



Wireless Data Performance in Multi-Cell Scenarios

Thomas Bonald, Sem Borst, Nidhi Hegde, Alexandre Proutière

► To cite this version:

Thomas Bonald, Sem Borst, Nidhi Hegde, Alexandre Proutière. Wireless Data Performance in Multi-Cell Scenarios. Sigmetrics, 2004, New York, United States. hal-01244814

HAL Id: hal-01244814

<https://hal.science/hal-01244814v1>

Submitted on 16 Dec 2015

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Wireless Data Performance in Multi-Cell Scenarios

T. Bonald*, S. Borst†, N. Hegde‡, A. Proutière*

ABSTRACT

The performance of wireless data systems has been extensively studied in the context of a single base station. In the present paper we investigate the flow-level performance in networks with multiple base stations. We specifically examine the complex, dynamic interaction of the number of active flows in the various cells introduced by the strong impact of interference between neighboring base stations. For the downlink data transmissions that we consider, lower service rates caused by increased interference from neighboring base stations result in longer delays and thus a higher number of active flows. This in turn results in a longer duration of interference on surrounding base stations, causing a strong correlation between the activity states of the base stations. Such a system can be modelled as a network of multi-class processor-sharing queues, where the service rates for the various classes at each queue vary over time as governed by the activity state of the other queues. The complex interaction between the various queues renders an exact analysis intractable in general. A simplified network with only one class per queue reduces to a coupled-processors model, for which there are few results, even in the case of two queues. We thus derive bounds and approximations for key performance metrics like the number of active flows, transfer delays, and flow throughputs in the various cells. Importantly, these bounds and approximations are insensitive, yielding simple expressions, that render the detailed statistical characteristics of the system largely irrelevant.

Categories and Subject Descriptors

C.2.1 [Computer-Communication Networks]: Network Architecture and Design—*Wireless communication*;

*France Telecom R&D, France,
{thomas.bonald,alexandre.proutiere}@francetelecom.com

†CWI, The Netherlands, Sem.Borst@cw.nl, also affiliated with Bell laboratories, Lucent Technologies

‡CWI, The Netherlands and currently with EURANDOM, The Netherlands, Hegde@eurandom.tue.nl

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

SIGMETRICS/Performance'04 June 12–16, 2004, New York, NY, USA.
Copyright 2004 ACM 1-58113664-1/04/0006 ...\$5.00.

G.3 [Mathematics of Computing]: Probability and Statistics—*Queueing theory, Stochastic processes*

General Terms

Performance

Keywords

Wireless data networks, elastic traffic, time-varying service, fluid regime, quasi-stationary regime, multi-class processor-sharing, stability, insensitivity.

1. INTRODUCTION

Next-generation wireless networks are expected to support a wide variety of high-speed data applications, in addition to conventional voice services and current low-bandwidth data services such as short messaging. The integration of these heterogeneous applications on a common transmission infrastructure raises similar challenges as in wireline integrated networks. In wireless environments, these issues are further exacerbated by interference problems, the intrinsically limited bandwidth, and the highly variable and unpredictable propagation characteristics. Specifically, the channel quality may vary widely among spatially distributed users due to distance-related attenuation. In addition, the channel conditions for a given user may vary dramatically over time because of fading effects.

Wireless circuit-switched voice networks rely on power control mechanisms for adjusting the transmit power to compensate for the varying channel quality and maintain a fixed transmission rate. Various data applications on the other hand, such as file transfers and Web browsing sessions, are less sensitive to packet-level delays, and do not have a stringent rate requirement. Such *elastic* applications are well-suited for rate control algorithms which dynamically adapt the transmission rate over time so as to match the fluctuations in the channel quality and available transmission resources (transmit power and time slots). In addition, the fraction of transmission resources allocated to the individual data users can be adapted as the number of active users varies over time. In case the transmissions are ‘orthogonal’ (e.g. time sharing or power sharing with orthogonal code words) and the available resources are constant over time and evenly shared among the active data users, the flow-level performance may be described by a Processor-Sharing model, with the service requirements of the various users normalized by their time-average transmission rates. Notably, the performance is then *insensitive*, in the sense that

the distribution of the number of active flows, the delays, and the throughputs only depend on the statistical characteristics of the system through easily calculated load factors.

The variations in the channel quality in fact open up the possibility of scheduling the data transmissions to the various users when their instantaneous transmission rates are relatively high [2, 5, 15, 16, 17]. While fading is considered to have a predominantly adverse impact for constant-rate voice connections, it thus provides the opportunity to achieve throughput gains for elastic data transfers. An important example of such a *channel-aware* scheduling policy is the Proportional Fair algorithm for the CDMA 1xEV-DO system [14]. In certain cases the flow-level performance of the Proportional Fair algorithm may also be modeled by a Processor-Sharing type queue, but with a state-dependent service rate that accounts for the throughput gains from channel-aware scheduling [6, 8].

All the above papers study scenarios with just a *single* isolated base station (BS). In the present paper we investigate the flow-level performance in data networks with *multiple* BSs. We assume that the BSs transmit at full power as long as there are any active users, and reduce the power to zero otherwise. We further suppose that within each cell the transmissions are orthogonal and that the available resources are shared in a fair manner among the active users. Between cells, however, the transmissions do interfere, and we explicitly take into account the complex, dynamic interaction of the number of active flows in each cell caused by the changing activity patterns of neighboring BSs. The resulting model amounts to a network of multi-class Processor-Sharing queues, where the service rates for the various classes at each queue vary over time as governed by the activity state of the other queues. The intricate correlation among the various BSs renders an exact analysis elusive in general. In the single-class case, the model reduces to a so-called *coupled-processors* model, which even for two queues is barely tractable [11, 13], reflecting the complexity of the model in general. Therefore, we focus on the derivation of bounds and approximations.

We obtain two types of bounds: (i) the first-degree bounds assume minimum and maximum interference incurred by the cell under consideration itself; (ii) the second-degree bounds assume minimum and maximum interference experienced by the *neighboring* cells, but allow the cell under consideration to be influenced by its neighbors. While the first-degree bounds are relatively rough and can be explicitly computed, the second-degree bounds are sharper, but correspond to a Processor-Sharing queue in a time-varying environment, and cannot be calculated analytically. In order to obtain closed-form estimates for the latter type of bounds, we introduce two limit regimes, termed *fluid* and *quasi-stationary* regime, where the time-varying environment evolves on a very fast and a very slow time scale, respectively. These two limit regimes provide explicit formulas for the distribution of the number of active flows, the mean transfer delays, and the flow throughputs. Importantly, the performance in both limit regimes is insensitive, and only depends on easily computed load factors, thus yielding simple estimates that render the detailed traffic properties and propagation features largely irrelevant. We will specifically consider lower bounds for the number of active flows evaluated in the fluid regime and upper bounds based on the quasi-stationary regime. Some theoretical arguments as well as the numerical ex-

periments that we perform for various network topologies indicate that the former tend to be optimistic, while the latter tend to be conservative and quite often remarkably sharp approximations.

The remainder of the paper is organized as follows. In Section 2 we present a detailed model description. We collect some useful preliminaries in Section 3. In Section 4 we derive sample-path bounds for the number of active flows in the various cells. In Section 5 we introduce the fluid and quasi-stationary regimes mentioned above. We examine some stability issues in Section 6. In Section 7 we discuss the numerical experiments that we conducted to illustrate the results. In Section 8 we make some concluding remarks.

2. MODEL DESCRIPTION

We consider a network of BSs, indexed by the set \mathcal{N} . Each BS provides service to data users in the corresponding cell through a shared downlink.

Radio environment. In evaluating the flow-level performance, it is useful to characterize the radio environment in terms of feasible transmission rates of the various users. By feasible rate, we mean the long-term transmission rate that a user would receive if it were the only user in the cell. The feasible rate for a given user depends on its channel quality, which in turn depends on its relative location to the serving BS in conjunction with the propagation characteristics. Because of inter-cell interference, the channel quality additionally depends on the relative position of the user to neighboring BSs, and whether these surrounding BSs are transmitting or not.

