Queueing models for bandwidth-sharing disciplines
Lieshout, P.M.D.

Citation for published version (APA):
Chapter 2

Gaussian methodology

Part I of this thesis is mainly devoted to differentiated bandwidth sharing in network links. In particular, our goal is to study a single-node system that operates under the GPS discipline. As explained in Chapter 1, such systems are typically analyzed at the burst-level. At this time scale traffic that enters the system approximately behaves as a continuous stream of work, i.e., as fluid. We consider a general and versatile class of fluid input processes, viz. the class of Gaussian inputs. The present chapter serves as an introductory chapter, in which we present the basic concepts and machinery that are needed in this part. Furthermore, we illustrate the use of this machinery on a simple system. We refer to [127] for an overview on queues with Gaussian input, so-called Gaussian queues.

2.1 Preliminaries on Gaussian random variables

Before introducing Gaussian inputs in Section 2.2, we first explain the concept of Normal (or Gaussian) random variables. In addition, we discuss some well-known properties of these variables in this section. Below we present the results that are of importance in Part I.

A Normal random variable $X$ with mean $\mu$ and variance $\sigma^2$ has density

$$
\frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}},
$$

which is denoted by $X \sim N(\mu, \sigma^2)$. The Moment Generating Function (MGF) of $X$ is given by

$$
\mathbb{E} e^{sX} = e^{\mu s + \frac{s^2}{2}}, \quad s \in \mathbb{R}.
$$

In the special case of $\mu = 0$ and $\sigma = 1$, we call $X$ standard Normal. Throughout Part I, we denote the density function of a standard Normally distributed variable $X$ by

$$
\phi(x) := \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}},
$$
its distribution function by

\[ \Phi(x) := \int_{-\infty}^{x} \phi(y) dy, \]

and the corresponding tail distribution by \( \Psi(x) := 1 - \Phi(x) \). A well-known (double) inequality is (see page 5 of [137]):

\[ \frac{1}{x + 1/x} \phi(x) \leq \Psi(x) \leq \frac{1}{x} \phi(x), \quad x > 0. \] (2.1)

It follows that, for \( x \to \infty \),

\[ \Psi(x) \sim \frac{1}{\sqrt{2\pi x}} e^{-x^2/2} =: \zeta(x), \] (2.2)

where we write \( f(x) \sim g(x) \) when \( f(x)/g(x) \to 1 \) if \( x \to \infty \).

A random variable \( X \) is \( d \)-variate Normally distributed, \( d \in \mathbb{N} \), with \( d \)-dimensional mean vector \( \mu \) and (non-singular) \( d \times d \) covariance-matrix

\[
\begin{bmatrix}
\text{Var}X_1 & \text{Cov}(X_1, X_2) & \ldots & \text{Cov}(X_1, X_d) \\
\text{Cov}(X_1, X_2) & \text{Var}X_2 & \ldots & \text{Cov}(X_2, X_d) \\
\vdots & \vdots & \ddots & \vdots \\
\text{Cov}(X_1, X_d) & \text{Cov}(X_2, X_d) & \ldots & \text{Var}X_d \\
\end{bmatrix},
\]

i.e., \( X \sim N_d(\mu, \Sigma) \), if \( X \) has density

\[
\frac{1}{(2\pi)^{d/2} \sqrt{\det(\Sigma)}} \exp\left(-\frac{(x - \mu)^T \Sigma^{-1} (x - \mu)}{2}\right),
\]

where \( \det(\Sigma) \) is the determinant of the matrix \( \Sigma \), and \( \Sigma^{-1} \) denotes the inverse of \( \Sigma \).

Suppose now that \( (Y, X) \) is \( (q + d) \)-variate Normally distributed, where \( Y \) is \( q \)-dimensional and \( X \) is \( d \)-dimensional. The mean vector \( \mu \) and covariance matrix \( \Sigma \) are partitioned as follows:

\[
\mu = \begin{pmatrix} \mu_Y \\ \mu_X \end{pmatrix}, \quad \text{with sizes } \begin{pmatrix} q \times 1 \\ d \times 1 \end{pmatrix};
\]

\[
\Sigma = \begin{pmatrix} \Sigma_{YY} & \Sigma_{YX} \\ \Sigma_{XY} & \Sigma_{XX} \end{pmatrix}, \quad \text{with sizes } \begin{pmatrix} q \times q & q \times d \\ d \times q & d \times d \end{pmatrix}.
\]

Then the random variable \( (Y | X = x) \), for some \( x \in \mathbb{R}^d \), is Normally distributed with mean vector \( \overline{\mu} \) and covariance matrix \( \overline{\Sigma} \), where

\[
\overline{\mu} = \mu_Y + \Sigma_{YX} \Sigma_{XX}^{-1} (x - \mu_X), \quad \text{and}
\]

\[
\overline{\Sigma} = \Sigma_{YY} - \Sigma_{YX} \Sigma_{XX}^{-1} \Sigma_{XY}.
\]

(2.4)
2.2 Gaussian input

and

\[ \Sigma = \Sigma_{YY} - \Sigma_{YX} \Sigma^{-1}_{XX} \Sigma_{XY}, \]

i.e., \( (Y|X = x) \sim N_q(\overline{\mu}, \Sigma) \). The above indicates that knowing the value of \( X \) to be \( x \) alters the mean vector and the covariance matrix of \( Y \), as \( Y \sim N_q(\mu_Y, \Sigma_{YY}) \). It is noted that, as opposed to the conditional mean \( \mu \), the conditional variance \( \Sigma \) does not depend on the value \( x \).

2.2 Gaussian input

In the previous section we introduced Normal random variables. Below we explain the connection between these and Gaussian input.

Let \( \{A(t), t \in \mathbb{R}\} \) be an input process, with \( A(0) \equiv 0 \). Also, let \( A(s, t) := A(t) - A(s) \) denote the amount of traffic arriving in \([s, t), s < t\]. Note that \( A(t) (\equiv -A(t)) \) denotes the amount of traffic generated in the interval \([0, t]\) (\([t, 0]\)) if \( t \geq 0 \) (\( t \leq 0 \)).

The input process \( A(t) \) is called a Gaussian process with stationary increments, if for all \( s < t \), \( A(s, t) \) is Normally distributed with mean \( \mu \cdot (t - s) \) and variance \( v(t - s) \), where \( \mu := \mathbb{E}A(1) \) and \( v(t - s) := \text{Var}A(s, t) \). Hence, the entire probabilistic behavior of the Gaussian input process can be expressed in terms of a mean traffic rate \( \mu \) and a variance function \( v(\cdot) : \mathbb{R}_+ \to \mathbb{R}_+ \). The assumption of stationary increments entails that the law of \( A(s, t) \) only depends on the length of the interval, and not on its position. We also define the centered process \( \overline{A}(t) := A(t) - \mu t \).

The class of Gaussian inputs is extremely rich, and this richness is best illustrated by the variety of possible choices for the variance function \( v(\cdot) \), see Chapter 2 of [127] for more details. The variance function fully determines the correlation structure. To see this, first note that, assuming \( 0 < s < t \), we have

\[ \text{Cov}(A(s), A(t)) = \text{Cov}(A(s), A(s) + A(s, t)) = \text{Var}A(s) + \text{Cov}(A(s), A(s, t)). \]

Then, using that

\[ \text{Var}A(t) = \text{Var}(A(s) + A(s, t)) = \text{Var}A(s) + \text{Var}A(s, t) + 2\text{Cov}(A(s), A(s, t)), \]

we find that

\[ \Gamma(s, t) := \text{Cov}(A(s), A(t)) = \text{Cov}(\overline{A}(s), \overline{A}(t)) = \frac{1}{2}(v(t) + v(s) - v(t - s)). \]

Throughout Part I we impose the following (weak) assumptions on \( v(\cdot) \).

**Assumption 2.2.1** The variance function \( v(\cdot) \) satisfies:

1. \( v(\cdot) \in C_1([0, \infty)) \);
2. For some \( \alpha < 2 \) it holds that \( v(t)t^{-\alpha} \to 0 \), as \( t \to \infty \);
3. \( v(\cdot) \) is strictly increasing.
We need the first two assumptions to apply certain techniques, which will be defined in Section 2.3. Assumption A3 is needed in the proofs of some lemmas and theorems. It is noted that various measurement studies have confirmed that A1-A3 are natural assumptions.

The class of Gaussian inputs covers a broad range of correlation structures. Importantly, Gaussian models include both SRD and LRD traffic. We say that $A(\cdot)$ exhibits long-range dependence if

$$\sum_{k=1}^{\infty} \text{Cov}(A(0, 1), A(k, k+1)) = \infty,$$

and that $A(\cdot)$ is SRD otherwise ($< \infty$).

Closely related to the notion of LRD, is the notion of self-similarity. A process $A(\cdot)$ is self-similar with Hurst parameter $H$, $H \in (0, 1)$, if

$$a^{-H} A(at) \overset{d}{=} A(t), \quad \forall a > 0,$$

where $\overset{d}{=}$ denotes equality in distribution.

We now mention two examples of Gaussian inputs that are of importance in this monograph, both satisfying A1-A3. We start with a fractional Brownian motion (fBm), which has variance function $v(t) = t^{2H}$, with $H \in (0, 1)$, implying that fBm exhibits self-similarity with Hurst parameter $H$. For $H \in (0, 1/2)$ it is easy to show that fBm is LRD, whereas for $H \in (1/2, 1)$ fBm is SRD. In the special case of $H = 1/2$, fBm reduces to an ordinary Brownian motion, which has independent increments. This illustrates that the notions of self-similarity and LRD are related in some cases, but not equivalent. Another example of Gaussian inputs is the integrated Ornstein-Uhlenbeck (iOU) process, which has variance function $v(t) = t - 1 + e^{-t}$. It is an easy exercise to show that iOU exhibits short-range dependence.

We remark that Gaussian inputs are often useful as approximations of well-known non-Gaussian input processes. We say that an input process $\{\tilde{A}(t), t \in \mathbb{R}\}$ with stationary increments has the Gaussian counterpart $\{A(t), t \in \mathbb{R}\}$ if $A(\cdot)$ is Gaussian and furthermore $E\tilde{A}(t) = EA(t)$ and $\text{Var}\tilde{A}(t) = \text{Var}A(t)$ for all $t$. In other words, $\tilde{A}(\cdot)$ inherits the correlation structure of $A(\cdot)$. Typical examples are the Gaussian counterpart of the Poisson stream, the M/G/$\infty$ input model, and the purely periodic stream. For more details we refer to Section 2.5 of [127].

A natural question that arises is whether Gaussian traffic describes real traffic accurately. Before addressing this question, let us first recall four main characteristics of real traffic. Typically, 1) it can be assumed that the real traffic input process is stationary, at least over suitable time periods. Furthermore, 2) the aggregation level at the core network is usually quite high, as the total input stream to each node in the network consists of a superposition of a large number of individual streams. We already argued in the previous chapter that extensive measurements showed that
3) network traffic exhibits significant positive correlation over a wide range of time scales, and 4) the traffic rate is bursty over a wide range of time scales, i.e., it exhibits extreme irregularity.

Let us now verify whether Gaussian inputs can capture these four properties. Clearly, Gaussian sources have stationary increments, so property 1) is fulfilled. Gaussian traffic arises as limiting process of the superposition of a large number of independent traffic sources, and is thus appropriate if the aggregation level is sufficiently large. In [64] it was empirically shown that a relatively low aggregation level is already sufficient for Gaussianity. A complicating issue is the fact that elastic traffic is controlled through feedback loops like TCP. In [97] it was, however, argued that (non-feedback) Gaussian traffic models are still justified as long as the level of aggregation is sufficiently large (both in time and number of flows), implying that property 2) is satisfied. Properties 3) and 4) are respectively satisfied if the arrival process is LRD and if the traffic rate process behaves irregularly at small time scales (i.e., it could have non-differentiable trajectories). Clearly, not all Gaussian inputs satisfy these last two properties, but for example fBm with $H > 1/2$ is a suitable candidate.

An issue associated with Gaussian traffic is that the cumulative input process will be (locally) decreasing, whereas the amount of real traffic generated in some interval cannot be negative. This fact may seem troublesome at first sight, however, similar problems appear in different settings. For example, consider the situation where the number of successes in $n$ Bernoulli trials is approximated by a Gaussian random variable for large $n$. In this case the real distribution of the number of successes is also closely approximated by a Normal distribution, which has $\mathbb{R}$, i.e., also negative values, as support. Although the cumulative input process will be (locally) decreasing, this occurs with a small probability, as the cumulative input process typically has a positive drift. In addition, as we will see in Section 2.4, the steady-state buffer content of a Gaussian queue can always be evaluated and lives on $[0, \infty)$.

Above we provided qualitative arguments suggesting that Gaussian inputs can describe real network traffic accurately. For more validation and justification of this claim we refer to [138] and Chapter 3 of [127].

2.3 Large deviations for Gaussian processes

In this section we consider large deviation results for Gaussian processes. As we will see below, large deviations are closely related to rare events. To explain the concept of large deviations in general, we start with a number of results relating to a finite-dimensional setting. Next we consider the infinite-dimensional framework, which is the one corresponding to Gaussian processes, and present the theorems that are of interest in this monograph. This section is based on Chapter 4 in [127].
2.3.1 Finite-dimensional framework: Cramér’s theorem

Consider a sequence of i.i.d. random variables \( X_1, \ldots, X_n \) that are distributed like a random variable \( X \), which has mean \( \mu := \mathbb{E}X \), with \( -\infty < \mu < \infty \). The law of large numbers states that the sample mean \( n^{-1} \sum_{i=1}^{n} X_i \) converges to \( \mu \) almost surely as \( n \to \infty \). Let us now focus on the probability that, although \( n \to \infty \), this sample mean does deviate severely from \( \mu \). Below we wish to analyze:

\[
P \left( \frac{1}{n} \sum_{i=1}^{n} X_i > x \right),
\]

for \( x > \mu \), where \( n \) and \( x \) are fixed.

Define the MGF of \( X \) by \( M(\theta) := \mathbb{E}e^{\theta X} \), and assume that the MGF is finite in a neighborhood of 0, so that all moments of \( X \) are finite. It is now straightforward to show that

\[
P \left( \frac{1}{n} \sum_{i=1}^{n} X_i > x \right) = P \left( e^{\theta \sum_{i=1}^{n} X_i} > e^{\theta x} \right) \leq e^{-n\theta x} \mathbb{E}e^{\theta \sum_{i=1}^{n} X_i} = e^{-n\theta x} (M(\theta))^n,
\]

for any \( \theta \geq 0 \), where we use the Markov inequality: \( P(Y \geq y) \leq \mathbb{E}Y/y \) for any non-negative random variable \( Y \), where \( \mathbb{E}Y < \infty \). As (2.5) holds for any \( \theta \geq 0 \), it also holds for the \( \theta \) that gives the tightest upper bound:

\[
P \left( \frac{1}{n} \sum_{i=1}^{n} X_i > x \right) \leq \inf_{\theta \geq 0} e^{-n\theta x} (M(\theta))^n = \exp \left( -n \sup_{\theta \geq 0} (\theta x - \log M(\theta)) \right). \tag{2.6}
\]

Equation (2.6) is known as the Chernoff bound and shows that the probability that the sample mean exceeds \( \mu \) decays exponentially as \( n \) increases, i.e., the decay rate, or equivalently, the rate function equals:

\[
J(x) := \sup_{\theta \geq 0} (\theta x - \log M(\theta)),
\]

where \( J(x) \) is referred to as the Fenchel-Legendre transform of \( \log M(\theta) \). Here \( J(x) \) can be interpreted as a cost function: the larger the distance to the mean \( \mu \) is, the higher the cost are. It is an easy exercise to show that \( J(x) > 0 \) if \( x \neq \mu \), \( J(\mu) = 0 \), and \( J(\cdot) \) is convex, see Exercise 4.1.1 in [127].

In turns out that the Chernoff bound is tight on a logarithmic scale. Before we state this result, known as Cramér’s theorem [52], we first need the following definition.

**Definition 2.3.1** A sequence \( Y_1, Y_2, \ldots \) obeys the large deviations principle (LDP) with rate function \( K(\cdot) \) if:
(a) For any closed set $F$,
\[ \limsup_{n \to \infty} \frac{1}{n} \log P \left( \frac{1}{n} \sum_{i=1}^{n} Y_i \in F \right) \leq - \inf_{x \in F} K(x); \]

(b) For any open set $G$,
\[ \liminf_{n \to \infty} \frac{1}{n} \log P \left( \frac{1}{n} \sum_{i=1}^{n} Y_i \in G \right) \geq - \inf_{x \in G} K(x). \]

**Theorem 2.3.2 [Cramér]** Let $X_i \in \mathbb{R}$ be i.i.d. random variables, distributed as a random variable $X$ with mean $\mu$ and MGF $M(\theta) = \mathbb{E} e^{\theta X}$ that is finite in a neighborhood of 0. Then $X_1, X_2, \ldots$ obeys the LDP with rate function $J(\cdot)$.

Using Theorem 2.3.2 and the fact that $J(\cdot)$ is convex, it can be proved that
\[ \lim_{n \to \infty} \frac{1}{n} \log P \left( \frac{1}{n} \sum_{i=1}^{n} X_i > x \right) = -J(x). \] (2.7)

Hence, Cramér’s theorem gives information on the logarithm of the probability, rather than the probability itself. From (2.7) we conclude that
\[ P \left( \frac{1}{n} \sum_{i=1}^{n} X_i > x \right) = f(x, n) e^{-nJ(x)}, \]
where $f(x, n)$ is not given explicitly, but known to be subexponential, i.e.,
\[ \lim_{n \to \infty} \frac{\log f(x, n)}{n} = 0. \]

In the absence of an explicit expression for $f(x, n)$, one may use the approximation
\[ P \left( \frac{1}{n} \sum_{i=1}^{n} X_i > x \right) \approx e^{-nJ(x)}. \] (2.8)

We remark that in some cases this approximation may be inaccurate, as some polynomial function $n^\alpha$, where $\alpha$ can be both positive and negative, or a function of the type $\exp(n^{1-\epsilon})$, where $\epsilon$ is a small positive number, can be part of $f(x, n)$. However, often this approximation is useful to gain insight.

Let us now consider the probability that $n^{-1} \sum_{i=1}^{n} X_i$ is contained in some set $B$. Then we find the approximation
\[ P \left( \frac{1}{n} \sum_{i=1}^{n} X_i \in B \right) \approx e^{-n \inf_{x \in B} J(x)}. \]
That is, the rate function is determined by the most likely point in the set $B$, i.e., the point $x^* := \arg \inf_{x \in B} J(x)$. Clearly, if $\mu \in B$ then $x^* = \mu$ and $\inf_{x \in B} J(x) = 0$.

As may be expected, Theorem 2.3.2 can also be extended to a multivariate, say $d$-dimensional with $d \in \mathbb{N}$, version. Let $\langle a, b \rangle$ denote the inner product $\sum_{i=1}^{d} a_i b_i$.

**Theorem 2.3.3 [Multivariate Cramér]** Let $X_i \in \mathbb{R}^d$ be i.i.d. $d$-dimensional random variables, distributed as a random variable $X$ with mean $\mu$ and MGF $M(\theta) = \mathbb{E}e^{\langle \theta, X \rangle}$ that is finite in a neighborhood of 0. Then the sequence $X_1, X_2, \ldots$ obeys the LDP with rate function $J_d(\cdot)$, where

$$J_d(x) := \sup_{\theta \in \mathbb{R}^d} \left( \langle \theta, x \rangle - \log M(\theta) \right).$$

(2.9)

Considering the specific case that $X$ is $d$-dimensional Normally distributed with mean vector $\mu$ and non-singular covariance matrix $\Sigma$, see Section 2.1, we find that

$$\log M(\theta) = \log \mathbb{E}e^{\langle \theta, X \rangle} = \langle \theta, \mu \rangle + \frac{1}{2} \theta^T \Sigma \theta.$$ Consequently, with $(x - \mu)^T := (x_1 - \mu_1, \ldots, x_d - \mu_d)$, we deduce that

$$\theta^* = \Sigma^{-1}(x - \mu) \quad \text{and} \quad J_d(x) = \frac{1}{2} (x - \mu)^T \Sigma^{-1}(x - \mu),$$

where $\theta^*$ denotes the optimizer in (2.9). The following theorem follows directly from the above, and will be used in Chapter 3. We refer to Exercise 4.1.9 in [127] for more details.

**Theorem 2.3.4** Let $(X, Y) \sim N_2(0, \Sigma)$, for a non-degenerate 2-dimensional covariance-matrix $\Sigma$. Then,

(i) $\lim_{n \to \infty} \frac{1}{n} \log \mathbb{P} \left( \frac{1}{n} \sum_{i=1}^{n} X_i \geq x \right) = \frac{1}{2} x^2 / (\Sigma_{XX})^2$;

(ii) $\lim_{n \to \infty} \frac{1}{n} \log \mathbb{P} \left( \frac{1}{n} \sum_{i=1}^{n} X_i \geq x, \frac{1}{n} \sum_{i=1}^{n} Y_i \geq y \right) = \inf_{a \geq x} \inf_{b \geq y} \Lambda(a, b)$,

where $\Lambda(a, b) := \frac{1}{2} \left( \begin{array}{cc} a & b \\ b & a \end{array} \right) \Sigma^{-1} \left( \begin{array}{c} a \\ b \end{array} \right)$ and $x, y > 0$.

**2.3.2 Infinite-dimensional framework: Schilder’s theorem**

Below we present an extension of Cramér’s theorem that relates to an infinitely-dimensional setting: the generalized version of Schilder’s theorem [15]. Whereas ‘Cramer’ can be applied to describe the likelihood of a sample mean of Normal random variables or vectors attaining a rare value, ‘Schilder’ describes the large deviations of the sample mean of Gaussian processes.

Let $A_1(\cdot), A_2(\cdot), \ldots$ be a sequence of i.i.d. Gaussian processes, distributed as a Gaussian process with variance function $v(\cdot)$. For large values of $n$ it is clear that the sample mean path $n^{-1} \sum_{i=1}^{n} A_i(t)$ approaches $\mu t$ almost surely, where $\mu := \mathbb{E}A_1(1)$. 

2.3 Large deviations for Gaussian processes

'Schilder' can be applied to determine the probability that the sample mean path deviates from some mean path. In particular, it characterizes the exponential decay rate of the sample mean path being contained in some specific set.

We continue with a description of the framework of Schilder’s sample-path large deviations principle (LDP) (see [15], and also Theorem 1.3.27 of [53] for a more detailed treatment). Below we assume that the processes \( A_1(\cdot), A_2(\cdot), \ldots \) are centered, but it is clear that the results for centered processes can be translated immediately into results for non-centered processes. Define the path space \( \Omega \) as

\[
\Omega := \left\{ \omega: \mathbb{R} \to \mathbb{R}, \text{ continuous, } \omega(0) = 0, \lim_{t \to \pm \infty} \frac{\omega(t)}{1 + |t|} = 0 \right\},
\]

which is a separable Banach space by imposing the norm

\[
||\omega||_{\Omega} := \sup_{t \in \mathbb{R}} |\omega(t)|.
\]

We note that in [7] it was pointed out that \( A_i(\cdot) \) can be realized on \( \Omega \) under Assumption A2. Then one can construct a reproducing kernel Hilbert space \( R \subseteq \Omega \), consisting of elements that are roughly as smooth as the covariance function \( \Gamma(s, \cdot) \); for details, see [8]. We start from a 'smaller' space \( R^* \), defined by

\[
R^* := \left\{ \omega: \mathbb{R} \to \mathbb{R}, \ \omega(\cdot) = \sum_{i=1}^{n} a_i \Gamma(s_i, \cdot), \ a_i, s_i \in \mathbb{R}, n \in \mathbb{N} \right\}.
\]

The inner product on this space \( R^* \) is, for \( \omega_a, \omega_b \in R^* \), defined as

\[
\langle \omega_a, \omega_b \rangle_R := \left\{ \sum_{i=1}^{n} a_i \Gamma(s_i, \cdot), \sum_{j=1}^{n} b_j \Gamma(s_j, \cdot) \right\} = \sum_{i=1}^{n} \sum_{j=1}^{n} a_i b_j \Gamma(s_i, s_j);
\]

notice that this implies \( \langle \Gamma(s, \cdot), \Gamma(\cdot, t) \rangle_R = \Gamma(s, t) \). This inner product has the following useful property, which is known as the reproducing kernel property,

\[
\omega(t) = \sum_{i=1}^{n} a_i \Gamma(s_i, t) = \left\{ \sum_{i=1}^{n} a_i \Gamma(s_i, \cdot), \Gamma(t, \cdot) \right\} = \langle \omega(\cdot), \Gamma(\cdot, t) \rangle_R.
\]

From this we introduce the norm \( ||\omega||_R := \sqrt{\langle \omega, \omega \rangle_R} \). The closure of \( R^* \) under this norm is defined as space \( R \). Now we can define the rate function:

\[
I(\omega) := \begin{cases} \frac{1}{2} ||\omega||_R^2 & \text{if } \omega \in R; \\ \infty & \text{otherwise.} \end{cases}
\]
Theorem 2.3.5 [Generalized Schilder] Let \( A_i(\cdot) \in \Omega \) be i.i.d. centered Gaussian processes, with variance function \( \nu(\cdot) \) satisfying Assumptions A1 and A2. Then the sequence \( A_1(\cdot), A_2(\cdot), \ldots \) obeys the LDP with rate function \( I(\cdot) \).

Recall that an LDP consists of an upper and lower bound, which apply to closed and open sets, respectively. We will use Theorem 2.3.5 for certain open sets (to be defined in the next chapters). It can be verified that these sets \( U \) are such that

\[
\inf_{\omega \in U} I(\omega) = \inf_{\omega \in \overline{U}} I(\omega),
\]

where \( \overline{U} \) is the closure of \( U \). The way to prove this is to show that an arbitrarily chosen path in \( \overline{U} \) can be approximated by a path in \( U \), see [142] and Appendix A of [130].

Now consider the probability that the sample mean path of \( n \) i.d.d. Gaussian processes is contained in some set of paths \( E \). Then ‘Schilder’ yields the approximation

\[
P \left( \frac{1}{n} \sum_{i=1}^{n} A_i(\cdot) \in E \right) \approx \exp \left( -n \inf_{f \in E} I(f) \right).
\]  \hspace{1cm} (2.13)

Hence, the decay rate is dominated by the path in the set \( E \) that minimizes \( I(f) \), i.e., the path \( f^* = \arg\inf_{f \in E} I(f) \). We refer to \( f^* \) as the most probable path (MPP), as the decay rate of (2.13) is fully determined by the likelihood of this most likely path in \( E \). That is, given that the sample mean path is contained in the set \( E \), with overwhelming probability this happens by a path that is close to \( f^* \).

A problem that arises is that, as we saw above, there is only an explicit expression for \( I(f) \) available if \( f \) corresponds to a linear combination of covariance functions. Another difficulty is that the optimization should be performed over all paths \( f \in E \), which are infinitely dimensional objects. Nevertheless, if we find such a minimizing path \( f^* \), then this is useful in order to gain insight into the dynamics of a problem. In Section 2.5 we explicitly derive the MPPs in a simple system.

There exists also a version of Schilder’s theorem relating to multi-dimensional Gaussian processes. In particular, we will use the framework that corresponds to two-dimensional Gaussian processes in Chapters 3 and 4. The formulation of this framework is nearly identical to the above, but more involved, and is therefore left out.

### 2.4 Gaussian queues

Consider the process \{\( A(t) - ct, t \geq 0 \}\), where \( A(t) \) is a Gaussian process and \( c > 0 \) is a scalar. The reflection of this process at zero is referred to as a Gaussian queue. Due to the stationary increments, it is clear that a sufficient condition for stability of this system is that \( \mu < c \). In Chapter 1 we already argued that the steady-state
buffer content of such a queue can be represented as

\[ Q := \sup_{t \geq 0} \{ A(t) - ct \}, \]

given that this stability condition is satisfied.

As mentioned before, an inherent conceptual problem of Gaussian queues is that the input process can be negative. However, irrespective of whether \( A(t) \) corresponds to negative traffic or not, \( Q \) can always be evaluated and lives on \([0, \infty)\).

We remark that Gaussian queues are in general hard to analyze. In particular, only the cases of the Brownian motion and the Brownian bridge (that is, a standard Brownian motion conditioned on \( B(1) = 0 \)) result in explicit expressions for the steady-state buffer content distribution, see Section 2.5. To gain insight, one often resorts to either approximations or asymptotics.

### 2.4.1 Approximation

As the steady-state buffer content distribution of Gaussian queues is intractable in general, this has motivated the derivation of approximations for the situation of a general correlation structure. Let us focus on the overflow probability \( \mathbb{P}(Q > b) \), with \( b \geq 0 \). In e.g. [65, 127] the following approximation was suggested:

\[ \mathbb{P}(Q > b) \approx \exp \left( -\inf_{t \geq 0} \frac{(b + (c - \mu)t)^2}{2v(t)} \right). \]

The above approximation is obtained by using that \( A(t) \sim N(\mu t, v(t)) \), and subsequently applying the Chernoff bound. It is noted that the analysis of Chapter 5 is based on this approximation. Interestingly, it turns out that the approximation is exact for the Brownian motion and the Brownian bridge, see Example 5.4.2 in [127].

### 2.4.2 Asymptotics

The relevance of asymptotics can best be illustrated by considering two examples of interest. We already mentioned in Chapter 1 that both packet losses (due to buffer overflow) and packet delay strongly determine the QoS as perceived by users. Particularly for data applications, the loss is only allowed to exceed some specific value with extremely small probability. Hence, the (exponential) decay rate of the loss probability is an important performance measure, as it can be used to approximate the loss probability. Similarly, for most real-time applications the delay can only
Gaussian methodology

exceed some specific threshold with extremely small probability, implying that the decay rate of the delay probability is also a useful measure.

Asymptotics may clearly serve as approximations of the probabilities of interest, and they have the important additional advantage that they often provide useful qualitative insights, while they remain computationally tractable. Recall that asymptotics are closely connected to the probabilities of rare events. Typically, the most likely way for a rare event to occur is fairly simple, and can be directly deduced from the results, as will be illustrated in the next section.

Two types of asymptotics are widely used, namely: large-buffer asymptotics and many-sources asymptotics. Within each of these two regimes, we also distinguish between logarithmic and exact asymptotics. Below we briefly discuss each of the four cases of asymptotics. As a side remark we mention that in practice most (real-time) applications do not tolerate large delays, hence the large-buffer asymptotics are not always appropriate. It can be argued that in those situations the many-sources asymptotic regime is more justified.

Logarithmic large-buffer asymptotics

In order to find the logarithmic large-buffer asymptotics, we need to derive a function $f_1(b) \in \mathbb{R}_+$, such that

$$\log P(Q > b) \sim -f_1(b), \quad b \to \infty,$$

(2.14)
i.e., we need to find the decay rate.

Exact large-buffer asymptotics

In case the logarithmic large-buffer asymptotics are characterized, i.e., $f_1(b)$ is known, it follows from (2.14) that

$$P(Q > b) \sim g_1(b)e^{-f_1(b)}, \quad b \to \infty,$$

where the function $g_1(b)$ is such that

$$\lim_{b \to \infty} \frac{\log g_1(b)}{f_1(b)} = 0.$$

If the functions $f_1(b)$ and $g_1(b)$ are both explicitly found, then we say that one has determined the exact large-buffer asymptotics. It is clear that the exact asymptotics are more refined than the logarithmic asymptotics, i.e., if the exact asymptotics are known, then they effectively also yield the logarithmic asymptotics. Exact asymptotics are often considerably harder to obtain though.
2.5 Brownian queues

In the remainder of this chapter we consider the reflection of the process \( \{B(t) - ct, t \geq 0\} \), where \( B(t) \) is a standard Brownian motion (with \( B(0) \equiv 0 \)), denoting the amount of traffic entering the system in the interval \([0, t]\), and \( c > 0 \) is the service capacity of the node. The reflection of \( \{B(t) - ct, t \geq 0\} \) at zero is referred to as a Brownian queue, which is a special kind of Gaussian queue, see Figure 2.1 for an illustration.

Ordinary Brownian input plays an important role in this monograph, as the use of Brownian input often results in explicit expressions for performance measures, thereby providing valuable insight. Brownian motions can be used to approximate weakly-dependent traffic streams, cf. also the celebrated ‘Central Limit Theorem in functional form’. Its mean and variance function are characterized through \( \mu = 0 \) and \( \nu(t) = t \), respectively. It can be verified that \( \Gamma(s, t) = \text{Cov}(B(s), B(t)) = \min\{|s|, |t|\} \) if \( s, t \geq 0 \) or \( s, t < 0 \), and \( \Gamma(s, t) = 0 \) otherwise. Let \( B(s, t) = B(t) - B(s) \) denote the amount of traffic generated in the interval \([s, t]\), \( s < t \). The goal of this section is to show how
the machinery that was presented earlier in this chapter can be used. We remark that some of the results that are derived below are already known, as the Brownian queue has been well-studied in the past, see e.g. [2, 3, 4, 75]. However, those results were obtained in a completely different and perhaps less transparent manner.

2.5.1 Useful properties

We already mentioned that the steady-state buffer content distribution of a Brownian queue is tractable. In fact, in [156] it was shown that it is exponentially distributed with mean $1/(2c)$. That is,

$$P(Q \leq b) = P \left( \sup_{t \geq 0} \{B(-t, 0) - ct\} \leq b \right) = P \left( \sup_{t \geq 0} \{B(0, t) - ct\} \leq b \right) = P \left( \forall t \geq 0 : B(t) \leq b + ct \right) = 1 - e^{-2bc}, \quad (2.16)$$

with $b, c \geq 0$, i.e., the probability that a standard Brownian motion stays below the function $b + ct$ for all $t \geq 0$, equals $1 - \exp(-2bc)$.

Another useful property is that

$$P(\forall t \in [0, 1] : B(t) \leq b + ct | B(1) = 0) = 1 - e^{-2b(b+c)}, \quad (2.17)$$

with $b, c \geq 0$, i.e., the probability that a Brownian bridge stays below the function $b + ct$, for all $t \in [0, 1]$, equals $1 - \exp(-2b(b+c))$. We can exploit (2.17) to derive that

$$P(\forall t \in [0, u] : B(t) \leq b + ct | B(u) = x) = P(\forall s \in [0, 1] : B(su) \leq b + csu | B(u) = x) = P(\forall s \in [0, 1] : \sqrt{u}B(s) \leq b + csu | \sqrt{u}B(1) = x) = P(\forall s \in [0, 1] : B(s) \leq \frac{b}{\sqrt{u}} + cs\sqrt{u} | B(1) = \frac{x}{\sqrt{u}}) = P(\forall s \in [0, 1] : B(s) \leq \frac{b}{\sqrt{u}} + \frac{cs\sqrt{u} - \frac{x}{\sqrt{u}}}{\sqrt{u}} | B(1) = 0),$$

$$= P(\forall s \in [0, 1] : B(s) \leq \frac{b}{\sqrt{u}} + \left(c\sqrt{u} - \frac{x}{\sqrt{u}}\right)s | B(1) = 0) = 1 - e^{-2\frac{b}{\sqrt{u}}(\frac{1}{2\sqrt{u}} + c\sqrt{u} - \frac{x}{\sqrt{u}})}, \quad (2.18)$$

with $b, c, u \geq 0$ and $x \in [0, b + cu]$, where we use that a standard Brownian motion is self-similar with $H = 1/2$. The above results can also easily be extended to general Brownian input, with drift $\mu > 0$ and variance $\nu(t) = \lambda t$, $\lambda > 0$, as we will see below and in the next chapters.
2.5 Brownian queues

The rate function given in Equation (2.12) simplifies considerably in case of standard Brownian input. Using (2.11) and the definition of \( \Gamma(s, t) \) for standard Brownian input, it can be shown (see Theorem 5.2.3 of [52]) that (2.12) is equivalent to

\[
I(\omega) = \begin{cases} 
\frac{1}{2} \int_{-\infty}^{\infty} (\omega'(t))^2 dt & \text{if } \omega \in R; \\
\infty & \text{otherwise.} 
\end{cases}
\] (2.19)

In the remainder of this section we analyze the transient behavior of a Brownian queue. In particular, we explicitly derive the joint distribution function

\[ p(b, T) = \mathbb{P}(Q_{0} > b_{0}, Q_{T} > b_{T}), \]

where \( b_{0}, b_{T} \geq 0, \) \( b = (b_{0}, b_{T}), \) \( T > 0, \) and \( Q_{t} \) denotes the workload at time \( t \geq 0, \) assuming that the workload process is in stationarity at \( t = 0. \) This also allows us to explicitly calculate the covariance of \( Q_{0} \) and \( Q_{T}. \) By setting \( b_{0} = b, b_{T} = \alpha b, \) and \( T = \gamma b, \) with \( \alpha, \gamma \geq 0, \) and letting \( b \to \infty, \) we also obtain the exact large-buffer asymptotics, i.e., we find a function \( f(\cdot) \) such that

\[
\mathbb{P}(Q_{0} > b_{0}, Q_{\gamma b} > \gamma b) / f(b) \to 1 \quad \text{as} \quad b \to \infty.
\]

It turns out that the nature of the asymptotics depends on the value of \( \alpha, \gamma, \) and the service rate \( c \) of the queue, i.e., there are various regimes. These regimes can be further interpreted relying on Schilder’s sample-path large deviations theorem. In particular, we obtain the MPP, i.e., the most likely way for the buffer to fill.

2.5.2 Joint distribution function

In this subsection we derive a closed-form expression for \( p(b, T). \) It turns out that it is easier to first calculate \( p(b, T) := \mathbb{P}(Q_{0} \leq b_{0}, Q_{T} \leq b_{T}). \) Recall that \( \Phi(\cdot) \) denotes the distribution function of a standard Normal random variable. According to Reich’s formula [155],

\[
Q_{0} = \sup_{t \geq 0} \{B(-t, 0) - ct\} \quad \text{and} \quad Q_{T} = \sup_{s \geq 0} \{B(T - s, T) - cs\}. \] (2.20)

Hence, we find that

\[
p(b, T) = \mathbb{P}\left( \sup_{t \geq 0} \{B(-t, 0) - ct\} \leq b_{0}, \sup_{s \geq 0} \{B(T - s, T) - cs\} \leq b_{T} \right)
= \mathbb{P}(\forall s, t \geq 0 : B(-t, 0) \leq b_{0} + ct, B(T - s, T) \leq b_{T} + cs)
= \mathbb{P}(\forall s, t \geq 0 : B(T, t + T) \leq b_{0} + ct, B(0, s) \leq b_{T} + cs),
\]

where the last line is obtained by using time reversibility of Brownian motion. Now, conditioning on the value of \( B(0, T), \) we obtain that

\[
p(b, T) = \int_{-\infty}^{b_{T} + ct} \mathbb{P}(\forall s \in [0, T) : B(0, s) \leq b_{T} + cs | B(0, T) = x) \times \mathbb{P}(\forall t \geq 0 : \forall s \geq T : B(T, t + T) \leq b_{0} + ct, B(0, s) \leq b_{T} + cs | B(0, T) = x) \, dx .
\]
Let us first focus on the first term in the above integral. Using (2.18), we obtain that
\[ P(\forall s \in [0, T) : B(0, s) \leq b_T + csB(0, T) = x) = 1 - \exp\left(-2b_Tc - 2b_T(b_T - x)/T\right). \] (2.21)

Proceeding with the second term in the integral, we find that
\[ P(\forall t \geq 0 : \forall s \geq T : B(T, T + t) \leq b_0 + ct, B(0, s) \leq b_T + cs|B(0, T) = x) = \]
\[ = P(\forall s, t \geq 0 : B(T, T + t) \leq b_0 + ct, B(T, T + s) \geq b_T + (s + T)c - x) = \]
\[ = P(\forall s, t \geq 0 : B(0, t) \leq b_0 + ct, B(0, s) \leq b_T + (s + T)c - x) = \]
\[ = P(\forall t \geq 0 : B(0, t) \leq \min\{b_0, b_T + cT - x\} + ct). \]

Exploiting (2.16), we deduce that
\[ P(\forall t \geq 0 : \forall s \geq T : B(T, T + t) \leq b_0 + ct, B(0, s) \leq b_T + cs|B(0, T) = x) = \]
\[ = P(\forall t \geq 0 : B(0, t) \leq \min\{b_0, b_T + cT - x\} + ct) = \]
\[ = \begin{cases} 
1 - \exp(-2b_0c) & \text{if } x \leq b_T + cT - b_0; \\
1 - \exp(-2(b_T + cT - x)c) & \text{if } x > b_T + cT - b_0. 
\end{cases} \] (2.22)

**Theorem 2.5.1** For each \( b_0, b_T, T \geq 0, \)
\[ p(\overline{b}, T) = -\Phi\left(k_1(\overline{b}, T)\right) + e^{-2b_Tc}\Phi\left(k_2(\overline{b}, T)\right) + e^{-2b_0c}\Phi\left(k_3(\overline{b}, T)\right) + e^{-2(b_0+b_T)c}\Phi\left(k_4(\overline{b}, T)\right), \]
where
\[ k_1(\overline{b}, T) = \frac{-b_T - cT - b_0}{\sqrt{T}}; \quad k_2(\overline{b}, T) = \frac{b_T - cT - b_0}{\sqrt{T}}; \]
\[ k_3(\overline{b}, T) = \frac{-b_T - cT + b_0}{\sqrt{T}}; \quad k_4(\overline{b}, T) = \frac{-b_T + cT - b_0}{\sqrt{T}}. \]

**Proof:** Using (2.21) and (2.22), we obtain that \( p(\overline{b}, T) \) equals
\[ \int_{-\infty}^{b_T - cT - b_0} \left(1 - \exp\left(-2b_Tc - 2\frac{b_T(b_T - x)}{T}\right)\right) \times \]
\[ (1 - \exp(-2b_0c)) \, dP(N(0, T) \leq x) + \]
\[ \int_{b_T + cT - b_0}^{b_T + cT - b_0} \left(1 - \exp\left(-2b_Tc - 2\frac{b_T(b_T - x)}{T}\right)\right) \times \]
\[ (1 - \exp(-2(b_T + cT - x)c)) \, dP(N(0, T) \leq x). \]
It is a straightforward exercise to show that the first integral is equal to
\[
(1 - \exp(-2b_0c)) \left( \Phi \left( \frac{b_T + cT - b_0}{\sqrt{T}} \right) - \exp(-2b_Tc) \Phi \left( \frac{-b_T + cT - b_0}{\sqrt{T}} \right) \right),
\]
whereas the second integral equals
\[
1 - \Phi \left( \frac{-b_T - cT - b_0}{\sqrt{T}} \right) - \Phi \left( \frac{b_T + cT - b_0}{\sqrt{T}} \right) + 
\exp(-2b_Tc) \left( \Phi \left( \frac{-b_T + cT - b_0}{\sqrt{T}} \right) + \Phi \left( \frac{b_T - cT - b_0}{\sqrt{T}} \right) - 1 \right).
\]
Using that \( \Phi(Q_i \leq b_i) = 1 - \exp(-2b_i c), i = 0, T, \) see (2.16), and that \( 1 - \Phi(x) = \Phi(-x), \) the stated follows from
\[
p(b, T) = 1 - P(Q_0 \leq b_0) - P(Q_T \leq b_T) + \Phi(b, T).
\]

2.5.3 Covariance function

In the previous subsection we derived a closed-form expression for \( p(b, T). \) This result also allows us to calculate the covariance of \( Q_0 \) and \( Q_T, \) i.e., \( \text{Cov}(Q_0, Q_T), \) which we present in the next theorem.

**Theorem 2.5.2** For each \( T \geq 0, \)
\[
\theta(T) := \text{Cov}(Q_0, Q_T) \tag{2.23}
\]
\[
= \left( -\frac{c^2 T^2}{2} - T + \frac{1}{2c^2} \right) \left( 1 - \Phi(c\sqrt{T}) \right) + \phi(c\sqrt{T}) \left( \frac{cT\sqrt{T}}{2} + \frac{\sqrt{T}}{2c} \right).
\]

*Proof:* First recall that \( \text{Cov}(Q_0, Q_T) = \text{E}Q_0Q_T - \text{E}Q_0\text{E}Q_T. \) Then use the well-known fact that \( Q_0 \) and \( Q_T \) are both exponentially distributed with mean \( 1/(2c), \) i.e., \( \text{E}Q_0\text{E}Q_T = 1/(4c^2). \) Hence, we are left with the computation of \( \text{E}Q_0Q_T. \) Using Theorem 2.5.1, we find that
\[
\text{E}Q_0Q_T = \int_0^\infty \int_0^\infty p(b, T)db_0db_T
\]
\[
= -\int_0^\infty \int_0^\infty \Phi(k_1(b, T)) \, db_0db_T + \int_0^\infty \int_0^\infty e^{-2b_Tc} \Phi(k_2(b, T)) \, db_0db_T
\]
\[
+ \int_0^\infty \int_0^\infty e^{-2b_0c} \Phi(k_3(b, T)) \, db_0db_T
\]
\[
+ \int_0^\infty \int_0^\infty e^{-2(b_0+b_T)c} \Phi(k_4(b, T)) \, db_0db_T.
\]
By interchanging the order of integration, and applying integration by parts, straightforward (though tedious) calculus yields that
\[
-\int_0^\infty \int_0^\infty \Phi(k_1(b, T)) \, db_0db_T \tag{2.24}
\]
\[= - \left( \frac{T}{2} + \frac{c^2T^2}{2} \right) \left( 1 - \Phi(c\sqrt{T}) \right) + \frac{cT\sqrt{T}}{2} \phi(c\sqrt{T}); \]
\[
\int_0^\infty \int_0^\infty e^{-2bTc} \Phi \left( k_2(b, T) \right) db_0 db_T \]
\[= \left( \frac{1}{2c^2} - \frac{T}{2} \right) \left( 1 - \Phi(c\sqrt{T}) \right) + \frac{\sqrt{T}}{2c} \phi(c\sqrt{T}); \]
\[
\int_0^\infty \int_0^\infty e^{-2b_0c} \Phi \left( k_3(b, T) \right) db_0 db_T \]
\[= \left( \frac{1}{2c^2} - \frac{T}{2} \right) \left( 1 - \Phi(c\sqrt{T}) \right) + \frac{\sqrt{T}}{2c} \phi(c\sqrt{T}); \]
\[
\int_0^\infty \int_0^\infty e^{-2(b_0+cT)c} \Phi \left( k_4(b, T) \right) db_0 db_T \]
\[= \left( \frac{T}{2} - \frac{1}{4c^2} \right) \left( 1 - \Phi(c\sqrt{T}) \right) + \frac{1}{4c^2} \Phi(c\sqrt{T}) - \frac{\sqrt{T}}{2c} \phi(c\sqrt{T}). \]  

(2.25)

Adding up (2.24)-(2.27), and subtracting 1/(4c^2) yields the stated. \[\square\]

It is noted that \(\theta(0) = \text{Var} Q_0 = 1/(4c^2)\), which is equivalent to the variance of an exponentially distributed variable with mean 1/(2c), as required. Also, note that \(\lim_{T \to \infty} \theta(T) \to 0\), as \(Q_0\) and \(Q_T\) become less and less correlated as \(T \to \infty\). The following proposition summarizes three properties of \(\theta(\cdot)\). This proposition implies that \((1 - \theta(\cdot))/\text{Var} Q_0\) is a distribution function on \([0, \infty)\).

**Proposition 2.5.3** \(\theta(\cdot)\) is non-increasing, convex and non-negative on \([0, \infty)\).

**Proof:** \(\theta(T)\) is non-increasing on \([0, \infty)\) if \(\theta'(T) \leq 0\), i.e.,
\[\frac{\phi \left( c\sqrt{T} \right)}{1 - \Phi \left( c\sqrt{T} \right)} \leq c\sqrt{T} + \frac{1}{c\sqrt{T}}; \]

(2.28)

which is equivalent to
\[\frac{\phi \left( c\sqrt{T} \right)}{1 - \Phi \left( c\sqrt{T} \right)} \leq c\sqrt{T} + \frac{1}{c\sqrt{T}}. \]

Likewise, \(\theta(T)\) is convex on \([0, \infty)\) if \(\theta''(T) \geq 0\), i.e.,
\[c^2 \left( 1 - \Phi(c\sqrt{T}) \right) + \frac{c}{\sqrt{T}} \phi \left( c\sqrt{T} \right) \geq 0, \]

or equivalently,
\[\frac{\phi \left( c\sqrt{T} \right)}{1 - \Phi \left( c\sqrt{T} \right)} \geq c\sqrt{T}. \]  

(2.29)
Recalling the standard equality (2.1), it is easily seen that both (2.28) and (2.29) hold. Since \( \theta(T) \) is non-increasing and \( \lim_{T \to \infty} \theta(T) \to 0 \), we also must have that \( \theta(T) \) is non-negative.

The next proposition presents the exact asymptotics of \( \theta(T) \).

**Proposition 2.5.4** If \( T \to \infty \),

\[
\theta(T) \sim \frac{4}{c^5T\sqrt{T}} \phi \left( c\sqrt{T} \right).
\]  

**Proof:** First use that \([5]\)

\[
(1 - \Phi(g(x))) \sim \left( \frac{1}{g(x)} - \frac{1}{(g(x))^3} + \frac{3}{(g(x))^5} - \frac{15}{(g(x))^7} \right) \phi(g(x))
\]  

if \( g(x) \) is increasing and \( x \to \infty \). Using (2.31) and Theorem 2.5.2, it can then be verified that

\[
\theta(T) \sim \left( \frac{4}{c^5T\sqrt{T}} + \frac{16\frac{1}{2}}{c^3T^2\sqrt{T}} - \frac{7\frac{1}{2}}{c^3T^3\sqrt{T}} \right) \phi \left( c\sqrt{T} \right) \sim \frac{4}{c^5T\sqrt{T}} \phi \left( c\sqrt{T} \right).
\]

We note that the correct exact asymptotics of \( \theta(T) \) can only be obtained, if all four terms of the right-hand side of (2.31) are used.

**Remark:** The correlation coefficient of \( Q_0 \) and \( Q_T \) is given by

\[
\rho(T) := \text{Cor}(Q_0, Q_T) = \frac{\text{Cov}(Q_0, Q_T)}{\sqrt{\text{Var}Q_0}\sqrt{\text{Var}Q_T}} = 4c^2\theta(T),
\]

as both \( Q_0 \) and \( Q_T \) are exponentially distributed with mean \( 1/(2c) \). Note that \( \rho(0) = 1 \) and \( \lim_{T \to \infty} \rho(T) \to 0 \). Due to (2.32), we also have that \( \rho(T) \) is non-increasing, convex and non-negative on \( [0, \infty) \), and that

\[
\rho(T) \sim \frac{16}{c^3T\sqrt{T}} \phi \left( c\sqrt{T} \right).
\]

Hence, the exponential decay rate of both \( \theta(T) \) and \( \rho(T) \) equals \( (c^2T)/2 \).

It is noted that Theorem 2.5.2 and Propositions 2.5.3 and 2.5.4 have already (partly) appeared (for \( \rho(T) \), instead of \( \theta(T) \)) in [4]. However, it is noted that our derivations are completely different compared to the ones given in [4]. We rely on Reich’s formula to obtain the results, whereas [4] does not use this formula implicitly. It turns out that Proposition 2.5.3 also extends to the class of Lévy inputs, i.e., arrival processes with stationary, independent increments, see Theorem 3.6 in [61]. This class comprises, besides Brownian input, also compound Poisson input as special case.
2.5.4 Exact large-buffer asymptotics

In this subsection we derive the exact asymptotics of \( p(\tilde{b}, T) \). We first present the following useful lemma.

**Lemma 2.5.5** Let \( b_0 = b \), \( b_T = \alpha b \) and \( T = \gamma b \), with \( \alpha, \gamma \geq 0 \). If \( b \to \infty \), then

\[
\begin{align*}
\Phi(k_1(\tilde{b}, T)) &\sim -\zeta(k_1(\tilde{b}, T)) ; \\
\Phi(k_2(\tilde{b}, T)) &\sim \begin{cases} 
-\zeta(k_2(\tilde{b}, T)) & \text{if } \alpha < 1 + c\gamma; \\
1/2 & \text{if } \alpha = 1 + c\gamma; \\
1 & \text{otherwise};
\end{cases} \\
\Phi(k_3(\tilde{b}, T)) &\sim \begin{cases} 
-\zeta(k_3(\tilde{b}, T)) & \text{if } \alpha > 1 - c\gamma; \\
1/2 & \text{if } \alpha = 1 - c\gamma; \\
1 & \text{otherwise};
\end{cases} \\
\Phi(k_4(\tilde{b}, T)) &\sim \begin{cases} 
-\zeta(k_4(\tilde{b}, T)) & \text{if } \alpha > c\gamma - 1; \\
1/2 & \text{if } \alpha = c\gamma - 1; \\
1 & \text{otherwise},
\end{cases}
\end{align*}
\]

where \( \zeta(\cdot) \) is as defined in (2.2).

**Proof:** First determine for which values of \( b_T/b_0 = \alpha, k_i(\tilde{b}, T), i \in \{1, 2, 3, 4\} \), is positive or negative. Note that \( k_1(\tilde{b}) \) is always negative. Hence, we obtain \( 1 + c\gamma, 1 - c\gamma \) and \( c\gamma - 1 \) as critical values from \( k_i(\tilde{b}), i = 2, 3, 4 \), respectively. Next use the fact that \( \Phi(-u) \sim \zeta(u) \) and \( \Phi(u) \sim 1 \) as \( u \to \infty \). Observe that \( \Phi(0) = 1/2 \).

We remark that \( -\zeta(k_i(\tilde{b}, T)) \) is positive in Lemma 2.5.5, as \( \zeta(k_i(\tilde{b}, T)) \) is negative in the listed cases, \( i = 1, \ldots, 4 \). Define

\[
\gamma(\tilde{b}, T) := 2b_0c + \frac{(-b_T - cT + b_0)^2}{2T}.
\]

**Theorem 2.5.6** Let \( b_0 = b, b_T = \alpha b, T = \gamma b \), with \( \alpha, \gamma \geq 0 \). Suppose \( c\gamma > 1 \).

For \( b \to \infty \),

\[
p(\tilde{b}, T) \sim \begin{cases} 
e^{-2(b_0 + b_T)c} & \text{if } 0 \leq \alpha < \left(\sqrt{c\gamma} - 1\right)^2; \\
\left(1 - \frac{1}{\sqrt{2\pi k_1(\tilde{b}, T)}} - \frac{1}{\sqrt{2\pi k_3(\tilde{b}, T)}}\right)e^{-2(b_0 + b_T)c} & \text{if } \alpha = \left(\sqrt{c\gamma} - 1\right)^2; \\
\left(1 - \frac{1}{\sqrt{2\pi k_2(\tilde{b}, T)}} - \frac{1}{\sqrt{2\pi k_3(\tilde{b}, T)}}\right)e^{-\gamma(\tilde{b}, T)} & \text{if } \left(\sqrt{c\gamma} - 1\right)^2 < \alpha < 1 + c\gamma; \\
\left(\frac{1}{2} - \frac{1}{\sqrt{2\pi k_3(\tilde{b}, T)}}\right)e^{-2b_Tc} & \text{if } \alpha = 1 + c\gamma; \\
e^{-2b_Tc} & \text{if } \alpha > 1 + c\gamma.
\end{cases}
\]
Proof: We only prove the last statement, as the other four statements follow in a similar way. We have to prove that
\[ p(\bar{b}, T)e^{2b_T c} \to 1 \text{ as } b \to \infty, \text{ for } \alpha > 1 + c\gamma. \]

From Lemma 2.5.5 we obtain that for \( \alpha > 1 + c\gamma \),
\[ \Phi(k_1(\bar{b}, T)) \sim -\zeta(k_1(\bar{b}, T)); \quad \Phi(k_3(\bar{b}, T)) \sim -\zeta(k_3(\bar{b}, T)). \]

Now straightforward calculus shows that, as \( b \to \infty \),
\[ \Phi(k_1(\bar{b}, T)) = o(e^{-2b_0 c}) \]
and the same applies for \( \Phi(k_3(\bar{b}, T))e^{-2b_0 c} \) and \( \Phi(k_4(\bar{b}, T))e^{-2(b_0 + b_T)c} \). Using that \( \Phi(k_2(\bar{b}, T)) \sim 1 \), Theorem 2.5.1 implies the stated. \( \Box \)

The following two theorems can be proven in a similar fashion as Theorem 2.5.6.

**Theorem 2.5.7** Let \( b_0 = b, \quad b_T = \alpha b, \quad T = \gamma b \), with \( \alpha, \gamma \geq 0 \). Suppose \( c\gamma = 1 \).

For \( b \to \infty \),
\[ p(\bar{b}, T) \sim \begin{cases} e^{-2b_0 c} & \text{if } \alpha = 0; \\ \left( \frac{1}{\sqrt{2\pi k_2(\bar{b}, T)}} - \frac{1}{\sqrt{2\pi k_3(\bar{b}, T)}} \right) e^{-\gamma(\bar{b}, T)} & \text{if } 0 < \alpha < 1 + c\gamma; \\ \frac{1}{2} - \frac{1}{\sqrt{2\pi k_3(\bar{b}, T)}} e^{-2b_T c} & \text{if } \alpha = 1 + c\gamma; \\ e^{-2b_T c} & \text{if } \alpha > 1 + c\gamma. \end{cases} \]

**Theorem 2.5.8** Let \( b_0 = b, \quad b_T = \alpha b, \quad T = \gamma b \), with \( \alpha, \gamma \geq 0 \). Suppose \( c\gamma < 1 \).

For \( b \to \infty \),
\[ p(\bar{b}, T) \sim \begin{cases} e^{-2b_0 c} & \text{if } 0 \leq \alpha < 1 - c\gamma; \\ \left( \frac{1}{2} - \frac{1}{\sqrt{2\pi k_2(\bar{b}, T)}} \right) e^{-2b_0 c} & \text{if } \alpha = 1 - c\gamma; \\ \left( \frac{1}{\sqrt{2\pi k_2(\bar{b}, T)}} - \frac{1}{\sqrt{2\pi k_3(\bar{b}, T)}} \right) e^{-\gamma(\bar{b}, T)} & \text{if } 1 - c\gamma < \alpha < 1 + c\gamma; \\ \frac{1}{2} - \frac{1}{\sqrt{2\pi k_3(\bar{b}, T)}} e^{-2b_T c} & \text{if } \alpha = 1 + c\gamma; \\ e^{-2b_T c} & \text{if } \alpha > 1 + c\gamma. \end{cases} \]

**2.5.5 Most probable path**

In the previous subsection it was shown that the nature of the large-buffer asymptotics strongly depends on the model parameters \( \alpha, \gamma \), and \( c \), i.e., there are multiple regimes. In this subsection we will interpret these regimes by exploiting sample-path large
deviations results. Schilder’s theorem, as introduced in Section 2.3, implies that the
exponential decay rate of the joint overflow probability is characterized by the path
that minimizes the decay rate. Among all paths such that the buffer exceeds \( b_0 \) and
\( b_T \) at time 0 and \( T \) respectively, this is the MPP: informally speaking, given that this
rare event occurs, with overwhelming probability \((b_0, b_T)\) is reached by a path ‘close
to’ the MPP.

In order to apply ‘Schilder’, we feed the single-node network by \( n \) i.i.d. standard
Brownian sources. The link rate and buffer thresholds are also scaled by \( n \): \( nc \), \( nb_0 \)
and \( nb_T \), respectively. Using (2.20), \( p_n(\bar{b}, T) \) can be expressed as
\[
P \left( \frac{1}{n} \sum_{i=1}^{n} B_i(\cdot) \in S \right),
\]
where
\[
S := \{ f \in \Omega | \exists s, t \geq 0 : -f(-t) > b_0 + ct, f(T) - f(T - s) > b_T + cs \},
\]
and \( \Omega \) is as defined in Equation (2.10).

We already argued in Section 2.3 that we can replace ‘\( > \)’ by ‘\( \geq \)’ in \( S \), which
is denoted as the set \( \overline{S} \), without any impact on the decay rate of \( p_n(\bar{b}, T) \). From
‘Schilder’ it then follows that
\[
J(\bar{b}, T) := -\lim_{n \to \infty} \frac{1}{n} \log p_n(\bar{b}, T) = \inf_{f \in \overline{S}} I(f) = \inf_{f \in \overline{S}} I(f).
\]
(2.33)
As we will see below, depending on the value of \( b_0, b_T, c, \) and \( T \), various regimes of
asymptotics exist. Recall from Section 2.3 that knowledge of the MPP automatically
implies that the decay rate is characterized, as the MPP translates in the decay rate
through Equation (2.19). In the remainder of this subsection we explicitly derive
\( J(\bar{b}, T) \) by determining the MPPs corresponding to the various regimes.

Let us first define
\[
\overline{U} := \{ f \in \Omega | \exists t \geq 0 : -f(-t) \geq b_0 + ct \};
\]
\[
\overline{V} := \{ f \in \Omega | \exists s \geq 0 : f(T) - f(T - s) \geq b_T + cs \},
\]
i.e., \( \overline{U} (\overline{V}) \) is the collection of all paths that yield a buffer content of at least \( b_0 \) \( (b_T) \)
at time 0 \( (T) \). It follows that \( \overline{S} \) is a subset of both \( \overline{U} \) and \( \overline{V} \), i.e., \( \overline{S} \subseteq \overline{U} \) and \( \overline{S} \subseteq \overline{V} \),
implying that
\[
J(\bar{b}, T) \geq \inf_{f \in \overline{U}} I(f); \quad J(\bar{b}, T) \geq \inf_{f \in \overline{V}} I(f).
\]
(2.33)
From the above it follows that there is equality in one of the inequalities of (2.33), if
either the MPP in \( \overline{U} \) or \( \overline{V} \) (or both) is also contained in the set \( \overline{S} \).
Fortunately, the MPPs in $\mathcal{U}$ and $\mathcal{V}$ are already available, see e.g. [7]. The MPP in $\mathcal{U}$ is given by, for $r \in [-b_0/c, 0]$,

$$f^*(r) = \mathbb{E}(B(r)) - B(-b_0/c) = 2b_0.$$  

The MPP is only specified on the interval $[-b_0/c, 0]$, because outside this interval the MPP generates traffic with mean rate $\mu = 0$. Using (2.4), it can then be verified that, for $r \in [-b_0/c, 0]$, $(f^*)'(r) = 2c$, whereas the derivative of this path is equal to zero outside this interval. In other words, the buffer starts to build up with constant rate $2c - c = c$ at time $-b_0/c$, which leads to $Q_0 = (b_0/c)c = b_0$, as required. Let us now determine the cost of this MPP. Using (2.19), we find that

$$I(f^*) = \frac{1}{2} \frac{b_0}{c} (2c)^2 = 2b_0c.$$  

The MPP in $\mathcal{V}$ has a similar structure as the one above, but now the buffer grows with constant rate $c$ in the interval $[T - b_T/c, T]$, which eventually gives $Q_T = b_T$, as required. The cost of this path can be derived in a similar manner and equal $2b_Tc$.

We are now ready to provide some explanation for each of the regimes of Theorems 2.5.6-2.5.8. Let us start with the regime $\alpha \geq 1 + c\gamma$ in Theorems 2.5.6-2.5.8. Using that $\alpha = b_T/b_0$ and $\gamma = T/b_0$, it is easily seen that we can rewrite $\alpha \geq 1 + c\gamma$ as $b_T - cT \geq b_0$. Subsequently, it is straightforward to show that the MPP in $\mathcal{V}$ is also contained in $\mathcal{S}$ under this regime, i.e., overflow at time $T$ implies overflow at time 0 without any additional effort. As the MPP in the set $\mathcal{V}$ is contained in $\mathcal{S}$, it is also the MPP in the set $\mathcal{S}$. In other words, $J(\bar{b}, T)$ is equal to $2b_Tc$, given that $b_T - cT \geq b_0$. The MPP is depicted in Figure 2.2 (top, left).

Next consider the regime $0 \leq \alpha \leq 1 - c\gamma$ in Theorems 2.5.7-2.5.8, or equivalently $b_T \leq b_0 - cT$. In this case one can verify that the MPP in the set $\mathcal{U}$ is also contained in the set $\mathcal{S}$, and therefore it is the MPP in $\mathcal{S}$. Thus, overflow at time 0 implies overflow at time $T$ without any extra effort. We conclude that $J(\bar{b}, T)$ equals $2b_0c$, given that $b_T \leq b_0 - cT$. The MPP is depicted in Figure 2.2 (top, right).

We proceed with the regime $0 \leq \alpha \leq (\sqrt{c\gamma} - 1)^2$ in Theorem 2.5.6, or equivalently $T \geq (\sqrt{b_0} + \sqrt{b_T})^2/c$. Consider the path that is such that the buffer builds up with rate $c$ in the interval $[-b_0/c, 0]$, empties with rate $c$ in the interval $(0, b_0/c)$, is empty in the interval $[b_0/c, T - b_T/c)$, and is growing again with rate $c$ in the interval $[T - b_T/c, T]$, i.e., the MPP of $\mathcal{U}$ and $\mathcal{V}$ combined. It can be verified that this path is contained in the set $\mathcal{S}$ if $T \geq (\sqrt{b_0} + \sqrt{b_T})^2/c$. In Section 2.5.6 we show that this path is in fact the MPP in $\mathcal{S}$. In that case, $J(\bar{b}, T)$ can be obtained by using (2.19), and equals $2b_0c + 2b_Tc$. Clearly, this is no surprise, as the path consists of the MPP of $\mathcal{U}$ and $\mathcal{V}$. Note that this suggests that $Q_0$ and $Q_T$ behave (almost) independently if, compared to $b_0$ and $b_T$, $T$ is large enough, as may be expected. The MPP is depicted in Figure 2.2 (bottom, left).
We now focus on the remaining regimes of Theorems 2.5.6-2.5.8. Consider the path that is such that the buffer builds up with rate $c$ in the interval $[-b_0/c, 0]$, and grows with rate $(b_T - b_0)/T$ in the interval $(0, T]$. Clearly, this path yields $Q_0 = b_0$ and $Q_T = b_T$, and is thus contained in $\mathcal{S}$. In Section 2.5.6 we show that this path is in fact the MPP for the remaining regimes. Assuming that this is indeed the case, $J(b, T)$ is obtained by using (2.19), and equals $\gamma(b, T)$. The MPP is depicted in Figure 2.2 (bottom, right).

The following two theorems are presented without proof, as they summarize the above-mentioned statements.

**Theorem 2.5.9** Suppose $cT > b_0$. Then it holds that

$$J(\bar{b}, T) \sim \begin{cases} 
  2(b_0 + b_T)c & \text{if } 0 \leq b_T \leq \left(\sqrt{cT} - \sqrt{b_0}\right)^2; \\
  \gamma(\bar{b}, T) & \text{if } \left(\sqrt{cT} - \sqrt{b_0}\right)^2 < b_T < b_0 + cT; \\
  2b_Tc & \text{if } b_T \geq b_0 + cT.
\end{cases}$$

**Theorem 2.5.10** Suppose $cT \leq b_0$. Then it holds that

$$J(\bar{b}, T) \sim \begin{cases} 
  2b_0c & \text{if } 0 \leq b_T \leq b_0 - cT; \\
  \gamma(\bar{b}, T) & \text{if } b_0 - cT < b_T < b_0 + cT; \\
  2b_Tc & \text{if } b_T \geq b_0 + cT.
\end{cases}$$
2.5 Brownian queues

2.5.6 Discussion

Using Theorems 2.5.6-2.5.8, the logarithmic large-buffer asymptotics can easily be derived as well. That is, we need to find a function $J^*(\bar{b}_\alpha, T_\gamma)$, with $\bar{b}_\alpha \equiv (b, \alpha b)$ and $T_\gamma \equiv \gamma b$, such that

$$
\lim_{b \to \infty} -\log \mathbb{P}(Q_0 > b, Q_{\gamma b} > \alpha b) / J^*(\bar{b}_\alpha, T_\gamma) = 1,
$$

where $\alpha, \gamma \geq 0$. With $b = b_0$, $\alpha b = b_0 T$, i.e., $\bar{b}_\alpha = \bar{b}$, and $\gamma b = T_\gamma = T$, it is not hard to see that $J^*(\bar{b}_\alpha, T_\gamma)$ equals $J(\bar{b}, T)$; compare Theorems 2.5.6-2.5.8 with Theorems 2.5.9-2.5.10, respectively. Indeed, since we assumed that in the many-sources framework the standard Brownian sources are i.i.d., and because a standard Brownian motion is characterized by independent increments, $J^*(\bar{b}_\alpha, T_\gamma)$ and $J(\bar{b}, T)$ should match, see for instance Example 7.4 in [65]. Recall that in the previous subsection we argued that the paths depicted in Figure 2.2 are MPPs in the set $\mathcal{S}$. From the above we conclude that this is indeed correct.

In the analysis we assumed that the input process was a standard Brownian motion, i.e., no drift and $v(t) = t$. We now show how the results can be extended to general Brownian input, with drift $\mu > 0$ and variance $v(t) = \lambda t$, $\lambda > 0$. Clearly, we should have that $c > \mu > 0$ to ensure stability. We denote the input process of a general Brownian motion by $\{B^*(t), t \in \mathbb{R}\}$. Then

$$
p(b, T) = \mathbb{P} \left( \sup_{t \geq 0} \{B^*(-t, 0) - ct\} > b_0, \sup_{s \geq 0} \{B^*(T - s, T) - cs\} > b_T \right)
= \mathbb{P} \left( \exists s, t \geq 0 : B^*(-t, 0) > b_0 + ct, B^*(T - s, T) > b_T + cs \right)
= \mathbb{P} \left( \exists s, t \geq 0 : B(-t, 0) > \frac{b_0 + (c - \mu) t}{\sqrt{\lambda}}, B(T - s, T) > \frac{b_T + (c - \mu) s}{\sqrt{\lambda}} \right).
$$

Hence, in order to generalize the results of this section, it follows that we have to set $c \leftarrow (c - \mu)/\sqrt{\lambda}$ and $b_i \leftarrow b_i/\sqrt{\lambda}$, $i = 0, T$ there. In order to generalize the results of Section 2.5.3 on the covariance, in addition we need to multiply the right-hand side of (2.23) and (2.30) by $\sqrt{\lambda}/\sqrt{\lambda} = \lambda$. The results on the correlation coefficient can be generalized in a similar way.

In this section we studied the joint distribution function of the workloads at time 0 and time $T$, the covariance of these workloads, large-buffer asymptotics, and the MPP leading to overflow. It is noted that one may also derive an explicit expression for

$$
q(\bar{b}, T) := \mathbb{P}(Q_T > b_T | Q_0 = b_0),
$$

by using $p(\bar{b}, T)$, see [115] for more details.