On the Large Deviations Approximation for the Stationary Distribution of Skip Free Regulated Queueing Networks

Florin Avram
Department of Statistics and Actuarial Science
Heriot Watt University,Edinburgh, Scotland and
Laboratoire de Mathèmatiques Appliquèes
Université de Pau
F.Avram@ma.hw.ac.uk

Abstract: Large deviations papers like Ignatyouk, Malyshev and Scherbakov [19] have shown that asymptotically, the stationary distribution of homogeneous regulated networks is of the form

$$\mathbb{P}\{X(t) \approx n\} \approx Ke^{-\alpha n},$$

with the coefficient α being different in various "boundary influence domains" and also depending in some of these domains on n.

In this paper we focus on the case of constant exponents α and on a subclass of networks we call "strongly skip-free" (which includes all Jackson and all two dimensional skip free networks). We conjecture that an asymptotic exponent α is constant iff it corresponds to a large deviations escape path which progresses gradually (from the origin to the interior) through boundary facets whose dimension always increases by one. Solving the corresponding large deviations problem for our subclass of networks leads to a family of "local large deviation systems" LLDS (for the constant exponents), which are expressed entirely in terms of the cumulant generating function of the network. In this paper, we show that at least for "strongly skip free" Markovian networks with independent transition processes, the LDDS is closely related to some "local boundary equilibrium systems" LES obtained by retaining from the equilibrium equations only those valid in neighborhoods of the boundary.

Since asymptotic results require typically only that the cumulant generating function is well defined over an appropriate domain, it is natural to conjecture that these LLDS systems will provide the asymptotic constant exponents regardless of any distributional assumptions made on the network.

Finally, we outline a practical recipe for combining the local approximations to produce a global large deviations approximation $\mathbb{P}\{X_t \approx n\} \approx \sum_j K_j e^{-\alpha^{(j)}\, n}$, with the coefficients K_j determined numerically.

Keywords: Large deviations, stationary distribution, skip free networks, phase type, compensation expansion.

1 Introduction

A. Motivation. One of the most influential results of qeueing theory is the fact that Jackson Markovian networks on $\mathbf{Z}_+^I$ have stationary distributions with product form. This case is however quite a special one, in view of the fact that a product form $\prod_{i=1}^I x_i^{n_i}$, determined by I parameters $(x_i, i = 1, ..., I)$ ends up satisfying simultaneously 2^I equilibrium equations (one for each boundary facet). However, it is quite typical that product form solutions may be found which approximate locally the stationary distribution. If these local product solutions are also asymptotically dominant in the subset of the state space where they were generated, they are computable via the large deviations approach. For Markovian and phase-type networks, however, they may also be found directly by simply solving algebraically the subset of the equilibrium equations valid in the part of the state space where they were constructed.

Combining local product approximations appears as an attractive way of producing global approximations for the stationary distribution of networks; under the name of "compensation approximation", complete results for this problem have been obtained by Adan, Wessels and Zijm [1] for the subclass of walks without transitions towards the North East corner. Also, important foundational work for general processes on the two dimensional quadrant has recently been laid by Borovkov, Mogulskii and Rogozin [6], [7], [28], [29].

The first step towards developping such an approximation for general regulated networks on the I'th dimensional octant consists in determining the local "tail exponents" in various subdomains of the state space, and these may be found as solutions of a "large deviations variational problem for escape paths", which has been formulated in Atar and Dupuis [3].

In this paper we focus on the simplest case of **constant tail exponents**, conjectured to correspond to large deviations escape path which progress gradually from the origin to the interior, through boundary facets whose dimension always increases by one. The solution of the Atar Dupuis variational problem for this subclass of escape paths and networks lead to a family of "local large deviation systems" LLDS, which are of polynomial type for the case of Markovian and phase-type networks.

To avoid various foundational issues related to the large deviations approach, in this paper we just formulate the LLDS (as a conjectured recipe for computing the constant exponents of networks with independent transitions), and show that in the case of "strongly skip-free" Markovian networks its solutions include those of the local equilibrium systems LES. We also discuss global approximations, to be obtained by further taking linear combinations of the local approximations with coefficients chosen to optimize various objectives, to be investigated in future work.

B. Framework. We will consider the number $X(t)$ of customers in a "free queueing network" process on $\mathbf{Z}^I$ with "independent transitions". More precisely:

Definition 1. *A "free" queueing network process $X(t)$ with independent transitions is a jump process $X(t) = \sum_{k=1}^{K} d_k N_k(t)$ on $\mathbf{Z}^I$, where $N_k(t)$ denote* **independent** *counting processes with stationary independent increments and $d_k, k = 1, ..., K$ denote fixed jump directions (i.e., transitions which occur at the increase times of the counting processes $N_k(t)$).*

We will denote by T and K the set of all the possible jumps (transitions) and their number and by r_k their average rates (=reciprocals of expected inter transition times).

Queuing network processes, or more precisely **polyhedrally constrained homogeneous processes**, are obtained from free queueing processes by constraining them in addition to remain in some polyhedral state space via some boundary mechanisms which have constant rates (just as the interior mechanism). The boundary mechanisms are typically different on each boundary facet. For example, in the case of networks with unbounded buffers the state space is the I dimensional quadrant $\mathbf{Z}_+^I$ and we might have up to $2^I - 1$ different boundary mechanisms. Additional mechanisms will appear at each boundary of the state space caused by limited buffer capacities or other constraints[1].

In this paper we consider only the **regulation** boundary mechanism (also known as Skorohod mapping), which consists in simply dropping jumps which would take the process outside the state space.

Definition 2. *A "polyhedrally regulated" or simply queueing process $\tilde{X}_t$ is the process obtained by cancelling all the transitions of the free process $X(t)$ which would take it outside a polyhedral state space S.*

C. Examples: Markovian and renewal networks.

Definition 3. *The free queueing network process $X(t) = \sum_{k=1}^{K} d_k N_k(t)$ will be called Markovian if the inter transition times of all the renewal processes $N_k(t)$ have exponential distribution and renewal-type if the inter transition times are independent.*

[1] There is nothing fundamentally different about the state space $\mathbf{Z}_+^I$ to distinguish it from other polyhedral state spaces, beyond the convenience of having only one corner. For a different (maybe bounded) state space S, one should consider a sequence of state spaces of the form $S^{(n)} = nS \bigcap \mathbf{Z}_+^I$ chosen so that after scaling by n they reduce to the fixed polyhedron S. If S has more than one vertex, some of the results below would need to be applied separately at each vertex of the state space, in terms of the local variables defining it.

We assume the existence of the moment generating functions $Ee^{s\,N_k(t)}$ for some strictly positive number s and let

$$C(\alpha) = \lim_{t \to \infty} \frac{\log Ee^{\alpha\,X(t)}}{t}$$

denote the cumulant generating function of the process $X(t)$, for all values α for which it is well defined. **Note:** We adopt the convention that α is a row vector and $X(t)$ a column vector.

By the independence assumption, this reduces to:

$$C(\alpha) = \sum_{k \in T} c_k(s_k)$$

where $c_k(s) = \lim_{t \to \infty} \frac{\log Ee^{s\,N_k(t)}}{t}$ is the cumulant generating function of the k'th counting process and $s_k = \alpha\, d_k$.

In the Markovian case, the cumulant generating functions are given explicitly by $c_k(s) = r_k(e^s - 1)$. Furthemore, explicit equations determining the cumulant generating functions $c_k(s)$ are available for several other types of stationary counting processes. For renewal-type networks for example, the asymptotic cumulant generating function of the renewal counting process $N_k(t)$ is provided by the classical Cramer-Lundberg equation which connects them to the Laplace transform $\hat{F}$ of the interrarrival period.

Proposition: (Cramer-Lundberg) The asymptotic cumulant generating function $c(s) = lim_{t \to \infty} \frac{logEe^{s\,N_t}}{t}$ of a renewal counting process $N(t)$ satisfies the equation

$$\hat{F}(c(s)) = e^{-s} \tag{1}$$

Definition 4. Skip free networks *are networks for which no transition may change any coordinate by more than one unit.*

Thus, for skip free networks the transitions may be either single arrivals of customers, single departures, transfers, splits, merges, joint arrivals, joint departures or joint merge/splits. They include the famous Jackson networks.

Definition 5. Jackson networks *are skip free Markovian networks on* $\mathbf{Z}_+^I$ *for which only single arrivals, single departures and transfers are allowed. The respective rates are traditionally denoted by* $\lambda_i, \mu_i p_{i,0}, \mu_i p_{i,j}$, *where* $p_{i,0} = 1 - \sum_{j=1}^{I} p_{i,j}$. *Thus,* μ_i *represents the sum of all the rates which decrease the* i' *th coordinate,* $P = \{p_{i,j}\}$ *is the matrix giving the proportions of inner routings and* $p_{i,0}$ *give the proportions who leave the system.*

D. The large deviations approximation. Product form stationary distributions for Markovian queueing networks like in the famous Jackson case are a quite rare occurrence. Typically, the systems of difference/differential equations encountered when studying the equilibrium of Markovian/renewal queuing models and their continuous RBM diffusion limits (regulated Brownian motion) do not admit explicit solutions. However, approximate explicit solutions often become available when the system is viewed from far away; this asymptotic approach is known in probablility as the **large deviations approximation**.

The key fact about the large deviations approximation for the "tail" stationary probability $\mathbb{P}(n)$ that a network has very large components $n = (n_1, n_2, ..., n_I)$ is that it is obtained by solving a variational problem which determines the most likely "filling" or "escape" path by which the network filled up from 0 (which is the position of the network in equilibrium when viewed from far away) to n.

To give a flavor of the nature of this variational escape problem in the multidimensional case, we sketch now the results obtained by Ignatyuk, Malyshev and Scherbakov [19] for the escape from equilibrium of two dimensional random walks:

Example: The quadrant case.

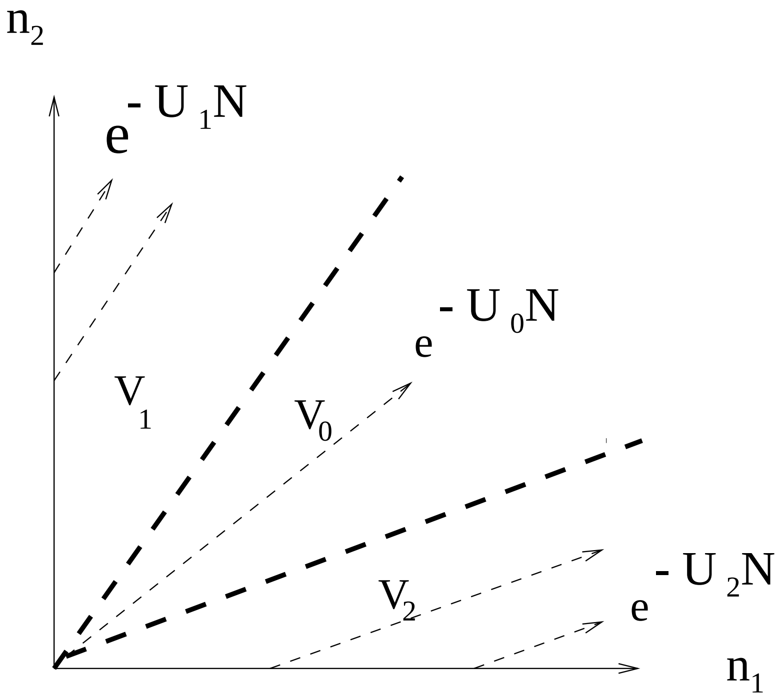

Figure 1: **Two dimensional "Far Away" escape diagram**

We note in Figure 1 that depending on the type of escape path, there may be up to three different regimes. The two boundaries regimes correspond to escape paths which follow in the beginning the respective boundary before turning towards the interior. In between there may be a third "interior" regime corresponding to direct escape paths. (While more complicated piecewise constant escape paths are in principle possible, it turns out here that they can never be asymptotically dominant).

In the two "boundary influence domains" adjoining the axes, the asymptotic behavior of the stationary distribution is of product form $\mathbb{P}(n) \approx Ke^{-\alpha n}$ (where $n = (n_1, n_2, ..., n_I)$ is a column vector and α denotes a row vector of the same dimension), however with different exponents in each boundary influence domain. In the middle domain however, the large deviations approximation is of the form $\mathbb{P}(n) \approx Ke^{-\alpha(n)n}$, with α depending on the angle of n.

In conclusion, the state space above ends up being partitioned in up to three possible subsets, each with different locally optimal escape velocities and exponents, which may all be found explicitly, with the ones adjoining the boundary being even constant.

Recently, Borovkov, Mogulskii and Rogozin [6], [7], [28], [29] have further proved under quite general conditions the convergence of the scaled stationary distributions of random walks on the quadrant to a "global large deviations approximation" which is a linear combination of the three possible locally optimal solutions:

$$\mathbb{P}\{X(t) \approx n\} \approx \sum_j K_j e^{-\alpha^{(j)(n)}\, n}, \tag{2}$$

and provided an explicit (quite complicated) procedure for computing the coefficients K_j.

In this paper we chose to retain from the general large deviations approximation above only the simplest terms with **constant exponents** $\alpha^{(j)}$; they will also be referred to as **primary boundary exponents**. We conjecture that they are related to the simplest possible type of escape paths:

Definition 6. *A gradual escape path is a path from the origin to an interior point* n *with exactly* I *pieces and for which the dimension of the traversed facet increases at each step by* 1.

Conjecture 1. *The exponent* $\alpha(n)$ *dominant at an interior point* n *is constant iff the escape path to* n *is gradual.*

Note: While this conjecture motivates our work, it is not used explicitly below.
We chose to study in this paper the constant boundary exponents for several reasons:

- Quite often the "middle regimes" with escape paths which "jump dimension" are all void, like for example in the case of Jackson networks.
- Even in cases when "middle regimes" do exist, including them might in principle be replaced by a combination of (more tractable) terms with constant exponents, like for example in the "compensation expansion".
- Constant boundary exponents, as illustrated here, may be obtained directly, without using the heavy machinery of large deviations theory.

Finally, we note that approximations of the form (2) have already been constructed in the past, for example in the pioneering work of Boxma and Konheim [8], starting from heuristics arguments which interpreted K_i precisely as probabilities of the network being in the vicinity of some boundaries.

Note: The large deviations approximation seems very well suited for networks with finite buffers and blocking mechanisms, which are very widely used in applications.

E. Local Equilibrium Systems. Local product form approximations may be obtained algebraically simply by solving subsets of the equilibrium equations which are valid locally. For example, consider an edge of the boundary, which is defined as the intersection of $I-1$ hyperpanes; we might look for product form solutions $\prod_{i=1}^{I} x_i^{n_i}$ by trying to satisfy this $I-1$ equations together with the interior equilibrium equation.

However, a more appropriate candidate for local equilibrium equations emerges from the large deviations theory insight that tail probabilities are crucially related to the order in which the components became huge, also known as "saturation" order. Inspired by this insight, we define:

Definition 7. *For a skip free Markovian network and a fixed given ("saturation") order $j_1, j_2, ..., j_I$, let $J_1 = \{j_1\}, J_2 = \{j_1, j_2\}, ...J_I = \{j_1, j_2, ..., j_I\}$ and let E_{J_k} denote the equilibrium equation associated with the facet on which all the coordinates in J_k are positive and the other coordinates are 0.*

The **local equilibrium system LES** *associated to the given saturation order is the system of I polynomial equations with I unknowns obtained by plugging product form solutions $\prod_{i=1}^{I} x_i^{n_i}$ into the system of equations $E_{J_I}, ..., E_{J_2}, E_{J_1}$.*

Notes: 1) Obviously, we should only retain solutions which satisfy the "validity" inequalities $x_i < 1, i = 1, 2, ..., I$.

2) Other additional conditions, discussed further below, are necessary for a solution to be asymptotically optimal. However, we are interested in any valid solutions, even in the case when they are not asymptotically optimal, since they may still be useful for refining the first order large deviations approximation.

3) For networks with phase type intertransition times, the systems $E_{J_1}, E_{J_2}, ..., E_{J_I}$ considered above are still well defined (viewed however over a vector space whose dimension equals the product of the orders of all the phase representations). However, looking for "product form solutions" ceases to be a straightforward procedure in this case. As illustrated in Section 4, for phase type networks a certain amount of case-dependent algebraic exploration may be necessary to process the LES.

Based on the solution of the local large deviation variational problem of Atar and Dupuis [3], provided for Jackson networks by Ignatyouk [20], we formulate below a second type of system for producing the constant boundary exponents, to be called **local large deviations system LLDS**.

The LLDS is formulated in terms of the cumulant generating functions of the transitions processes only, which suggests that as long as these are well defined, it will work independently of specific distributional assumptions; for example, we conjecture that it provides the boundary exponents also for networks whose transition times are modeled by Markovian arrival processes

MAP and for fluid networks. Bypassing the difficulties inherent in a large deviations proof of this conjecture (see for example [15], [12], [13], [14] for some large deviations results for networks and [30] and [25], [27] for some foundational difficulties for establishing large deviation results), we only give a simple algebraic justification which works in the Markovian case and also for particular examples of networks with phase-type transitions (the particular case of tandem networks is already known from the work of Fujimoto, Makimoto and Takahashi [17]).

F. Contents. The paper is structured as follows. In Section 2 we introduce the LLDS and show that its solutions include those of the LES in the "strongly skip free" Markovian case. In sections 3 and 4 we illustrate the LLDS by computing the boundary exponents for some concrete two dimensional Markovian and renewal type networks, respectively. Section 5 discusses the implementation of a global large deviations approximation obtained by combining the various local approximations.

2 The local large deviation systems LLDS

Works like for example Ignatyouk, Malyshev and Scherbakov [19], Knessl and Tier [21], [23], Borovkov and Mogulskii [6], Atar and Dupuis [3], Avram, Dai and Hasenbein [4] and Ignatyouk [20] suggested that the variational problem for the escape of networks from equilibrium may be reduced to a sequence of calculus problems and solved explicitly in terms of the cumulant generating functions of the transition processes involved. Indeed, solving the the large deviations variational problem (for the most likely escape velocities and their associated tail exponents) in the case of gradual escape paths lead us to the **local large deviations systems LLDS** formulated below. These systems have been also identified in the particular case of tandem networks by Fujimoto, Makimoto and Takahashi [17].

The LLDS is defined for skip free networks, via a recursive recipe we call **shorting** (which removes by averaging one "fastly changing, non saturated component").

Definition 8. *Let n_j denote the last coordinate to become positive on the escape path to the interior in a skip free network. Let $E(x) = C(\alpha) = 0$ denote the free cumulant generating function of the network, expressed in terms of variables x_i defined by $x_i = e^{-\alpha_i}$. In the Markovian case, this substitution puts the cumulant generating function in the rational form:*

$$E(x) = \sum_k r_k(w_k(x) - 1),$$

where $w_k(x) = e^{\alpha\, d_k} = \prod_{i=1}^{I} x_i^{-d_{k,i}}$. We will call this rationalized expression E the "fundamental form" of the network. We break now the fundamental form into $E = R' + E'$ where:

$$R' = \sum_k^{(j)} r_k(w_k - 1),$$

denotes the part of the fundamental form obtained by summing only over the transitions cancelled on $n_j = 0$, indicated by $\sum^{(j)}$. We call this the **reducing equation***.*

Finally, let $\tilde{E}$ be the expression obtained by eliminating the variable x_j (corresponding to the last component to become saturated) from the rest of the fundamental form $E' = E - R'$ by using the equation $R' = 0$ (Note that by the skip-free assumption, the reducing equation R' is linear in x_j and has therefore a unique solution). We call $\tilde{E}$ the **shorted equation***.*

Note: Formally, we replace the variable x_j in both equations by a new variable x'_j to be called "auxiliary twist", and then eliminate it.

Lemma 2.1. *If E is the fundamental form of a skip free Markovian network, then for any variable x_j the shorted equation $\tilde{E}$ is also the fundamental form of a network, which is obtained as follows: 1) The component j is removed and all transitions not involving it are kept unchanged. 2) The transitions involving the j'th component are partitioned into the sets $\mathcal{A}_j$ and $\mathcal{D}_j$ of transitions which increase and respectively decrease by one the component j. These transitions are then replaced in the new network by the set of all sums $d = d_i + d_{i'}$, where $d_i \in \mathcal{A}_j$ and $d_{i'} \in \mathcal{D}_j$, with new rates c_j determined from (3).*

Proof: In the Markovian skip-free case, each cumulant generating function is of the form $c_k(x) = r_k w_k(x) - r_k$, where $w_k(x) = \prod x_i^{-d_{k,i}}$ is the product of all the components x_i increased by the transition k with exponent -1 and of all the components x_i decreased by the transition k with exponent 1. Therefore, we may split the fundamental form as

$$E = (x_j D_1 - d_1) + (x_j^{-1} A_1 - a_1) + U,$$

with the three pieces corresponding respectively to transitions which decrease/increase/ leave unaffected the component n_j. More precisely, the terms $x_j D_1, x_j^{-1} A_1$ correspond to transitions into and d_1, a_1 correspond to transitions out of a given state $n = (n_1, ..., n_j, ...)$, d_1 and a_1 are constants and D_1, A_1 are functions of the remaining exponents $x_i, i \neq j$ which equal d_1, a_1 when all x_i equal 1.

The two equations involved in the shorting step of the fundamental form R', E' are therefore

$$R' = x'_j D_1 - d_1 = 0, \qquad E' = \frac{1}{x'_j} A_1 - a_1 + U = 0,$$

and eliminating x'_j yields

$$\tilde{E} = U + \frac{D_1 A_1 - a_1 d_1}{d_1}.$$

Both D_1 and A_1 are sums of monomials involving only the powers ± 1 and the strongly skip free assumption implies that their product may also be written in the form $\sum c_j w_j(x)$, with $w_j(x)$ being monomials with exponents as described in the Lemma. This leads to a decomposition

$$\frac{D_1 A_1 - a_1 d_1}{d_1} = \sum c_j (w_j(x) - 1) \tag{3}$$

for some coefficients c_j which may be interpreted as the rates of a new Markovian network ■

Let now $o = (j_1, j_2, \ldots, j_I)$ denote a fixed order in which the coordinates become saturated for a given type of gradual escape path. The coordinate j_I is thus the last to become huge according to the order o.

In order to apply the shorting procedure recursively for $i = 1, 2, ...I - 1$, we need to assume that at each step the shorted network is still skip free. We introduce therefore:

Definition 9. Strongly Skip Free networks. *A skip free network is called "strongly skip free" if either*

- $I = 2$ *(it is two dimensional) or*
- *for any order of shorting* $o = (j_1, j_2, \ldots, j_{I-1})$, *the consecutive shorted networks are all skip free.*

Strongly skip free networks include many networks, besides the classical Jackson ones; for example, introducing only joint arrivals or only joint departures is alright, as long as the network has no transfer transitions.

We introduce now the LLDS via a recursive recipe which builds $2I - 1$ equations with as many unknowns. It starts with the last segment which contributes only one equation, the free "fundamental form" $E_0 = C(\alpha) = 0$, and then adds two equations for each consecutive escape segment on the way back to the origin.

Step 1: We add the equations $R'_1 = \sum_k^{(j_I)} r_k(w_k - 1)$, and $E'_1 = E_0 - R'_1$ obtained by shorting the component n_{j_I}, (with x_{j_I} replaced by x'_{j_I}), and the equation $\tilde{E}_1$ obtained by eliminating the variable x'_{j_I} from E'_1 by using the equation $R'_1 = 0$.

Step 2. Next, we further break the equation $\tilde{E}_1$ in two equations to be denoted as $R'_2 = 0, E'_2 = 0$. R'_2 is given by $R'_2 = \sum_k^{(j_I, j_{I-1})} r_k(w_k - 1)$, where the summation $\sum^{(j_I, j_{I-1})}$ denotes all the terms in the equation $\tilde{E}_1$ corresponding to new transitions which become impossible when $n_{j_{I-1}} = 0$, (and have not already been included in R'_1); E'_2 is the remaining part $E'_2 = \tilde{E}_1 - R'_2$.

Also, $x_{j_{I-1}}$ is replaced by a new "auxiliary twist" $x'_{j_{I-1}}$ in both equations, and eliminated to obtain $\tilde{E}_2$.

Proceeding in this fashion, we construct pairs of equations $R'_j = 0, E'_j = 0, j = 1, ..., I-1$ which correspond to removing consecutively the coordinates $j_I, j_{I-1}..., j_2$ from the end of the escape path. The last equation $\tilde{E}_{I-1}$ involves only the exponent x_{j_1} corresponding to the first component which became positive.

Definition 10. *a) The extended LLDS for a skip free Markovian network associated to a sequence $o = (j_1, j_2, \dots j_I)$ (which describes the order in which the coordinates become positive along an escape path) is the system of $2I-1$ equations $E_0 = 0, R'_i = 0, E'_i = 0, i = 1, ..., I-1$ having as unknowns the I original variables x_i and the $I-1$ "auxiliary twists" $x'_{j_I}, x'_{j_{I-1}}, x'_{j_2}$.*

b) The shorted LLDS is the system of I equations $E_0 = 0, \tilde{E}_i, i = 1, ..., I-1$ obtained after the elimination of the auxiliary twists.

Theorem 1. *For strongly skip free Markovian networks, the solutions of the LES are included among those of the shorted LDDS.*

Proof: We put first the LES system of equations $E_{J_I}, ..., E_{J_2}, E_{J_1}$, in a more convenient (equivalent) form.

With $E_0 = E_{J_I}$ denoting again the "interior" equilibrium equations, let R_1 denote the difference of the equations E_{J_I} and $E_{J_{I-1}}$. Clearly, the system $R_1, E_{J_{I-1}}, ..., E_{J_2}, E_{J_1}$, is equivalent to the LES. Eliminating now x_{J_I} from $R_1 = E_0 - E_{JI-1} = x_{J_I}^{-1} A_1 - d_1 = 0$ we find that $E_{J_{I-1}}$ transforms precisely into the second LLDS equation

$$\tilde{E}_1 = U + \frac{D_1 A_1}{d_1} - a_1$$

Thus, the LES is equivalent to $R_1 = 0, \tilde{E}_1 = 0, ..., E_{J_2} = 0, E_{J_1} = 0$ and proceeding to replace the equations we find that it is equivalent to

$$\textbf{LES} \qquad R_k = 0, k = 1, ..., I-1, \tilde{E}_{I-1} = 0$$

where R_k, D_k, A_k, d_k, a_k are defined inductively at the k'th shorting step.

On the other hand, taking the difference of the first two LLDS equations yields:

$$-(x_{J_I} D_1 - d_1)(A_1 x_{J_I}^{-1} - d_1)/d_1 = -R'_1 R_1/d_1$$

and proceeding inductively we find that the LLDS equations are equivalent

$$\textbf{LDDS} \qquad R'_k R_k = 0, k = 1, ..., I-1, \tilde{E}_{I-1} = 0$$

and so its solutions include those of the LES ■

The LLDS provides the solution for strongly skip free networks of the large deviations escape variational problem formulated by Atar and Dupuis [3]. While a complete proof of this is not available yet, we provide now some notes on this relation, for a better appreciation of the particular cases presented in the next section.

The variational escape problem requires finding for each facet Λ traversed by the escape path the rates of change y_k of the renewal coordinates $N_k(t)$ and the probabilities p_k that all the components necessary for firing the transition k are positive, where y_k, p_k mimimize an "averaged" rate functional

$$\bar{L}(Y) = \sum_k p_k l_k(y_k)$$

The numbers p_k may be interpreted as proportions of time that the transition k is used. They must satisfy the constraint that after averaging with respect to p_k, the network has a mean drift directed along Λ.

For a given escape path, all transitions may uniquely be assigned to the first coordinate whose removal (shorting) makes that transition impossible and all transitions k made impossible by shorting the same coordinate x_j will have the same probability p_j of being active, satisfying

$$\sum_{k \in \mathcal{A}_j} y_k - p_j \sum_{k \in \mathcal{D}_j} y_k = 0$$

where $\mathcal{A}_j, \mathcal{D}_j$ denote the set of transitions increasing/decreasing j. We find

$$p_j = \frac{\sum_{k \in \mathcal{A}_j} y_k}{\sum_{k \in \mathcal{D}_j} y_k} \tag{4}$$

which may also be put in the form

$$p_j = \frac{x_j{}'}{x_j} = e^{\alpha_j' - \alpha_j}$$

obtained using the equations $R_j = 0, R_j' = 0$ satisfied by x_j, x_j'.

Notes: 1) It turns out when the condition $p_j < 1$ is satisfied at each shorting step, the coresponding product approximation is indeed asymptotically optimal.

2) Each pair of equations in the extended LDDS describes the escape path during a corresponding escape segment, with the expression $y_k = r_k w_k(x)$ yielding the "twisted" escape rate at which the transition k works while on that segment.

3) If during some shorting step only one transition k is regulated, the reducing equation $c_k(s_k) = y_k - r_k = 0$ implies that the twisted escape rate y_k equals the unconstrained expected rate r_k. (If more than one transition is regulated, the reducing equation may be interpreted as a conservation law $\sum y_k = \sum r_k$)

4) Transitions which do not affect the component being shorted will not modify their rate y_k when leaving the corresponding boundary.

5) The expressions $\tilde{E}_k$ in the shorted LLDS depend only on the first $I - k$ exponents $j_1, j_2, .., j_{I-k}$. They have an important probabilistic interpretation, turning out to be precisely the cumulant generating functions of "shorted free networks" in which the last coordinates $j_{I-k+1}, ...j_I$ have been averaged succesively, and the first $I - k$ coordinates $j_1, j_2, .., j_{I-k}$ are "free/saturated", i.e. so large that they evolve without any effect from the nonnegativity constraints.

In conclusion, while local product approximations may be computed in the Markovian (and phase-type) case via the standard LES, the extended LLDS which represents the solution of the large deviations variational problem has the advantage of being well defined outside the Markovian framework and also of providing additional information on the asymptotic optimality of the approximants.

3 Two dimensional examples of boundary exponents for skip free Markovian networks

In this section we illustrate the computation of boundary exponents for skip free regulated Markovian processes $X(t)$ whose state space S is the nonegative quadrant $Z_+^2 = \{(n_1, n_2) : n_1, n_2 \geq 0\}$. $X(t)$ is here a jump process $X(t) = \sum_{k=1}^K d_k N_k(t)$ on $\mathbf{Z}_+^2$ with $d_k, k = 1, ..., K$ denoting fixed jump directions which occur at the arrival times of independent Poisson processes $N_k(t)$ with rates r_k. We also assume for simplicity (w.l.o.g.) that the sum of the rates satisfies $\sum_{k=1}^K r_k = 1$, under which the model may also be interpreted as a discrete random walk whose summands $\xi(i)$ take values d_k with probability r_k. We recall that regulation at the boundaries means that the transitions d_k which would take the process outside the state space are replaced by staying on the spot ("canceled transitions"). The "fundamental form" of the process is the expression $E_0(x) = \sum_k r_k(w_k - 1)$, where $w_k = w_k(x) = \prod_{i=1}^I x_i^{-d_{k,i}}$. For proceses on the quadrant there are only two possible types of gradual escape paths: $\{1\} \Longrightarrow \{1,2\}$ and $\{2\} \Longrightarrow \{1,2\}$ (each facet is denoted by the set of its nonnegative coordinates).

We provide now a couple of examples of using the LDDS.

3.1 M/M/M tandem

This is the simplest Jackson network, and therefore the local product approximations are all equal, being given by the famous Jackson equations. Still, the LLDS solved below provides the escape rates, which don't seem to be available in the literature.

Let us denote by $a = (1,0), t = (-1,1), d = (0,-1)$ the vectors of change after a first station arrival, transfer and second station departure, respectively. The set of possible transitions for the M/M/M tandem is $T = \{a,t,d\}$ and their rates will be denoted by λ, μ_1, μ_2 respectively. The cumulant generating function of the free queueing process and the fundamental form are thus:

$$\lambda(e^{\alpha_1}-1)+\mu_2(e^{-\alpha_2}-1)+\mu_1(e^{-\alpha_1+\alpha_2}-1) = \lambda({x_1}^{-1}-1)+\mu_2(x_2-1)+\mu_1(x_1x_2^{-1}-1),$$

where we put $x_i = e^{-\alpha_i}, i = 1,2$.

The (common) equation E_0 for the inner segment, the reducing equation R_1, the auxiliary twist and the reduced equation $\tilde{E}_1$ are given in the table below for the two possible types of gradual escape paths $\{1\} \Longrightarrow \{1,2\}$ and $\{2\} \Longrightarrow \{1,2\}$.

$\boxed{E_0:}$	$\lambda({x_1}^{-1}-1)+\mu_2(x_2-1)+\mu_1(x_1x_2^{-1}-1)=0$	
	along $n_2=0: \{1\} \Longrightarrow \{1,2\}$	along $n_1=0: \{2\} \Longrightarrow \{1,2\}$
$\boxed{R_1':}$	$c_d = \mu_2(x_2'-1)=0,\ x_2'=1$	$c_t = \mu_1(x_1'x_2^{-1}-1)=0,\ x_1'=x_2$
$\boxed{E_1':}$	$c_a+c_t = \lambda({x_1}^{-1}-1)+\mu_1(x_1(x_2')^{-1}-1)=0$	$c_a+c_d = \lambda({x_1'}^{-1}-1)+\mu_2(x_2-1)=0$
$\boxed{\tilde{E}_1:}$	$\lambda({x_1}^{-1}-1)+\mu_1(x_1-1)=0$	$\lambda({x_2}^{-1}-1)+\mu_2(x_2-1)=0$

We note that the reduced equations $\tilde{E}_1$ are the fundamental forms of a $M/M/1$ queue with parameters λ, μ_1 and λ, μ_2 respectively, with clear probabilistic interpretations. The final solution in both cases coincide, being given by the classical Jackson solution $x_1 = \frac{\lambda}{\mu_1}, x_2 = \frac{\lambda}{\mu_2}$.

We can also compute the most likely escape rates; starting from the origin and moving towards the interior along the two types of broken paths, we have:

1. Along $n_2 = 0$
$$(y_a, y_t, y_d) = (\boxed{\mu_1}, \lambda, \tilde{\mu_2}) \text{ followed by} (\boxed{\mu_1}, \mu_2, \lambda)$$
2. Along $n_1 = 0$
$$(y_a, y_t, y_d) = (\mu_2, \tilde{\mu_1}, \boxed{\lambda}) \text{ followed by} (\mu_1, \mu_2, \boxed{\lambda}).$$

We note in both cases that the regulated transition on the first leg of the escape indicated by a tilde works at mean rate and that the "unaffected" transition indicated by a box stays unchanged. Also, the nonregulated rates are simply swapped. Finally, the interior escape rates for both cases are the same, which is equivalent with the exponents being equal.

Optimality check: The solutions obtained above are indeed asymptotically dominant only if the proportions of time p_j spent on the boundary $n_j = 0$ are strictly positive; p_j may be found from the equation (4) which ensures that the mean drift is 0 along the boundary (or from the unproved formula $p_j = \frac{x_j}{x_j'}$) yielding respectively:

$$\begin{aligned} p_2 &= \frac{y_t}{y_d} = \frac{\lambda}{\mu_2} \\ p_1 &= \frac{y_a}{y_t} = \frac{\mu_2}{\mu_1}. \end{aligned}$$

Since $p_2 < 1$ the first type of gradual escapes is always feasible, but the second is feasible only if $\mu_1 > \mu_2$. If $\mu_1 \leq \mu_2$ only escapes through the boundary $n_2 = 0$ are feasible (see Figure 2).

The asymptotical optimality condition $p_1 < 1$ was also found in the form $\frac{x_1}{x_2} < 1$ by Fujimoto, Makimoto and Takahashi [17] for the renewal tandem. From the point of view of the algebraic "equilibrium approach" the necessity of this conditions is rather puzzling (see [17], Remark 2 following Theorem 3.2), and its meaning is only revealed via the large deviations approach which interprets the quantity above as a probability which needs to be strictly less than 1.

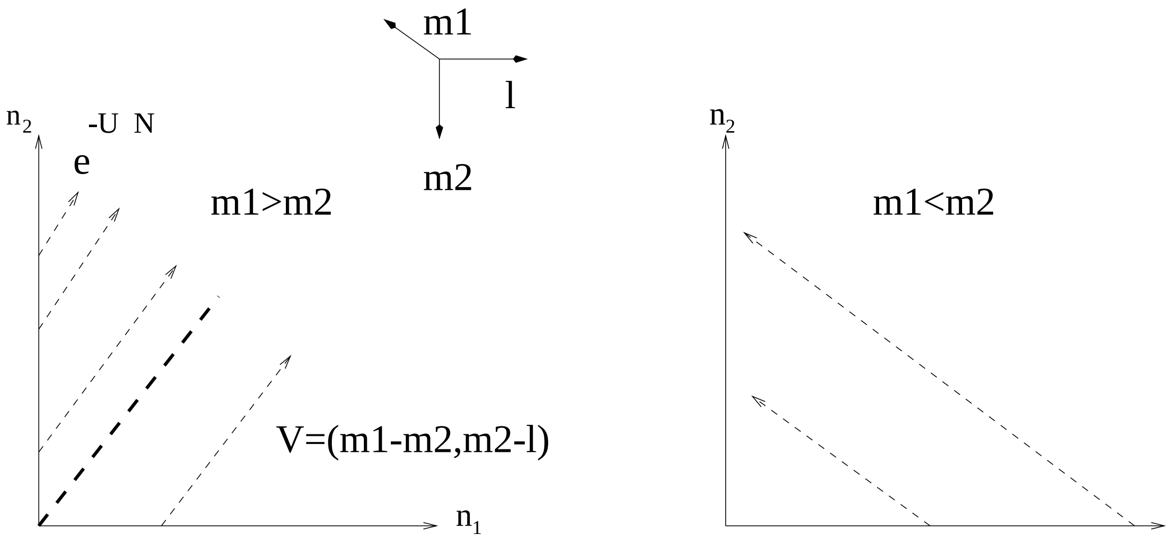

Figure 2: **Escape directions for Jackson tandem**

3.2 Joint departures Markovian network

We consider now a non Jackson Markovian network obtained from two parallel servers by adding the possibility of joint departures.

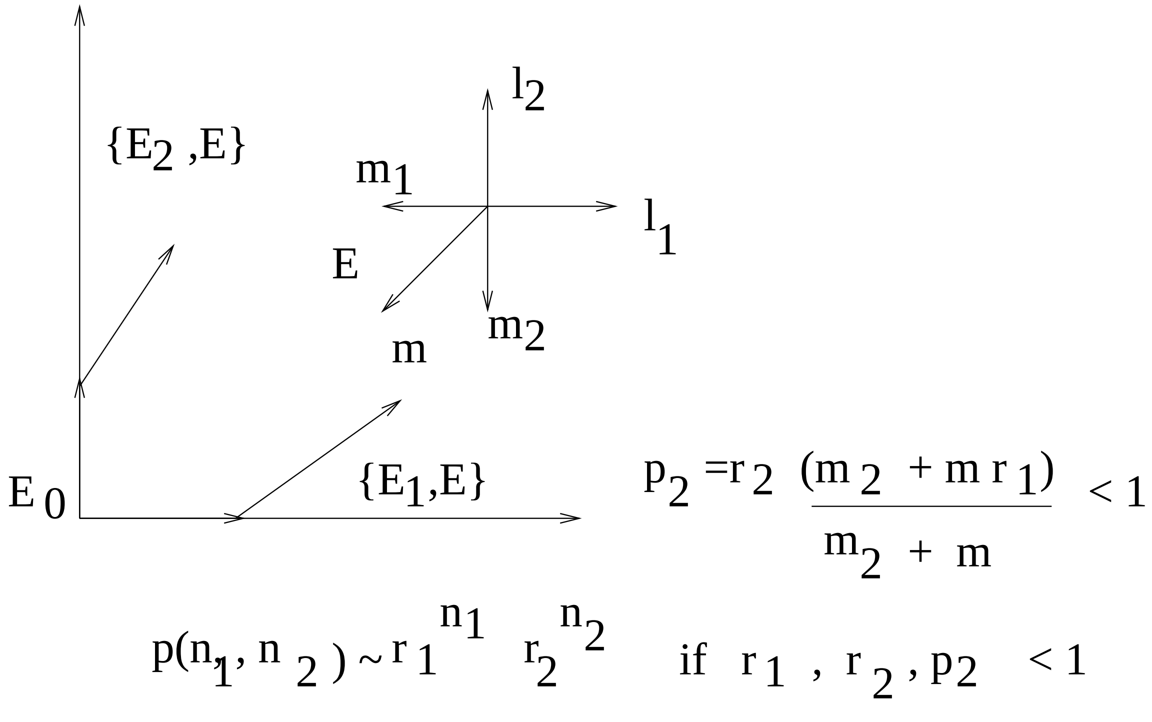

Figure 3: **Joint departures network**

Let us denote by $a_1 = (1,0), a_2 = (0,1), d_1 = (-1,0), d_2 = (0,-1), d = (-1,-1)$ the vectors of change after arrivals, departures and joint departure, with respective rates denoted by $\lambda_1, \lambda_2, \mu_1, \mu_2$ and μ.

The cumulant generating function is:

$$c_{a_1} + c_{a_2} + c_{d_1} + c_{d_2} + c_d = \lambda_1(e^{\alpha_1} - 1) + \lambda_2(e^{\alpha_2} - 1) + \mu_1(e^{-\alpha_1} - 1) + \mu_2(e^{-\alpha_2} - 1)$$
$$+\mu(e^{-\alpha_1 - \alpha_2} - 1) = \lambda_1({x_1}^{-1} - 1) + \lambda_2({x_2}^{-1} - 1) + \mu_1(x_1 - 1) + \mu_2(x_2 - 1) + \mu(x_1 x_2 - 1)$$

The (common) equation E_0 for the inner segment, the reducing equation R_1, the auxiliary twist and the reduced equation $\tilde{E}_1$ are given in the table below for the gradual escape paths $\{1\} \Longrightarrow \{1,2\}$ near $n_2 = 0$.

E_0 :	$\lambda_1({x_1}^{-1}-1)+\lambda_2({x_2}^{-1}-1)+\mu_1(x_1-1)+\mu_2(x_2-1)+\mu(x_1x_2-1)$
	along $n_2 = 0 : \{1\} \Longrightarrow \{1,2\}$
R_1' :	$c_{d_2}+c_d=\mu_2(x_2'-1)+\mu(x_1x_2'-1)=0,\ x_2'=\frac{\mu_2+\mu}{\mu_2+\mu x_1}$
E_1' :	$c_{a_1}+c_{a_2}+c_{d_1}=\lambda_1({x_1}^{-1}-1)+\lambda_2({x_2'}^{-1}-1)+\mu_1(x_1-1)=0$
$\tilde{E}_1$:	$\lambda({x_1}^{-1}-1)+\left(\mu_1+\mu\frac{\lambda_2}{\mu_2+\mu}\right)(x_1-1)=0$

We note that the reduced equation $\tilde{E}_1$ is the fundamental form of the $M/M/1$ queue with the original rates λ_1, μ_1 at the first station with an added second departure stream of rate μ working only $\frac{\lambda_2}{\mu_2+\mu}$ of the time, with a clear probabilistic interpretation. The final solution is $x_1 = \frac{\lambda_1}{\mu_1+\mu\frac{\lambda_2}{\mu_2+\mu}}, x_2 = \frac{\lambda_2}{\mu_2+\mu}$.

Optimality check: For the approximation near the n_2 axis for example, the extra feasibility condition obtained from (4) turns out to be $p_2 = \frac{y_{a_2}}{y_d+y_{d_2}} = x_2\frac{\mu_2+\mu x_1}{\mu_2+\mu} < 1$.

4 The LLDS for renewal networks

The results for renewal networks to be described in this section, which increase considerably the power of the LD approach, are obtained merely by replacing the explicit formulas available for $c_k(s)$ in the Markovian case by the solution of the implicit Cramer-Lundberg equation they must satisfy in the renewal case. We note first an immediate corollary of the Cramer-Lundberg equation:

Lemma 4.1. *The asymptotic cumulant generating function* $c(\alpha) = \lim_{t\to\infty} \frac{\log\left(Ee^{\alpha\, X(t)}\right)}{t}$ *of a free renewal queueing process* $X(t)$ *is given by*

$$c(\alpha) = \sum_k c_k(\alpha'\, d_k) \tag{5}$$

where $c_k(s)$ *are the unique solutions of*

$$\hat{F}_k(c_k(s)) = e^{-s} \tag{6}$$

Note: If the Laplace transform $\hat{F}_k$ of the inter transition time is of phase type, the corresponding Cramer-Lundberg equation is algebraic in c_k. As a corollary, the local product form approximations may be found by solving algebraic systems in this case.

We conjecture that the formally guessed analog of the local large deviations systems LLDS from the Markovian case continues to work also in the renewal case; in the next two subsections we illustrate first its use in the tandem example, provide then the algebraic derivation of the LLDS from the local equilibrium equations LES in this case and finally provide a converse to a result of Knessl [22] concerning the existence of an exact compensation expansion with precisely three terms. The existence of exact compensation expansions illustrates the point that the solutions of the LLDS maybe useful even without being asymptotically optimal.

4.1 LLDS for Ph/M/M tandem

We consider now a Ph/Ph/Ph renewal tandem with cumulant generating functions $c_k(s_k)$ where k can stand for a(rival), t(ransfer) or d(eparture) and respective rates λ, μ_1, μ_2. This problem was extensively studied from a large deviations point of view in [31], [18] and [11] and from the matrix-geometric point of view in [17] and [5], and we bring it here for the purpose of comparing the two approaches.

For the convenience of explicit computations, we will assume as in [18] exponential transfer and service times, so that $c_t = \mu(x_1 x_2^{-1} - 1), c_d = \mu_2(x_2 - 1)$, and furthermore focus on the case of Erlang arrivals ER_n, in which $\hat{F}_a(c_a) = \hat{F}(c_a) = \left(\frac{\lambda n}{\lambda n + c_a}\right)^n, \hat{F}^{-1}(w) = \lambda n(w^{-1/n} - 1)$; Knessl [22] studied deterministic arrivals which maybe obtained as the limiting case $n \to \infty$, yielding $\hat{F}^{-1}(w) = -\lambda \log w$.

The Cramer Lundberg fundamental system E_0 for the inner escape path is:

$$
\begin{aligned}
c_a + c_t + c_d &= 0 \\
c_a &= \hat{F}^{-1}(x_1) = \lambda n(x_1^{-1/n} - 1) \approx -\lambda \log x_1 \\
c_t &= \mu_1(x_1 x_2^{-1} - 1) \\
c_d &= \mu_2(x_2 - 1)
\end{aligned}
$$

Further equations R_1, E_1 for each type of gradual escape are obtained by **shorting**:

$$
\begin{array}{lcc}
 & \text{for } n_2 = 0 & \text{for } n_1 = 0 \\
\boxed{R_1' :} & \mu_2(x_2' - 1) = 0 & \mu(x_1' x_2^{-1} - 1) = 0 \\
\boxed{E_1' :} & c_a' + c_t' = 0 & c_a' + c_d' = 0 \\
 & \hat{F}_a(c_a') = x_1 & \hat{F}_a(c_a') = x_1' \\
 & c_t' = \mu(x_1(x_2')^{-1} - 1) & c_d' = \mu_2(x_2 - 1)
\end{array}
$$

The two systems become $\hat{F}_a(\mu_1(1 - x_1)) = x_1,\ x_2 = x_1 \frac{\mu_1}{\mu_2}$ and $\hat{F}_a(\mu_2 - \mu_2 x_2) = x_2,\ \hat{F}_a(\mu_2 - \mu_2 x_2 + \mu_1 - \mu_1 \frac{x_1}{x_2}) = x_1$ respectively, as already shown in [18].

4.2 Algebraic approach for the Ph/M/M tandem

In this section we present a compensation approximation for the Ph/M/M tandem. The primary terms were first obtained for general Ph/Ph/Ph tandems in [17] (see also [24]).

This example explains our interest in studying the LLDS formulated in Section 2, as opposed to the LES. On one hand, it will illustrate the point that while the primary product approximations may be obtained by the LES, this requires considerable "specialized" effort, with methods which are dependent on the example. On the other hand, the existence of exact compensation expansions obtained by Knessl[22] in the D/M/M case (with three terms only) illustrates the usefulness of the LLDS even in cases when the feasibility conditions (not included in the LLDS) are not satisfied.

We denote by μ_1, μ_2 the rates of service in the two stations and by α, T be an irreducible representation of the phase type arrival process, with T denoting the transition matrix and α being a **row** vector which sums to 1 (the initial distribution of the phases). Let $\mathbf{o}$ denote a column vector of ones, $\mathbf{t}$ be the **column** vector of arrival rates $\mathbf{t} = -T\mathbf{o}$, let $\hat{F}(c) = \alpha(cI - T)^{-1}\mathbf{t}$ be the Laplace transform of the arrival density and let $\hat{F}^{-1}$ denote its inverse function (note that $\hat{F}(c)$ is a decreasing function).

The stationary probability is a vector function $\pi = \pi(n_1, n_2)$ with values in the vector space $\mathbb{R}^d$, where d is the dimension of the phase representation. Let T_a, T_t, T_d be operators of translation in the directions: $(-1, 0), (1, -1)$ and $(0, 1)$ respectively. The equilibrium equations are of the form $A\pi = 0$, with the operator A defined as follows:

- In the interior

$$A = T - (\mu_1 + \mu_2)I + \mathbf{t}\alpha\, T_a + \mu_1 T_t + \mu_2 T_d.$$

- Near the n_1 axis

$$A_1 = T - \mu_1 I + \mathbf{t}\alpha T_a + \mu_2 T_d.$$

- Near the n_2 axis

$$A_2 = T - \mu_2 I + \mu_1 T_t + \mu_2 T_d.$$

We will need repeatedly the following lemma of [17]:

Lemma 4.2. *(FMT) For any $w > 0$ and c in the domain of convergence of $\hat{F}$, the "modified eigenvector" problem*

$$\mathbf{v}(T + w^{-1}\mathbf{t}\alpha) = c\mathbf{v}$$

has a positive (row vector) solution iff $\hat{F}(c) = w$, which is given by $\mathbf{v} = \alpha(cI - T)^{-1}$, up to a multiplicative constant.

4.2.1 Primary terms

We consider first the LES for the order **(1,2)**. Looking for product form solutions $\pi(n_1, n_2) = x_1^{n_1} x_2^{n_2}\mathbf{v}$ satisfying $A\pi = 0, A_1\pi = 0$ leads to the following two equations:

$$\begin{aligned}\mathbf{v}(T + x_1^{-1}\mathbf{t}\alpha) &= (-c_t - c_d)\mathbf{v}\\ \mathbf{v}(T + x_1^{-1}\mathbf{t}\alpha) &= (\mu_1 - \mu_2 x_2)\mathbf{v}\end{aligned}$$

where we put $c_t = \mu_1(\frac{x_1}{x_2} - 1), c_d = \mu_2(x_2 - 1)$ (this notation becomes convenient in the large deviations approach later).

From Lemma 4.2, we conclude that

$$\hat{F}(\mu_1 - \mu_2 x_2) = x_1 = \hat{F}(\mu_1(1 - \frac{x_1}{x_2}) + \mu_2(1 - x_2)))$$

This LES is equivalent therefore with the equations $\mu_1 x_1 = \mu_2 x_2, \hat{F}(\mu_1 - \mu_1 x_1) = x_1$ (which were also derived in Section 4.1 via the LLDS) and $\mathbf{v} = \alpha\big((\mu_1 - \mu_1 x_1)I - T\big)^{-1}$.

We consider now the LES for the order **(2,1)**.This time we can only find a product form solutions $\pi_{n_1,n_2} = y_1^{n_1} y_2^{n_2}\mathbf{v}$ valid for $n_1 \geq 1,)$ with a different form $y_2^{n_2}\mathbf{v_0}$ valid for $n_1 = 0$. This leads to the following three equations:

$$\begin{aligned}\mathbf{v}(T + y_1^{-1}\mathbf{t}\alpha) &= (-c_t - c_d)\mathbf{v}\\ \mathbf{v_0}\mathbf{t} &= \mathbf{v}\mathbf{t}\\ \mathbf{v_0}\big(-c_d I - T\big) &= \mu_1\frac{y_1}{y_2}\mathbf{v}\end{aligned}$$

The first equation yields $\mathbf{v} = \alpha\big((-c_t - c_d)I - T\big)^{-1}$, and $\hat{F}(-c_t - c_d) = y_1$. The last equation yields

$$\begin{aligned}\mathbf{v_0} = \mu_1\frac{y_1}{y_2}\mathbf{v}\big(-c_d I - T\big)^{-1} &= \mu_1\frac{y_1}{y_2}\alpha\big((-c_t - c_d)I - T\big)^{-1}\big(-c_d I - T\big)^{-1}\\ &= \mu_1\frac{y_1}{y_2}\frac{1}{-c_t}\alpha\left(\big((-c_t - c_d)I - T\big)^{-1} - \big(-c_d I - T\big)^{-1}\right)\end{aligned}$$

Multiplied by $\mathbf{t}$ and using the second equation and the form of $\mathbf{v}$ this yields $\hat{F}(-c_t - c_d) = \frac{\frac{y_1}{y_2}}{1-\frac{y_1}{y_2}}\big(\hat{F}(-c_d) - \hat{F}(-c_t - c_d)\big) = \frac{1}{\frac{y_2}{y_1}-1}\big(\hat{F}(-c_d) - \hat{F}(-c_t - c_d)\big)$. Simplifying yields

$\hat{F}(-c_d) = \frac{y_2}{y_1}\hat{F}(-c_t - c_d) = y_2$.

We conclude that this LES is equivalent with the equations

$$\hat{F}(\mu_2 - \mu_2 y_2) = y_2$$
$$\hat{F}(\mu_1(1 - \frac{y_1}{y_2}) + \mu_2(1 - y_2))) = y_1$$

(which were also derived in Section 4.1 via the LLDS) and

$$\mathbf{v} = \alpha\big((-c_t - c_d)I - T\big)^{-1}$$
$$\mathbf{v_0} = \mu_1 \frac{y_1}{y_2}\mathbf{v}\big(- c_d I - T\big)^{-1}.$$

4.2.2 Compensation terms

Let K_1P_1, L_1Q_1 denote the primary product approximants constructed in the previous section, with two as yet undetermined constant prefactors. For each of them we may further develop a **compensation expansion** of the form:

$$K_1P_1 + K_2P_2 + K_3P_3 + ... + Q_1 + L_2Q_2 + L_3Q_3 + ...$$

where P_i, Q_i denote product approximants and K_i, L_i are constants, and each term "compensates" its predecessor.

Definition 11. *The compensating term for a term $p(N)$ on a boundary Λ is a term $q(N)$ such that*
1)$q(N)$ satisfies the "inner" Markovian equilibrium operator and
2) $p(N) + q(N)$ satisfy together the Λ boundary Markovian operator.

Both the product approximants P_i, Q_i and the preceding prefactors K_i, L_i are easily computable as functions of the preceding terms (see for example ([1])). Hence, the only remaining unknowns in a compensation approximation are the constants K_1, L_1 preceding the primary terms, for the determination of which one needs in general to appeal to numerical methods [2]

An interesting particular case with an explicit solution is the D/M/M tandem with exponential servers and deterministic arrivals solved in [22]. In this case, the solution is given by the sum of three products $p = K_1P_1 + L_1Q_1 + R$ where P_1, Q_1 denote the primary approximations constructed near the n_1, n_2 axes and the term R "compensates" simultaneously both of them on the opposite boundary. In this case, K_1, L_1 maybe determined from the fact that both primary terms are compensated by R and from the normalization condition, without recourse to numerical methods.

In the Erlang case, we will construct two compensation terms P_2, Q_2 for each of the primary trems and show that they coincide only in the $n = \infty$ case. Also, here the constants K_i, L_i are replaced by vectors.

The first compensation term P_2 of the $(\mathbf{1}, \mathbf{2})$ primary term on the $n_1 = 0$ boundary must be of the form $\mathbf{Z}(n_1, n_2) = \tilde{x}_1^{n_1} x_2^{n_2}\,\mathbf{z}$, where $\tilde{x}_1$ is the root of the equation $\tilde{x}_1 = \hat{F}(\mu_1(1 - \frac{\tilde{x}_1}{x_2}) + \mu_2(1 - x_2)))$ (so that the inner operator is satisfied).

The boundary operator $A_2(\pi + \mathbf{Z}) = 0$ yields $\mathbf{z} = -\mathbf{v}\big(T + \mu_2(x_2 - 1) + \mu_1\frac{x_1}{x_2}\big)\big(T + \mu_2(x_2 - 1) + \mu_1\frac{\tilde{x}_1}{x_2}\big)^{-1}$.

Similarly, the first compensation term Q_2 of the $(\mathbf{2}, \mathbf{1})$ primary term on the $n_2 = 0$ boundary must be of the form $\mathbf{z} y_1 \tilde{y}_2$, where $\tilde{y}_2$ is the root of the equation $\hat{F}(\mu_1(1 - \frac{y_1}{y_2}) + \mu_2(1 - y_2))) = \hat{F}(\mu_1(1 - \frac{y_1}{\tilde{y}_2}) + \mu_2(1 - \tilde{y}_2)))$ which simplifies to: $\tilde{y}_2\, y_2 = y_1\frac{\mu_1}{\mu_2}$.

Finally, $A_1(\pi + Z) = 0$ yields $\mathbf{z} = -\mathbf{v}\big(T + (\mu_2 y_2 - \mu_1)I + y_1^{-1}\mathbf{t}\alpha\big)\big(T + (\mu_2\tilde{y}_2 - \mu_1)I + y_1^{-1}\mathbf{t}\alpha\big)^{-1}$.

In general, further compensation terms are necessary. We note now a converse of Knessl result:

[2] We may choose a rectangle with the vertex at the origin, on which we aim to solve the equilibrium equations exactly; this gives rise to a system with an equal number of equations and unknowns; in addition, we have other equations on the boundary of the rectangle attempting to fit the "inner" stationary probabilities to the outer ones. For the latter, we assume the "compensation" form with parameters K_i. As long as the rectangle is large enough, this gives rise to an overdetermined system which may be solved by least squares, for example.

Lemma 4.3. *The Ph/M/M tandem has a compensation expansion with precisely three terms only in the (limiting) D/M/M case.*

Proof: For the compensation expansion to have three terms, the same compenstaing term must compensate both primary terms. This may only happen if $\tilde{x_1} = y_1$ and $\tilde{y_2} = x_2$ and leads to the system:

$$\begin{aligned} \mu_1 - \mu_1 x_1 &= \hat{F}^{-1}(x_1) \\ \mu_1(1 - \frac{y_1}{x_2}) + \mu_2(1 - x_2) &= \hat{F}^{-1}(y_1) \\ \mu_2 - \mu_2 y_2 &= \hat{F}^{-1}(y_2) \\ \mu_1 x_1 = \mu_2 x_2, x_2 y_2 &= y_1 \mu_1/\mu_2 \end{aligned}$$

From the last two equations we find that $y_1 = y_2 x_1$ and then from the second equation we find that

$$\hat{F}^{-1}(y_2 x_1) = (\mu_1 - \mu_2 x_2) + (\mu_2 - \mu_1 \frac{y_1}{x_2}) = \hat{F}^{-1}(x_1) + \hat{F}^{-1}(y_2)$$

This may happen iff the function $\hat{F}^{-1}$ is logarithmic and $\hat{F}$ is exponential, which happens iff the arrivals are deterministic.

In conclusion, in general more terms should be included in the compensation expansion and finally the two series should be combined via numerically determined parameters K_1, K_2. Another type of approximation in this problem (using only one primary term + unknown boundary stationary probabilities to be determined numerically) has been proposed by Luh [24].

5 Conclusions

Several methods of approximation for the stationary distribution of queueing networks have been developped in the last years. It would be quite interesting to compare these results with those obtained by the large deviations approach.

In particular, the compensation method proposed by Adan, Wessels and Zijm [1] may obtain from the large deviations approach:

- Explicit algebraic systems for finding the primary exponents.
- Criteria for when the primary terms are asymptotically dominant.
- Information on the order in which the boundaries are visited (and therefore where on where the primary exponents need to be compensated).

The large deviations-compensation approximation will result in a sum of several series of the form

$$\mathbb{P}\{X(t) \approx n\} \approx \sum_j K_j e^{-\alpha^{(j)(n)}\, n}$$

each of which starts with a primary coefficient. It is still not known currently how to find equations or analytic expressions for the primary coefficients K_i in multidimensional problems (this is called "sharp large deviations"); however, an easy numerical approach is available in the Markovian case. Namely, we assume the desired asymptotic form outside a fixed neighborhood of the origin and solve (for the primary coefficients K_i) the equilibrium equations of the network in this neighborhood (the least squares solution).

A second interesting line of research would be to investigate the solution of the Atar-Dupuis large deviations escape problem for more skip free or more general networks.

Acknowledgement: We thank D. Bertsimas, C. Knessl, I. Ignatyouk, A. Mogulskii, K. Majewski, P. Dupuis and the referee of this paper for valuable suggestions and comments.

References

[1] I. Adan, J. Wessels, W.H.M. Zijm, A compensation approach for two dimensional Markov processes, Adv. Appl. Prob. 25(1993) 783-817.

[2] I. Adan, Explicit solutions for queuing problems, Memorandum COSOR 97-05, 1997.

[3] R. Atar and P. Dupuis, Large deviations and queueing networks: methods for rate function identification, Stochastic Processes and their Applications 84(1999) 255-296.

[4] F. Avram, J. Hasenbein and J. Dai. Explicit Solutions for Variational Problems in the Quadrant, Queueing Systems 37(2001) 261-291.

[5] N. Bean, J. Li and P.G. Taylor, Some Asymptotic Properties of Two-stage Tandem Networks of PH/PH/1 Queues, in: A. Alfa and S. Chakravarthy (Editors), Matrix Analytic Methods for Stochastic Models, Notable Publications Inc, New Jersey, 1998, pp 115-134.

[6] A.A. Borovkov and A.A. Mogulskii, Large deviations for stationary Markov chains in the quarter plane, Proceedings of the Japan-Russia Symposium in probability, Tokyo, 1995.

[7] A.A. Borovkov and A.A. Mogulskii (2000). Large deviations for Markov chains in the quarter plane. To appear in "Uspehi Mat. Nauk"(in Russian).

[8] O. Boxma and A. Konheim, Approximate analysis of exponential queueing systems with blocking, Acta Inf. 15(1981) 19-66.

[9] C.S. Chang, Sample path large deviations and intree networks, Queueing Systems 20(1995) 7-36.

[10] A. Dembo and O. Zeitouni. *Large Deviations Techniques and Applications*, volume 38 of *Applications of Mathematics*. Springer, New York, 1998.

[11] R. Dobrushin and E. Pechersky Large deviations for tandem queueing systems *Journal of Applied Mathematics and Stochastic Analysis*, **7** (1994), 301-330.

[12] P. Dupuis, R. S. Ellis and A.Weiss. Large deviations for Markov processes with discontinuous statistics I. *Annals of Probability*, **19**(1991), 1280-1297.

[13] P. Dupuis and R. S. Ellis. Large deviations for Markov processes with discontinuous statistics II. *Probability Theory and Related Fields* **91** (1992), 153-194.

[14] P. Dupuis and R. S. Ellis. The large deviation principle for a general class of queueing systems. *Trans. Amer. Math. Soc.*, 347:2689–2751, 1996.

[15] P. Dupuis and H. Ishii. Large deviation behaviour of two competing queues. In *Proceedings of the 22nd Annual Conference on Information Sciences and Systems*, March 1988.

[16] K. Fujimoto and Y. Takahashi. Tail behavior of the steady state distribution in tandem queues. Journal of the O.R. Soc. of Japan, 39(4): pp 525-540, 1998.

[17] K. Fujimoto, Y. Takahashi and N. Makimoto, Asymptotic properties in tandem queues, Journal of the O.R. Soc. of Japan, 41, pp 118-141, 1998.

[18] A. Ghanesh and V. Anantharam, Stationary tail probabilities in exponential server tandems with renewal arrivals, Queueing Systems 22(1996), 203-247.

[19] I. Ignatyuk, V. Malyshev, and V. Scherbakov. Boundary effects in large deviations problems. *Russian Mathematical Surveys*, 49:41–99, 1994.

[20] I. Ignatyuk. Large deviations of Jackson networks, , Annals of Probability 10(2000), No.3, 962-1001.

[21] C. Knessl. On the diffusion approximation to a fork-join queuing model. *SIAM J. Appl. Math.*, 51:160–171, 1991.

[22] C. Knessl. An explicit solution to a tandem queueing model. *QUESTA*, 30:261-272, 1998.

[23] C. Knessl and C. Tier. A diffusion model for two tandem queues with general renewal input. *Comm. Stat. Stochastic Models*, 15:299–343, 1999.

[24] HP Luh. Product form solutions in Tandem Queues. *Journal of the O.R. Soc. of Japan*, 42:437–456, 1999.

[25] K. Majewski. Large deviations of stationary reflected Brownian motions. In F. P. Kelly, S. Zachary, and I. Ziedins, editors, *Stochastic Networks: Theory and Applications*. Royal Statistical Society, Oxford University Press, 1996.

[26] K. Majewski. Heavy traffic approximations of large deviations of feedforward queueing networks. *Queueing Systems: Theory and Applications*, 28:125–155, 1998.

[27] K. Majewski. Large deviations of the steady-state distribution of reflected processes with applications to queueing systems. *Queueing Systems: Theory and Applications*, 29:351–381, 1998.

[28] A.A. Mogulskii and B. Rogozin. Random Walks in the Positive Quadrant. I. Local Theorems". Matematicheskie Trudy, 1999, V.2, N2, P.57-97 (Engl. trans. in: SIBAM, 2000, V. ,N1, P.1-41.)

[29] A.A. Mogulskii and B. Rogozin. Random Walks in the Positive Quadrant. II. Integral Theorems". Matematicheskie Trudy, 2000, V.1, N1, P48—118. (Engl. trans. in: SIBAM, 2000, V. ,N2).

[30] A. Shwartz and A. Weiss. *Large deviations for performance analysis.* Chapman & Hall, New York, 1995.

[31] P. Tsoucas, Rare events in series of queues, J. Appl. Prob. 29, pp 168-175, 1992.

Vitae

Florin Avram was born on 5 June 1954, in Bucharest, Romania. He finished high school in Romania in 1973, the B.S. in Mathematics at the Tel Aviv University in 1976 and the Ph.D. in Applied Mathematics at Cornell University in 1986. His main interests are fluid and large deviations approximations for queueing networks, stochastic control and mathematical finance. Currently, he is affiliated with the Department of Actuarial Mathematics, Heriot Watt University, Edinburgh, Scotland and with Laboratoire de Mathèmatiques Appliquèes de l'Université de Pau et des Pays de l'Adour