Rather than modeling these system-specific aspects explicitly, it will be convenient to introduce a class structure to capture the dependence of the feasible rates on the propagation characteristics, the network topology and the spatial configuration of the users. The classes encapsulate these detailed properties, and implicitly correspond to specific sub-regions of the coverage area. Specifically, we assume that the users in cell i may be categorized into K_i classes, and denote by $C_{i,\mathcal{A}}(k)$ the feasible rate of class- k users in cell i when the set of active BSs is $\mathcal{A} \subseteq \mathcal{N}$. We will frequently omit the index k and just write $C_{i,\mathcal{A}}$ when the class identity is not essential.

We assume that the feasible rates satisfy a natural monotonicity property: the larger the set of active BSs, and thus the larger the inter-cell interference, the smaller the feasible rate, i.e.,

$$\mathcal{A} \subseteq \mathcal{B} \quad \Rightarrow \quad C_{i,\mathcal{A}} \geq C_{i,\mathcal{B}}. \quad (1)$$

In the definition of $C_{i,\mathcal{A}}$ it is immaterial whether or not cell i is included in the set \mathcal{A} , and it will be convenient to adopt the convention $C_{i,\mathcal{A}} \equiv C_{i,\mathcal{A} \cup \{i\}} \equiv C_{i,\mathcal{A} \setminus \{i\}}$.

Traffic characteristics. Class- k flows arrive to cell i as a Poisson process of rate $\lambda_i(k)$. Each flow is associated with a unique serving BS for its entire duration, and in particular we do not consider hand-offs between cells. Note that hand-offs will tend to be rare for data transfers, since they predominantly result from user mobility, which typically occurs on a relatively slow time scale compared to a typical flow duration. The flow sizes are independent and identically distributed with mean σ . (The results trivially extend to the case where the flow sizes are also cell- and class-dependent.)

Radio resource sharing. We recall here that we consider data traffic, which is not power-controlled, as is voice traffic, but instead is rate-controlled. In particular, we assume that the BSs operate at full power as long as there are any active flows to serve, and use zero power otherwise. Within each cell transmissions are orthogonal and the available radio resources (time slots or transmit power) are evenly shared among the active flows, obviously of the instantaneous channel conditions and the activity state of the neighboring BSs. Thus, when the set of active BSs is \mathcal{A} and the number of active flows in cell i is n_i , the actual transmission rate of class- k flows in cell i is

$$C_{i,\mathcal{A}}(k)/n_i. \quad (2)$$

The analysis may be generalized to the case where the actual transmission rate of class- k flows in cell i is $C_{i,\mathcal{A}}(k)G_i(n_i)/n_i$. The function $G_i(\cdot)$ may be interpreted as a gain factor accounting for the throughput improvements from channel-aware scheduling algorithms, which opportunistically allocate radio resources taking into account the instantaneous channel conditions of the various users [2, 5, 8, 15].

REMARK 1. *Since the nominal transmission rates in an actual system are constrained to a discrete set, we consider a finite number of user classes, i.e., $K_i < \infty$ for all $i \in \mathcal{N}$. However, the results readily extend to a continuum of user classes (corresponding to infinitesimally small subregions).*

REMARK 2. *In some current implementations such as the CDMA 1xEV-DO system, the selection of the transmission rate is not guided by the actual interference experienced on the traffic channel, but rather based on the full interference inferred from pilot channel measurements. However, the effectively received rate may still improve when interfering BSs are inactive due to the use of so-called incremental redundancy. This mechanism consists in transmitting packets in a succession of slots until either the packet is successfully decoded or a certain maximum number of slots is reached in which case an error is declared. Although the selection of the packet size is based on the measured interference on the pilot channel, the expected number of required slots and the probability of error mainly depend on the actual interference encountered on the traffic channel.*

REMARK 3. *The need to explicitly distinguish between several user classes arises from the fact that the feasible rates in different locations are impacted by inter-cell interference in a non-uniform manner. The latter property in fact provides a potential incentive to allocate radio resources based on the activity state of the neighboring BSs. For example, users on the boundary of the cell are affected disproportionately by inter-cell interference compared to users in the middle of the cell. Thus, it pays to schedule users on the periphery when few neighboring BSs are active, and schedule users in the center when many surrounding BSs are active. Note that the rationale behind such intra-cell scheduling strategies is somewhat reminiscent of the objective of channel-aware scheduling algorithms to exploit multi-user diversity caused by independent fading.*

REMARK 4. *The actual data performance over wireless links is also strongly affected by the interaction between the TCP congestion control mechanism and the high latency and*

high loss rates. The latter issue is beyond the scope of the present paper, but has been widely studied in the literature, where various approaches to improve the TCP performance in wireless networks have been proposed, see for instance [3, 9].

3. PRELIMINARIES

The model described in the previous section corresponds to a network of multi-class Processor-Sharing queues, where the service rates for the various classes at each queue vary over time as governed by activity state of the other queues. The complex interaction between the various queues renders an exact analysis impractical in general. In the single-class case, i.e., $K_i = 1$ for all $i \in \mathcal{N}$, the model reduces to a so-called *coupled-processors* model. The latter model has been studied in detail for the case of two queues. Fayolle & Iasnogorodski [13] showed that in the case of exponentially distributed service times the analysis of the joint queue length distribution may be formulated as a Riemann-Hilbert problem. Boxma & Cohen [11] considered the case of generally distributed service times, and showed that the joint workload distribution may be obtained as the solution to a boundary value problem. For the case of more than two queues, hardly any results are known [10].

The fact that even the single-class two-queue case is barely tractable, testifies to the complexity of the model in general. Therefore, we focus in the next sections on the derivation of bounds and approximations. The numerical experiments presented later indicate that the bounds and approximations are quite often remarkably tight.

As mentioned above, the complexity of the model stems from the fact that the service rate at each queue is time-varying, depending on the activity state of the other queues. In deriving the bounds and approximations, it will be convenient to consider a reference system for a cell in isolation where the activity state of the other cells is fixed, so that the service rates are fixed as well. Specifically, if the service rates of class- k flows in cell i are always $C_{i,\mathcal{A}}(k)$, then the cell in isolation behaves as a standard multi-class Processor-Sharing system with load $\rho_{i,\mathcal{A}} = \sum_{k=1}^{K_i} \rho_{i,\mathcal{A}}(k)$, where

$$\rho_{i,\mathcal{A}}(k) = \frac{\lambda_i(k)\sigma}{C_{i,\mathcal{A}}(k)}.$$

In particular, the reference system is stable if and only if $\rho_{i,\mathcal{A}} < 1$, in which case the stationary distribution of the number of active flows of the various classes is given by

$$\pi(m_1, \dots, m_{K_i}) = (1 - \rho_{i,\mathcal{A}}) \left(\prod_{k=1}^{K_i} \frac{m_k}{m_1 \dots m_{K_i}} \right) \prod_{k=1}^{K_i} \rho_{i,\mathcal{A}}(k)^{m_k}.$$

Thus, the stationary distribution of the total number of active flows is

$$\pi_{i,\mathcal{A}}(n) = (1 - \rho_{i,\mathcal{A}}) \rho_{i,\mathcal{A}}^n, \quad (3)$$

and the mean number of active class- k flows is

$$E[n_{i,\mathcal{A}}(k)] = \frac{\rho_{i,\mathcal{A}}(k)}{1 - \rho_{i,\mathcal{A}}}. \quad (4)$$

Using Little's law, we obtain that the mean class- k delay is

$$E[T_{i,\mathcal{A}}(k)] = \frac{E[n_{i,\mathcal{A}}(k)]}{\lambda_i(k)} = \frac{\sigma}{(1 - \rho_{i,\mathcal{A}})C_{i,\mathcal{A}}(k)},$$

and the class- k flow throughput, defined as the ratio of the mean flow size to the mean flow delay, is

$$\gamma_{i,A}(k) = \frac{\sigma}{E[T_{i,A}(k)]} = (1 - \rho_{i,A})C_{i,A}(k). \quad (5)$$

4. PERFORMANCE BOUNDS

We now present some relatively simple lower and upper bounds for the number of active flows in the various cells. The bounds rely on the basic property that increasing or decreasing the service rates in a given cell decreases or increases the number of active flows in this cell, as formalized in the next proposition. Let $n_i(t)$ be the number of active flows in cell i at time t . Let $\mathcal{A}(t)$ be the set of active BSs at time t . Let $n_i^-(t)$ and $n_i^+(t)$ be the number of active flows in two reference systems where the service rates of the users in cell i are given by $C_i^-(t)$ and $C_i^+(t)$, respectively.

PROPOSITION 1. *If $C_i^-(t) \geq C_{i,A(t)} \geq C_i^+(t)$ for all t , then $n_i^-(t) \leq n_i(t) \leq n_i^+(t)$ for all t .*

PROOF. The proof follows from (2) in combination with sample-path arguments. \square

The above proposition provides a method for constructing various bounds by making specific choices for $C_i^-(t)$ and $C_i^+(t)$.

First-degree bounds. We first derive rather crude bounds for the number of active flows, obtained by assuming minimum or maximum service rates in cell i , i.e., maximum or minimum interference from other cells. Let $n_{i,A}(t)$ be the number of active flows in cell i at time t when the service rates are always $C_{i,A}$.

PROPOSITION 2. *The number of active flows in cell i at time t is larger (resp. smaller) than that obtained when the service rates in cell i are always $C_{i,\emptyset}$ and $C_{i,\mathcal{N}}$, respectively, i.e.,*

$$n_{i,\emptyset}(t) \leq n_i(t) \leq n_{i,\mathcal{N}}(t), \quad \forall t.$$

PROOF. Since $\emptyset \subseteq \mathcal{A}(t) \subseteq \mathcal{N}$ at any time t , we deduce from (1) that

$$C_{i,\emptyset} \geq C_{i,A(t)} \geq C_{i,\mathcal{N}}, \quad \forall t.$$

The statement then follows from Proposition 1 by taking $C_i^-(t) = C_{i,\emptyset}$ and $C_i^+(t) = C_{i,\mathcal{N}}$. \square

Second-degree bounds. We now derive tighter bounds for the number of active flows by assuming minimum or maximum service rates in the other cells $j \neq i$, but allowing the service rates in cell i itself to be influenced by the activity state of the other cells.

Let $n'_{i,\emptyset}(t)$ and $n'_{i,\mathcal{N}}(t)$ be the number of active flows in cell i at time t , assuming that the service rates in the cells $j \neq i$ are always $C_{j,\emptyset}$ and $C_{j,\mathcal{N}}$, respectively. Let $\mathcal{A}_\emptyset(t) = \{j \in \mathcal{N} \setminus \{i\} : n_{j,\emptyset}(t) > 0\}$ and $\mathcal{A}_\mathcal{N}(t) = \{j \in \mathcal{N} \setminus \{i\} : n_{j,\mathcal{N}}(t) > 0\}$ be the sets of active BSs at time t when the service rates in the cells $j \neq i$ are always $C_{j,\emptyset}$ and $C_{j,\mathcal{N}}$, respectively. Noting that the cells $j \neq i$ then behave independently, we have that the stationary distribution of these sets is given by

$$\Pr[\mathcal{A}_\emptyset(t) = \mathcal{B}] = \prod_{j \in \mathcal{B}} (1 - \pi_{j,\emptyset}(0)) \times \prod_{j \notin \mathcal{B}} \pi_{j,\emptyset}(0),$$

and

$$\Pr[\mathcal{A}_\mathcal{N}(t) = \mathcal{B}] = \prod_{j \in \mathcal{B}} (1 - \pi_{j,\mathcal{N}}(0)) \times \prod_{j \notin \mathcal{B}} \pi_{j,\mathcal{N}}(0),$$

for all $\mathcal{B} \subseteq \mathcal{N} \setminus \{i\}$, with the convention $\pi_{j,A}(0) = 0$ if $\rho_{j,A} \geq 1$.

PROPOSITION 3. *The number of active flows in cell i at time t is larger (resp. smaller) than that obtained when the service rates in the cells $j \neq i$ are always $C_{j,\emptyset}$ and $C_{j,\mathcal{N}}$, respectively, i.e.,*

$$n'_{i,\emptyset}(t) \leq n_i(t) \leq n'_{i,\mathcal{N}}(t), \quad \forall t.$$

PROOF. Using Proposition 2 for all $j \neq i$, we find that $\mathcal{A}_\emptyset(t) \subseteq \mathcal{A}(t) \subseteq \mathcal{A}_\mathcal{N}(t)$ at any time t , and hence we deduce from (1) that

$$C_{i,\mathcal{A}_\emptyset(t)} \geq C_{i,A(t)} \geq C_{i,\mathcal{A}_\mathcal{N}(t)}, \quad \forall t.$$

The statement then follows from Proposition 1 by taking $C_i^-(t) = C_{i,\mathcal{A}_\emptyset(t)}$ and $C_i^+(t) = C_{i,\mathcal{A}_\mathcal{N}(t)}$. \square

5. LIMITING REGIMES

The bounds derived in the previous section are much simpler than the original performance measures since the actual set of active BSs is replaced by a virtual set that evolves independently of the number of active flows in the considered cell. The second-degree bounds are much finer than the first-degree ones but intractable in general. The difficulty arises from the fact that the number of active flows $n'_{i,\emptyset}(t)$ (resp. $n'_{i,\mathcal{N}}(t)$) evolves like the number of customers in a multi-class Processor-Sharing queue with correlated per-class service rate variations, driven by a random environment described by the process $\mathcal{A}_\emptyset(t)$ (resp. $\mathcal{A}_\mathcal{N}(t)$). In this section, we introduce approximations of the bounds based on two limit regimes, termed *fluid* and *quasi-stationary*, where the processes $\mathcal{A}_\emptyset(t)$ and $\mathcal{A}_\mathcal{N}(t)$ evolve on an infinitely fast and an infinitely slow time scale, respectively. Such approximations turn out to be powerful in analyzing multi-class Processor-Sharing systems (they were previously applied to systems modeling the integration of streaming and data traffic in wireline networks [12], and more recently to evaluate the performance impact of user mobility in wireless data systems [7]).

Here we consider an arbitrary stationary and ergodic process $\mathcal{A}'(t)$ describing the evolution of the set of active BSs. We assume this process is independent of the state of the considered cell i . This process might, for example, represent $\mathcal{A}_\emptyset(t)$ or $\mathcal{A}_\mathcal{N}(t)$. Consider now a family of systems, parametrized by $s \in (0, \infty)$ and obtained by replacing the process $\mathcal{A}'(t)$ by $\mathcal{A}'(s \times t)$. The parameter s represents the speed of variations in the set of active BSs.

Quasi-stationary regime. The quasi-stationary regime is obtained when the speed of variations s tends to 0. In the limit for $s \rightarrow 0$, the set of active BSs is frozen to its initial state. Thus the quasi-stationary regime corresponds to the system where the set of active BSs is constant and equal to \mathcal{B} with probability $\Pr[\mathcal{A}'(0) = \mathcal{B}]$.

Assuming $\rho_{i,\mathcal{N}} < 1$, we deduce from (3) the distribution of the number of active flows in cell i in the quasi-stationary regime:

$$\pi_i^{\text{qs}}(n) = E[\pi_{i,\mathcal{A}'}(n)] = E[(1 - \rho_{i,\mathcal{A}'}^n) \rho_{i,\mathcal{A}'}^n],$$

where the expectation is taken with respect to the distribution of $\mathcal{A}'(0)$.

In view of (4), the mean number of active class- k flows is:

$$\mathbb{E}[n_i^{\text{qs}}(k)] = \mathbb{E} \left[\frac{\rho_{i,\mathcal{A}'}(k)}{1 - \rho_{i,\mathcal{A}'}} \right].$$

We deduce the class- k flow throughput:

$$\gamma_i^{\text{qs}}(k) = \frac{\lambda_i(k)\sigma}{\mathbb{E}[n_i^{\text{qs}}(k)]} = \mathbb{E}[(C_{i,\mathcal{A}'}(k)(1 - \rho_{i,\mathcal{A}'})^{-1})^{-1}].$$

For example, the class- k flow throughput in the quasi-stationary regime of the upper bound of Proposition 3 is given by:

$$\gamma_i^{\text{qs}}(k) = \left(\sum_{\substack{\mathcal{B}: \\ \mathcal{N}' \subset \mathcal{B} \subset \mathcal{N} \setminus \{i\}}} \prod_{j \in \mathcal{B} \setminus \mathcal{N}'} \rho_{j,\mathcal{N}} \prod_{j \notin \mathcal{B}} (1 - \rho_{j,\mathcal{N}}) \times (C_{i,\mathcal{B}}(k)(1 - \rho_{i,\mathcal{B}}))^{-1} \right)^{-1},$$

where $\mathcal{N}' = \{j \in \mathcal{N} \setminus \{i\} : \rho_{j,\mathcal{N}} \geq 1\}$.

Fluid regime. The fluid regime is obtained when the speed of variations s tends to ∞ . In the limit for $s \rightarrow \infty$, the set of active BSs evolves so rapidly that each class- k flow in cell i sees a constant feasible rate, equal to his mean feasible rate $\mathbb{E}[C_{i,\mathcal{A}'}(k)]$. Thus the fluid regime is the system where the set of active BSs is constant and equal to $\mathbb{E}[\mathcal{A}'(0)]$.

The traffic load of cell i in the fluid regime is given by:

$$\rho_i^{\text{fl}} = \sum_{k \in \mathcal{K}_i} \frac{\lambda_i(k)\sigma}{\mathbb{E}[C_{i,\mathcal{A}'}(k)]}. \quad (6)$$

If $\rho_i^{\text{fl}} < 1$, the distribution of the number of active flows is then:

$$\pi_i^{\text{fl}}(n) = (1 - \rho_i^{\text{fl}}) \left(\rho_i^{\text{fl}} \right)^n.$$

The mean number of active class- k flows is:

$$\mathbb{E}[n_i^{\text{fl}}(k)] = \frac{\rho_i^{\text{fl}}(k)}{1 - \rho_i^{\text{fl}}}, \quad \text{with } \rho_i^{\text{fl}}(k) = \frac{\lambda_i(k)\sigma}{\mathbb{E}[C_{i,\mathcal{A}'}(k)]}.$$

We deduce the class- k flow throughput:

$$\gamma_i^{\text{fl}}(k) = \frac{\lambda_i(k)\sigma}{\mathbb{E}[n_i^{\text{fl}}(k)]} = \mathbb{E}[C_{i,\mathcal{A}'}(k)](1 - \rho_i^{\text{fl}}).$$

A question of importance is whether these quasi-stationary and fluid regimes provide actual bounds for the performance of the original system. From recent studies on single-class Processor Sharing queues with time-varying capacity [12], the performance in the quasi-stationary (resp. fluid) regime is worse (resp. better) than that of the actual system. Thus for any cell i that contains a single class, the quasi-stationary regime of the upper bound $n_{i,\mathcal{N}}(t)$ is itself an upper bound for the number of active flows in cell i , while the fluid regime of the lower bound $n_{i,\emptyset}(t)$ is itself a lower bound. It proves extremely difficult to extend this result to multi-class cells. The numerical results presented in Section 7 suggest however that the result still holds.

6. STABILITY CONDITION

We now examine the stability of the system. The following stability condition is based on the first-degree bounds derived in Section 4.

PROPOSITION 4. *If $\rho_{i,\mathcal{N}} < 1$, then cell i is stable.*

PROOF. As $\rho_{i,\mathcal{N}} < 1$ is the necessary and sufficient condition for the stability of cell i in the reference system where the set of active BSs is \mathcal{N} , the result is a direct consequence of Proposition 2. \square

The second-degree bounds provide a tighter stability condition. As in Section 5, we first consider a virtual system where the set of active BSs is given by an independent stationary and ergodic process $\mathcal{A}'(t)$. Let $n'_i(t)$ be the number of active flows in cell i at time t and ρ_i^{fl} the traffic load in the fluid regime, given by (6). The following result can be proved as Theorem 1 in [7].

LEMMA 1. *If $\rho_i^{\text{fl}} < 1$, then the stochastic process $n'_i(t)$ is stable.*

Let $\rho_{i,\mathcal{N}}^{\text{fl}}$ be the load of the second-degree upper bound in the fluid regime:

$$\rho_{i,\mathcal{N}}^{\text{fl}} = \sum_{k \in \mathcal{K}_i} \frac{\lambda_i(k)\sigma}{\mathbb{E}[C_{i,\mathcal{A}'}(k)]}.$$

Using Lemma 1 and Proposition 3, we immediately obtain:

PROPOSITION 5. *If $\rho_{i,\mathcal{N}}^{\text{fl}} < 1$, then cell i is stable.*

Note that Proposition 5 is stronger than Proposition 4 since:

$$\rho_{i,\mathcal{N}}^{\text{fl}} \leq \sum_{k \in \mathcal{K}_i} \frac{\lambda_i(k)\sigma}{C_{i,\mathcal{N}}(k)} = \rho_{i,\mathcal{N}}.$$

REMARK 5. *The sufficient stability conditions of Propositions 4 and 5 have their natural counterpart in the form of necessary conditions, but these are likely to be loose in practice.*

It follows from Proposition 5 that a sufficient condition for the stability of the overall network is:

$$\forall i \in \mathcal{N}, \quad \rho_{i,\mathcal{N}}^{\text{fl}} < 1.$$

For a two-cell network, it may be verified that this condition is in fact also necessary. In the case of a single class per cell, this follows from the stability condition of a two-processor coupled system [11]. It proves extremely difficult to derive the exact stability condition when the network contains more than two cells. This may notably be explained by the fact that the stability condition *does* generally depend on the flow size distribution. Consider a network of three cells for instance, with a very large number of active flows in one cell and a small number of active flows in the other two cells. The mean service rate of active flows in the “full” cell depends on the steady state of the other two cells, which is indeed sensitive to the flow size distribution (the two cells behave as a two-processor coupled system). Thus the stability condition is itself sensitive. We can give a general necessary stability condition, however.

PROPOSITION 6. *If the network is stable, then there is at least one cell i such that $\rho_{i,\mathcal{N}} < 1$.*

PROOF. Assume that $\rho_{j,\mathcal{N}} \geq 1$ for all cells j . Starting from any initial state where the set of active BSs $\mathcal{A}(0)$ is the entire set \mathcal{N} , i.e., all cells have at least one active flow, the number of active flows in cell i behaves independently of the number of active flows in other cells as long as $\mathcal{A}(t) = \mathcal{N}$.

Specifically, the number of active flows in cell i evolves like the number of customers in a multi-class Processor-Sharing queue of load $\rho_{i,\mathcal{N}}$ (cf. Section 3). As $\rho_{i,\mathcal{N}} \geq 1$, there is a positive probability that this queue never empties. Since the number of cells is finite, there is a positive probability that no cell empties. The system is unstable. \square

In view of Propositions 4–6, we deduce the stability condition of networks with homogeneous loads, i.e., such that the traffic load $\rho_{i,\mathcal{N}}$ of cell i when other cells are always active is the same for all i . In this case, the necessary and sufficient stability condition simply reads:

$$\forall i \in \mathcal{N}, \quad \rho_{i,\mathcal{N}} < 1.$$

Such *homogeneous* networks may represent regular topologies (like hexagonal networks) with uniform traffic distribution, or networks with non-regular topologies but planned in such a way that large cells cover areas with lower traffic density, which is indeed the case in well-engineered networks.

7. NUMERICAL RESULTS

The results of the previous sections hold in quite general settings. In this section we make some specific assumptions in order to conduct numerical experiments. In particular, we consider a set of feasible rates in line with the High Data Rate (HDR) standard [5, 14] which underlies the CDMA 1xEV-DO system. These results can, however, be easily extended to UMTS HSDPA as well.

7.1 Experimental setting

Radio environment. As mentioned in Section 2, the feasible rate of a user depends on his position and the set of active BSs through his signal-to-noise-and-interference ratio (SNR). As mentioned above, we will consider a discrete set of rates as defined in the HDR standard. Table 1 shows the typical SNR requirements for the various rates at a one percent frame-error rate based on AWGN channel assumptions [5]. This table also lists the corresponding E_b/N_0 values. Note that the target E_b/N_0 values are approximately constant for all data rates except the three highest rates. For convenience, we assume a constant E_b/N_0 target of 2.5 dB for the numerical experiments, thus linearizing the SNR thresholds. Since the users with the highest data rates are least demanding of transmission resources, it is reasonable to expect that slightly altering their target E_b/N_0 values should not drastically affect the overall performance characteristics of the system.

The SNR of a user is in turn a function of his position and the interference from active BSs, as follows:

$$\text{SNR}_{i,\mathcal{A}}(u) = \frac{P_i(u)}{\eta + I_{i,\mathcal{A}}(u)},$$

where $P_i(u)$ is the power received at position u from the serving BS i , η is the background noise, and $I_{i,\mathcal{A}}(u)$ is the interference caused by the set of active BSs \mathcal{A} . The received power $P_i(u)$ includes signal attenuation due to fading and path loss. We do not consider fading effects in this paper. The received power then is given by $P_i(u) = P\Gamma_i(u)$, where P is the power transmitted by a BS, assumed to be the maximum transmit power and identical for all BSs, and $\Gamma_i(u)$ is the path loss function. Following standard models, the path

v	Rate c_v kb/s	SNR dB	E_b/N_0 dB
0	2457.6	9.5	6.49
1	1843.2	7.2	5.44
2	1228.8	3.0	3.00
3	921.6	1.3	2.55
4	614.4	-1.0	2.01
5	307.2	-4.0	2.02
6	204.8	-5.7	2.08
7	153.6	-6.5	2.53
8	102.6	-8.5	2.28
9	76.8	-9.5	2.54
10	38.4	-12.5	2.55

Table 1: Feasible rates in HDR systems.

loss function is given by:

$$\Gamma_i(u) = \begin{cases} 1 & \text{if } d(i, u) \leq \epsilon, \\ \left(\frac{\epsilon}{d(i, u)}\right)^\alpha & \text{otherwise,} \end{cases} \quad (7)$$

where ϵ is a reference distance very close to BS i where full power is received, $d(i, u)$ is the distance from position u to BS i , and α is the path loss exponent. The path loss exponent depends on the propagation environment, and has typical values between 2 and 5. The interference term is given by:

$$I_{i,\mathcal{A}}(u) = P \sum_{\substack{j \in \mathcal{A} \\ j \neq i}} \Gamma_j(u).$$

Network topology. We will consider two types of networks: linear networks with equidistant BSs placed on a line, and regular hexagonal networks as shown in Figure 2. For the latter type of networks, we further distinguish between cells with omni-directional antennas, referred to as omni-cells here, as shown in Figure 2 and sectorized cells as shown in Figure 11. We first consider omni-cells, and sectorized cells will be discussed further at the end of this section. R denotes the farthest distance covered by a BS, so that for linear networks, the distance between adjacent BSs is $2R$ and for hexagonal networks, cells are regular hexagons with sides of length R and the distance between neighboring BSs is $\sqrt{3}R$. Note that the framework of our model is general enough to allow for different sized or shaped cells. For greater transparency we provide here examples of how this model can be used for certain symmetric topologies.

Linear networks as shown in Figure 1 model highway or street scenarios and users are placed on a line on either side of the BS. In hexagonal networks as shown in Figure 2, users are placed on the plane. A user's position is represented in polar coordinates by (r, θ) , with a reference BS at $(0, 0)$. For linear networks, $\theta = 0$ and $\theta = \Pi$ correspond to the two sides of a BS.

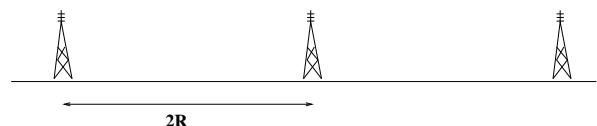


Figure 1: BSs on a line.

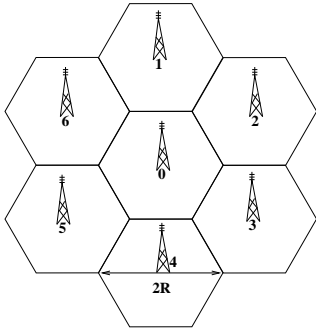


Figure 2: Hexagonal network.

Discrete rates and classes. Recall that the feasible rate of a user is dependent on his position as well as the set of active BSs. For a given set of active BSs then, the cell can be partitioned into regions such that all users in a given region have identical feasible rates. Denote $r_{v,\mathcal{A}}(\theta)$ to be the farthest distance from the BS where a rate of c_v can be achieved with the set of active BSs \mathcal{A} , along the angle θ from the BS. For a linear network $\theta = 0$ or Π and $r_{v,\mathcal{A}}(\cdot) \leq R$. When $\mathcal{A} = \emptyset$ this results in symmetric spacings of r_v on either side, and similarly for $\mathcal{A} = \mathcal{N}$. The sizes and number of these regions change with the set \mathcal{A} , as shown in Figure 3, with the normalization $r_{0,\emptyset}(\cdot) = 1$. We will assume throughout this section that inter-cell interference is due only to direct neighbors. This approximation is quite common, and is reasonable given the fall-off in path loss with distance under the power-law decay assumed in (7). For a cell with normalized radius $R = 3$ and path loss exponent $\alpha = 2$, Figure 3 shows the regions of possible feasible rates for all four possible sets of active BSs. The words ON and OFF on each side of the cell in Figure 3 denote the activity of the BS on that respective side. We observe that for a user near $r = 1.5$, the rate is reduced from 1228.8 kb/s to 921.6 kb/s as the active set changes from no neighbor being active to both being active. For a similar change in active set, a user near $r = 1.9$ experiences a feasible rate reduced in half from 614.4 kb/s to 307.2 kb/s. The effect of inter-cell interference is to shrink all regions toward the BS. Thus, changes in feasible rates are considerable for certain positions, especially users who are close to the edge of the cell, with lower data rates. These users are the most demanding on the transmission resources and are affected the most by the rate changes. Such rate changes, therefore, are expected to have a significant impact on the flow-level performance.

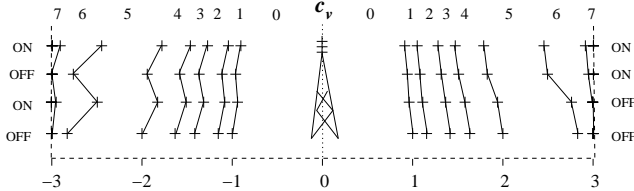


Figure 3: Impact of interference on feasible rates in a linear network ($R = 3$, $\alpha = 2$).

In hexagonal networks, due to the hexagonal shape of the cells, $r_{v,\mathcal{A}}(\theta) \leq \frac{\sqrt{3}R}{\sqrt{3 \cos \theta' + \sin \theta'}}$ where $\theta' = \theta \bmod \Pi/3$. When $\mathcal{A} = \emptyset$ the regions with identical feasible rates form

concentric rings around the BS. Figure 4 shows how these regions change with respect to \mathcal{A} , normalized with $r_{0,\emptyset}(\cdot) = 1$. As the set of active BSs, \mathcal{A} , changes from \emptyset , no interference, to $\{1, 2, 3, 4, 5, 6\}$, interference from all direct neighbors, the rings shrink considerably and feasible rates at positions away from the center change significantly. For example, a user at $r = 2.5, \theta = 0$ has a feasible rate of 307.2 kb/s when no other BS is active, and the feasible rate is halved to 153.6 kb/s when all direct neighbors are active. For a user at $r = 2.8, \theta = 0$, a similar change in \mathcal{A} reduces the feasible rate to a third, from 307.2 kb/s to 102.6 kb/s.

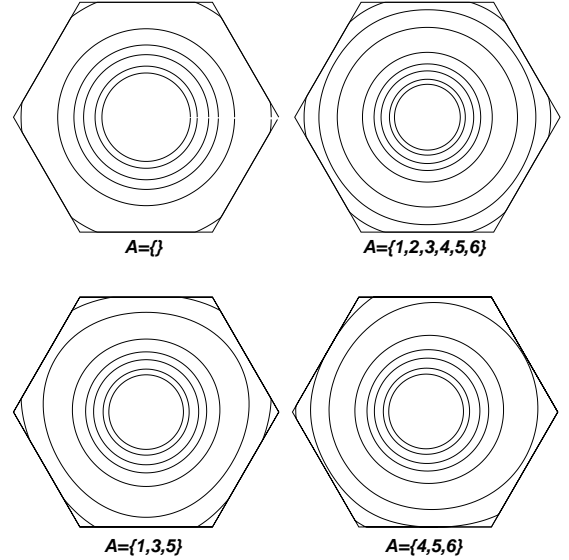


Figure 4: Impact of interference on feasible rates in a hexagonal network with omni-cells ($R = 3$, $\alpha = 2$).

The class structure described in Section 2 corresponds to subregions of the cell based on the feasible rate and variations in this rate due to changes in inter-cell interference. The regions described above are specific to a given set of active BSs. When the set of active BSs is not static, users may belong to different regions as their feasible rates change. We create user classes then, by partitioning the cell into subregions where feasible rates and their variations are identical. Two users belonging to the same class will thus have identical feasible rates for any set of active BSs.

Traffic characteristics. We assume that data flows arrive according to a Poisson process of rate λ_i in cell i , and that the arrivals are uniformly distributed throughout the cell. Under this assumption and the definition of classes above, the arrival rate of class- k data flows is $\lambda_i(k) = \lambda_i S_i(k)/S_i$, where $S_i(k)$ is the surface of the subregion k and S_i is the surface of cell i . The concept of surface corresponds to segments in linear networks and to areas in hexagonal networks.

7.2 Flow-level performance

We study flow-level performance through flow throughput as defined by (5) for some class- k flows as well as for an arbitrary flow in a given cell. We present throughput results of the first-degree bounds presented in Section 4 and the appropriate regimes of the second-degree bounds of Section 5, and compare these to simulation results. The

second-degree bounds for performance shown are the quasi-stationary regime of the lower bound and the fluid regime of the upper bound. The numerical results will show that the regimes of these second-degree bounds are indeed quite tight. Note that the lower and upper bounds for performance presented here correspond to the upper and lower bounds, respectively, for the number of active flows considered in Sections 4 and 5. The simulations are performed for exponentially distributed flow sizes. We will, however, also present a case of hyperexponential flow sizes in order to study how sensitive the throughput is to the flow size distribution.

For throughput results for an arbitrary flow, we will consider two cases of load for each network topology: homogeneous load among all cells and heterogeneous load where the load of all cells except the reference cell is always set to 0.8. Results are plotted for varying load, $\rho_{i,N}$, and it is this notion of load that we will refer to as *load* for the rest of this section. Values for the parameters η , P , and ϵ are normalized such that $r_{0,0}(\cdot) = 1$.

Linear networks. Figures 5 and 6 show throughput of an arbitrary flow in a linear network with homogeneous and heterogeneous load, respectively. A linear network was simulated as a ring of cells as opposed to a line in order to avoid the effects of cells on the edge. It has been shown in Section 6 that the system is unstable for $\rho_{i,N} \geq 1$ for homogeneous networks. We observe this in Figure 5 where the flow throughput of the actual system is zero for $\rho_{i,N} \geq 1$. In Figure 6 the throughput of the actual system is above zero past this point because the load in the neighboring cells is less than one. For both cases of load shown here, the first-degree bounds are not too rough because each cell has only two direct interferers. The second-degree bounds are quite tight and the quasi-stationary regime of the lower bounds serves as a good conservative approximation. Note that the duration of a file transfer tends to be fairly small relative to the idle or busy period of a BS, as is well captured by the quasi-stationary regime. Indeed, this has been found for simpler models in a time-varying environment [12], and apparently remains true in the more complicated model under consideration here.

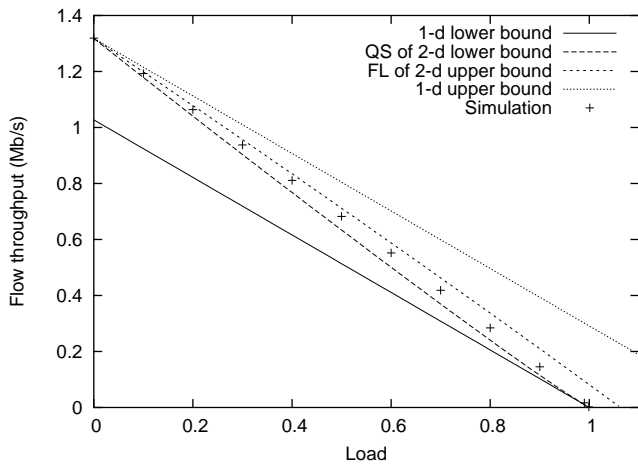


Figure 5: Throughput of an arbitrary flow in a linear network with homogeneous load ($R = 1.5$, $\alpha = 4$).

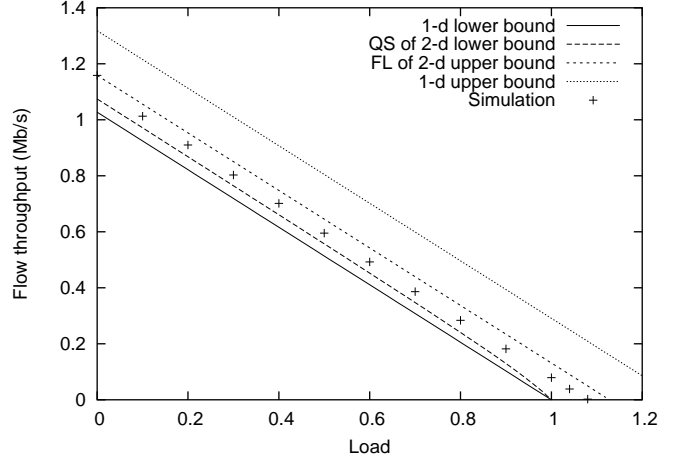


Figure 6: Throughput of an arbitrary flow in a linear network with heterogeneous load ($R = 1.5$, $\alpha = 4$).

Figure 7 shows flow throughput for five arbitrarily chosen classes in a linear network with homogeneous load. These five classes correspond, respectively, to the subregions surrounding the points at distances of 0.8, 1.0, 1.2, 1.4, and 1.5 from the BS, which has a coverage distance of $R = 1.5$. The throughput of flows is higher for users closer to the BS. The quasi-stationary regime of the lower bound is shown again to be an excellent approximation.

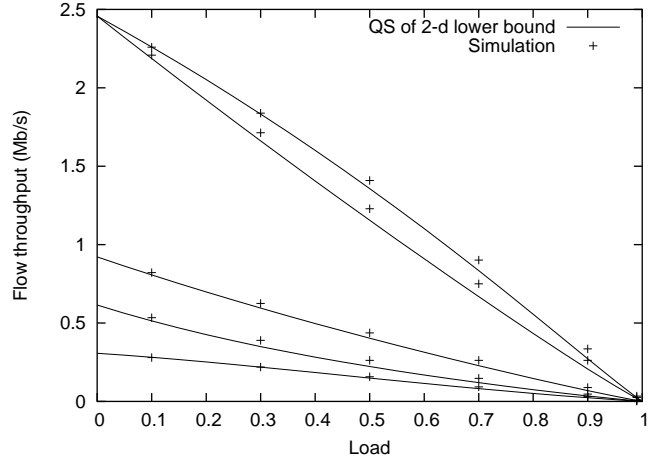


Figure 7: Per-class flow throughput in a linear network with homogeneous load ($R = 1.5$, $\alpha = 4$).

We now consider different flow size distributions in order to study the sensitivity of flow size distribution to flow throughput. In particular we consider increasingly variable flow sizes, characterized by a hyperexponential distribution defined as follows:

$$\Pr[\sigma > x] = \frac{a \exp\left(\frac{-ax}{E[\sigma]}\right) + \exp\left(\frac{-x}{aE[\sigma]}\right)}{a + 1}.$$

Such a distribution corresponds to a fraction $a/(a + 1)$ of small flows of size $E[\sigma]/a$ and a fraction $1/(a + 1)$ of large flows of size $aE[\sigma]$ and the variance in flow sizes increases with the parameter a . Figure 8 shows results of simulations

with hyperexponential flow sizes with $a = 5, 10$, and 50 , along with results of exponential flow sizes and the bounds. We observe that the system is fairly insensitive to the flow size distribution. It has been suggested in Section 6 that the stability condition is indeed sensitive to the flow size distribution. The fluid and quasi-stationary regimes of the bounds, however, are insensitive. The closeness of these bounds suggests that the performance may be relatively insensitive, as shown in the numerical results. For the rest of this paper we will continue with exponentially distributed flow sizes.

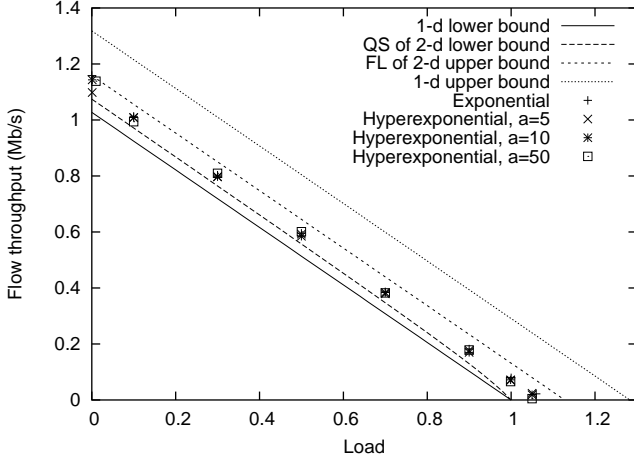


Figure 8: Throughput of an arbitrary flow with hyperexponential distribution with varying parameter a , in a linear network with heterogeneous load ($R = 1.5$, $\alpha = 4$).

Hexagonal networks: omni-cells. Figures 9 and 10 plot flow throughput for a hexagonal network of omni-cells with homogeneous and heterogeneous load, respectively. A hexagonal network is simulated with 19 cells on a plane in the pattern shown in Figure 2. In order to reduce the run time and computational complexity for the simulations, the BSs on the outer ring are set to be always active. We have performed simulations not presented here in which the BSs on the outer ring were set to be always idle, and simulations of a 37-cell network with the outer ring of BSs always active. In comparison with simulations of the 37-cell network, we find our original set-up of 19 cells with outer BSs always active to be sufficiently close and we continue with this set-up for further simulations of hexagonal networks with omni-cells. Note that we are interested in the flow performance in the reference cell at the center of such a network. The effect of the outer ring of BS is thus quite small for this reference cell, under the power-law decay assumed in (7).

The reference cell has six direct neighbors, and thus the difference between no direct neighbor being active to all such neighbors being active can be quite substantial, compared to linear networks with similar-sized cells. This effect would explain the relatively loose first-degree bounds in Figures 9 and 10. The second-degree bounds, especially the quasi-stationary regime of the lower bound are, however, remarkably close to the simulation results for $\rho_{i,\mathcal{N}} < 1$, which is a reasonable operating regime. Note that we show here re-

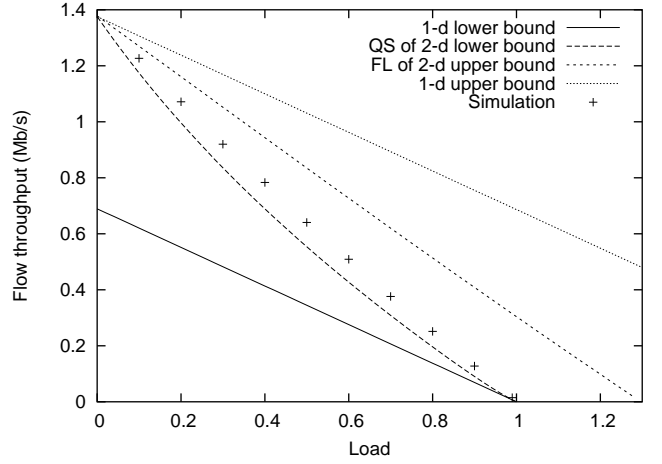


Figure 9: Throughput of an arbitrary flow in a hexagonal network with homogeneous load ($R = 1.5$, $\alpha = 4$).

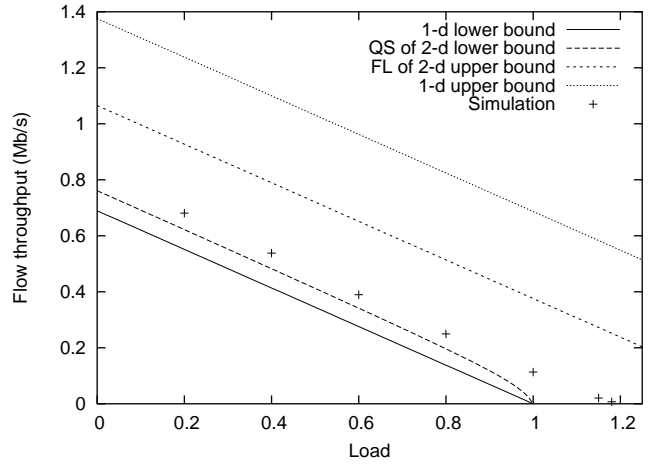


Figure 10: Throughput of an arbitrary flow in reference cell i in a hexagonal network with heterogeneous load ($R = 1.5$, $\alpha = 4$).

sults for $R = 1.5$, which may be quite small for hexagonal networks that are typically deployed in rural areas. As we increase R to be more realistic with rural scenarios, the bounds are closer because of the reduced effect of interference in larger cells. We point out, however, that for similar-sized cells, a larger number of interfering BSs leads to looser first-degree bounds.

Hexagonal networks: sectors. We have assumed above that cells have omni-directional antennas. In practice, cells are typically divided into sectors, with a directional antenna in each sector. Configurations for such network topologies differ in the number of sectors per cell (ranging from two to six) and the orientation of the directional antennas with respect to neighboring cells. We will show here an example of a symmetric network topology of cells with three 120° sectors each, oriented in a symmetric fashion, as shown in Figure 11.

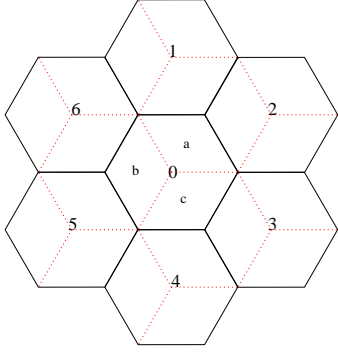


Figure 11: Hexagonal network with 120° sectors in each cell.

The power received at position u from sector i , $P_i(u)$, now includes signal attenuation as a function of the angle from the center beam of the transmitting directional antenna. The received power at position (r, θ) is now given by $P_i(r, \theta) = P\Gamma_i(r)\hat{H}_i(\theta)$ where $\hat{H}_i(\theta)$ depends on the antenna pattern. The signal attenuation due to angle separation from the center beam of the antenna can be reasonably estimated by [1]:

$$\hat{H}_i(\theta) = 10^{H_i(\theta)/10}, \quad H_i(\theta) = -\min \left\{ 12 \left(\frac{\theta'}{\theta_{i,3dB}} \right)^2, H_m \right\},$$

where θ' is the angle separation from the user and the center beam of the antenna at sector i , $\theta_{i,3dB}$ is the horizontal beam width of antenna i (typically 70° for 120° sectors), and $H_m = 20$ dB is the maximum loss.

Due to these directional antennas, a user in a given sector may experience lower interference, as compared to the network with omni-directional antennas. In the topology given in Figure 11, for a user in reference sector 0a, sectors 1c and 2b are primary interferers, while sectors 3b, 4a, 5a, and 6c are secondary interferers. Figure 12 shows the effect of the activity of these interfering sectors on the feasible rate regions in the reference sector 0a. For these results, we make an approximation by assuming that $\hat{H}_i(\theta) = 1$ if $\theta' > 60^\circ$, thus ignoring interference from other sectors of the same cell.

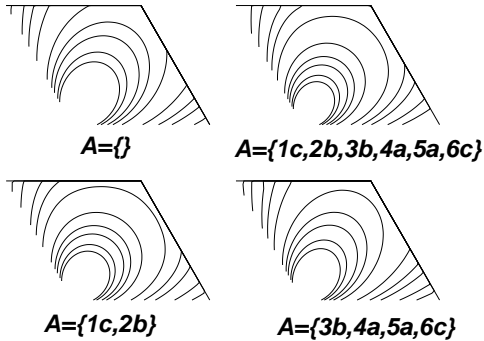


Figure 12: Impact of interference on feasible rates in a sectored hexagonal network ($R = 2$, $\alpha = 2$).

Figures 13 and 14 show throughput of an arbitrary flow in sectored hexagonal networks with homogeneous and het-

erogeneous loads, respectively. We use the topology shown in Figure 11 and assume that interference is due only to primary interferers, thus only sectors 1c and 2b interfere with transmissions in sector 0a. We observe that in comparison with hexagonal networks of omni-cells, the first-degree bounds are close, due to fewer interferers. As seen in other results presented so far, the quasi-stationary regime of the lower bound is a very good approximation.

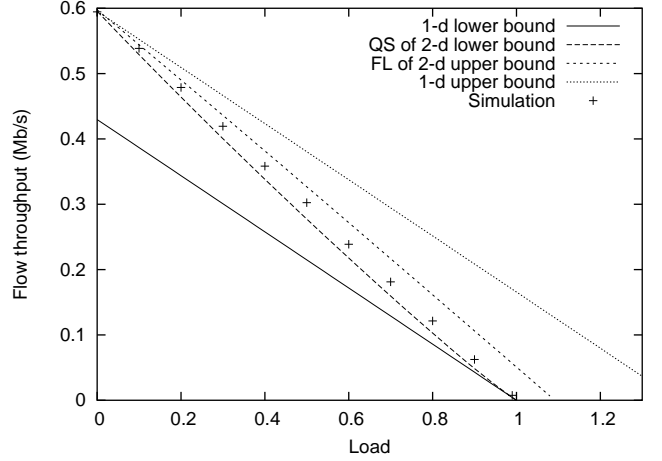


Figure 13: Throughput of an arbitrary flow in a sectored hexagonal network with homogeneous load ($R = 1.5$, $\alpha = 3$).

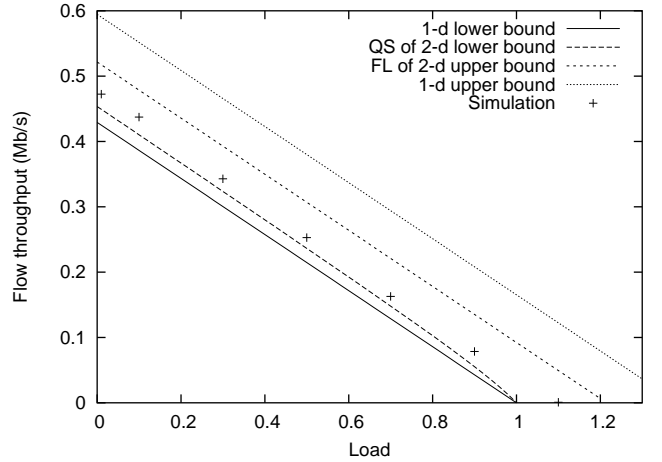


Figure 14: Throughput of an arbitrary flow in a sectored hexagonal network with heterogeneous load ($R = 1.5$, $\alpha = 3$).

8. CONCLUSION

We have modeled networks of multiple BSs, including the complex interactions of their activity periods through mutual interference. We derived first-degree bounds by assuming maximum and minimum interference in the cell under consideration. We also obtained closer bounds in the form of quasi-stationary and fluid regimes of second-degree bounds which assume maximum and minimum interference

in neighbors of the reference cell. Numerical experiments for regular network topologies showed that these bounds are indeed quite close, and further demonstrated that the quasi-stationary regime of the lower bound for performance is an extremely accurate approximation. The results highlight the importance of capturing the dynamic activity patterns of neighboring BSs. Models of a BS in isolation as considered in most previous work basically correspond to the first-degree bounds, which only provide a very rough approximation, at best.

The observation that the lower bound for performance is so tight, especially in hexagonal networks, may be explained as follows. In a system with strong interactions between the activity periods of the BSs, there is a tendency toward synchronization among the BSs. For example, when its neighboring BSs are active, a BS would serve its users at lower rates, thus causing it to stay active also, as its neighbors. Similarly, when the neighboring BSs are inactive, users enjoy higher rates, allowing flows to be served quicker, thus enabling the BS to return to an inactive state, as its neighbors. This coupling effect is stronger when there are more neighbors, as in hexagonal networks, as well as at higher loads. Recall that the lower bound assumes the neighbors of the reference cell or their neighbors to be always on. Thus, in an active cell users are served at a rate that assumes that the neighbors are also on, reflecting the above coupling of BSs. At the upper bound however, an active BS serves its users assuming the neighbors are all off, so the synchronization between the activity states of the BSs is not captured. Therefore the lower bound reflects this coalescing phenomenon better than the upper bound, which explains why the lower bound provides a more accurate approximation, especially at higher loads where the coupling effect is more pronounced.

It is worth recalling that we assumed that the BSs always transmit at full power when there are any active users. We further supposed that the transmission resources are shared in a fair manner among the active users, regardless of the activity state of neighboring BSs. In principle, the throughput may be improved by scheduling the transmissions to the various users based on the activity state of interfering BSs. For example, Figures 3, 4 and 12 indicate that edge users experience a strong increase in feasible rates when neighboring BSs are idle. This suggests that their throughputs may be increased by scheduling such users when the corresponding BSs are inactive. (Note that this form of scheduling requires not only that the activity state of neighboring BSs is known, but also that the locations of active users are known. Such information might be available in future wireless networks.) Further throughput gains may be obtained by coordinating the activity patterns of adjacent BSs and adjusting the transmit powers based on the configuration of active users. The throughput benefits from such intra-cell and inter-cell scheduling strategies in a static context have been explored in [4]. The above options provide useful instruments to enhance performance in a dynamic context as well, although wedded with potential pitfalls and challenging implementation issues. We leave these as interesting avenues for further research.

9. REFERENCES

- [1] 3GPP (2002). Spatial Channel Model Text Description. [Combined ad-hoc from 3GPP and 3GPP2.] <http://www.3gpp.org>
- [2] Agrawal, R., Bedekar, A., La, R.J., Subramanian, V. (2001). Class and channel condition based weighted proportional fair scheduler. In: *Teletraffic Engineering in the Internet Era, Proc. ITC-17*, Salvador da Bahia, eds. J.M. de Souza, N.L.S. da Fonseca, E.A. de Souza e Silva (North-Holland, Amsterdam), 553–565.
- [3] Balakrishnan, H., Padmanabhan, V.N., Seshan, S., Katz, R.H. (1997). A comparison of mechanisms for improving TCP performance over wireless links. *IEEE/ACM Transactions on Networking* **5**, 756–769.
- [4] Bedekar, A., Borst, S.C., Ramanan, K., Whiting, P.A., Yeh, E.M. (1999). Downlink scheduling in CDMA data networks. In: *Proc. Globecom '99 Conference*, 2653–2657.
- [5] Bender, P., Black, P., Grob, M., Padovani, R., Sindhushayana, N., Viterbi, A. (2000). CDMA/HDR: a bandwidth-efficient high-speed wireless data service for nomadic users. *IEEE Commun. Mag.* **38** (7), 70–77.
- [6] Bonald, T., Proutière, A. (2003). Wireless downlink channels: user performance and cell dimensioning. In: *Proc. ACM Mobicom 2003*.
- [7] Bonald, T., Borst, S.C., Proutière, A. (2004). How mobility impacts the flow-level performance of wireless data networks. In: *Proc. Infocom 2004*.
- [8] Borst, S.C. (2003). User-level performance of channel-aware scheduling algorithms in wireless data networks. In: *Proc. Infocom 2003*.
- [9] Chakravorty, R., Katti, S., Crowcroft, J., Pratt, I. (2003). Flow aggregation for enhanced TCP over wide-area wireless. In: *Proc. Infocom 2003*.
- [10] Cohen, J.W. (1984). On a functional relation in three complex variables; three coupled processors. