t I_{\{Q_2(s) > 0\}} ds$ is the amount of time that Q_2 was busy in $[0, t]$, it holds that

$$\mathbb{P}(Q_1 > u) \leq \mathbb{P}\left(\sup_{t \geq 0} \left\{ A_1(t) - ct + \phi_2 c \int_0^t I_{\{Q_2^{\phi_2 c}(s) > 0\}} ds \right\} > u\right),$$

where $\phi_2 c \int_0^t I_{\{Q_2^{\phi_2 c}(s) > 0\}} ds$ is an on-off process: the on times are the busy periods of a queue with service rate $\phi_2 c$ that is fed only by flow 2, the on rate is equal to $\phi_2 c$, and we denote the fraction of time that the process is on by p_{up} . Applying first Lemma 6.2, for which we need to assume that $\mu_1 < \phi_1 c$, and then Lemma 6.1 we obtain

$$\mathbb{P}(Q_1 > u) \leq p_{\text{up}} \mathbb{P}\left(Q_1^{\phi_1 c} > u\right) p_{\text{off}} \frac{\phi_2 c}{\phi_2 c - \mu_2} \mathbb{P}\left(T_{\text{on}}^r > \frac{1}{\phi_1 c} \frac{\alpha_1}{2 - \alpha_1} \frac{u(\phi_2 c - \mu_2)}{r_{\text{on}} - \mu_2}\right).$$

Lower bound. Observe that the event $\{Q_2(t) > 0\}$ is implied by the event that flow 2 is in the on state at time t . Hence,

$$\mathbb{P}(Q_1 > u) \geq \mathbb{P}\left(\sup_{t \geq 0} \left\{ A_1(t) - ct + \phi_2 c \int_0^t I_{\{d_s A_2(s) > 0\}} ds \right\} > u\right),$$

where $\phi_2 c \int_0^t I_{\{d_s A_2(s) > 0\}}$ is an on-off process with on times T_{on} and on rate $\phi_2 c$. Again applying Lemma 6.2 results in

$$\mathbb{P}(Q_1 > u) \geq (1 - p_{\text{off}}) \mathbb{P}\left(Q_1^{\phi_1 c} > u\right) \mathbb{P}\left(T_{\text{on}}^r > \frac{1}{\phi_1 c} \frac{\alpha_1}{2 - \alpha_1} u\right).$$

Logarithmic asymptotics. Because T_{on}^r is regularly varying, it is easily seen that in the upper bound, for any $\gamma > 0$,

$$\lim_{u \rightarrow \infty} \frac{1}{u^\gamma} \log \left(p_{\text{up}} p_{\text{off}} \frac{\phi_2 c}{\phi_2 c - \mu_2} \mathbb{P}\left(T_{\text{on}}^r > \frac{1}{\phi_1 c} \frac{\alpha_1}{2 - \alpha_1} \frac{u(\phi_2 c - \mu_2)}{r_{\text{on}} - \phi_2 c}\right) \right) = 0,$$

and that similarly in the lower bound, for any $\gamma > 0$,

$$\lim_{u \rightarrow \infty} \frac{1}{u^\gamma} \log \left((1 - p_{\text{off}}) \mathbb{P}\left(T_{\text{on}}^r > \frac{1}{\phi_1 c} \frac{\alpha_1}{2 - \alpha_1} u\right) \right) = 0,$$

which completes the proof. \square

References

- [1] S.C. Borst, O.J. Boxma, and P.R. Jelenković. Reduced-load equivalence and induced burstiness in GPS queues with long-tailed traffic flows. *Queueing Systems*, 43:273–306, 2003.
- [2] S.C. Borst, K. Dębicki, and A.P. Zwart. The supremum of a Gaussian process over a random interval. *Statistics and Probability Letters*, 68:221–234, 2004.
- [3] S.C. Borst, M. Mandjes, and M.J.G. van Uitert. Generalized Processor Sharing queues with heterogeneous traffic classes. *Advances in Applied Probability*, 35:806–845, 2003.
- [4] S.C. Borst, M. Mandjes, and M.J.G. van Uitert. Generalized Processor Sharing with light-tailed and heavy-tailed input. *IEEE/ACM Transactions on Networking*, 11:821–834, 2003.
- [5] O.J. Boxma and V. Dumas. The busy period in the fluid queue. *Performance Evaluation Review*, 26:100–110, 1998.
- [6] K. Dębicki. A note on LDP for supremum of Gaussian processes over infinite horizon. *Statistics and Probability Letters*, 44:211–219, 1999.
- [7] K. Dębicki. Ruin probability for Gaussian integrated processes. *Stochastic Processes and their Applications*, 98:151–174, 2002.
- [8] K. Dębicki and Z. Palmowski. On-off fluid models in heavy traffic environment. *Queueing Systems*, 33:327–338, 1999.
- [9] K. Dębicki and T. Rolski. A Gaussian fluid model. *Queueing Systems*, 20:433–452, 1995.
- [10] A.B. Dieker. Conditional limit theorems for queues with Gaussian input; a weak convergence approach. to appear in *Stochastic Processes and their Applications*, 2005.
- [11] C. Fraleigh, F. Tobagi, and C. Diot. Provisioning IP backbone networks to support latency sensitive traffic. In *Proceedings Infocom 2003*, San Francisco, USA, 2003.
- [12] J. Kilpi and I. Norros. Testing the Gaussian approximation of aggregate traffic. The 2nd Internet Measurement Workshop, Marseille, France, 2002.
- [13] C. Kotopoulos, N. Likhanov, and R.R. Mazumdar. Asymptotic analysis of GPS systems fed by heterogeneous long-tailed sources. In *Proceedings Infocom 2001*, pages 299–308, Anchorage, Alaska, USA, 2001.

- [14] V. Kulkarni and T. Rolski. Fluid model driven by an ornstein-uhlenbeck process. *Probability in the Engineering and Informational Sciences*, 8:403–417, 1994.
- [15] W.E. Leland, M.S. Taqqu, W. Willinger, and D.V. Wilson. On the self-similar nature of Ethernet traffic. *IEEE/ACM Transactions on Networking*, 2:1–15, 1994.
- [16] P. Lieshout and M. Mandjes. Many-sources delay asymptotics for GPS with Gaussian inputs. Work in progress, 2005.
- [17] P. Lieshout, M. Mandjes, and S. Borst. GPS scheduling: selection of optimal weights and comparison with strict priorities. Submitted, 2005.
- [18] M. Mandjes and M.J.G. van Uitert. Sample-path large deviations for Generalized Processor Sharing queues with Gaussian inputs. *Performance Evaluation*, 61:225–256, 2005.
- [19] M. Mandjes and M.J.G. van Uitert. Sample-path large deviations for tandem and priority queueus with Gaussian inputs. *Annals of Applied Probability*, 15:1193–1226, 2005.
- [20] P. Mannersalo and I. Norros. GPS schedulers and Gaussian traffic. In *Proceedings Infocom 2002*, pages 1660–1667, New York, New York, USA, 2002.
- [21] T. Mikosch, S. Resnick, H. Rootzén, and A. Stegeman. Is network traffic approximated by stable Lévy motion or fractional Brownian motion? *Annals of Applied Probability*, 12:23–68, 2002.
- [22] I. Norros. A storage model with self-similar input. *Queueing Systems*, 16:387–396, 1994.
- [23] V.I. Piterbarg. *Asymptotic methods in the theory of Gaussian processes and fields*, volume 148 of *Translations of mathematical monographs*. AMS, Providence RI, 1996.
- [24] E. Reich. On the integrodifferential equation of Takács I. *Annals of Mathematical Statistics*, 29:563–570, 1958.
- [25] M.S. Taqqu, W. Willinger, and R. Sherman. Proof of a fundamental result in self-similar traffic modeling. *Computer Communication Review*, 27:5–23, 1997.
- [26] M.J.G. van Uitert. *Generalized Processor Sharing Queues*. PhD thesis, Eindhoven University of Technology, Eindhoven, The Netherlands, 2003.
- [27] M.J.G. van Uitert and S.C. Borst. A reduced-load equivalence for Generalised Processor Sharing networks with long-tailed input flows. *Queueing Systems*, 41:123–163, 2002.
- [28] Z.-L. Zhang. Large deviations and the Generalized Processor Sharing scheduling for a two-queue system. *Queueing Systems*, 26:229–245, 1997.

- [29] Z.-L. Zhang. Large deviations and the Generalized Processor Sharing scheduling for a multiple-queue system. *Queueing Systems*, 28:349–376, 1998.
- [30] A.P. Zwart. Tail asymptotics for the busy period in the GI/G/1 queue. *Mathematics of Operations Research*, 26:485–493, 2001.
- [31] A.P. Zwart, S.C. Borst, and K. Değbicki. Subexponential asymptotics of hybrid fluid and ruin models. *Annals of Applied Probability*, 15:500–517, 2005.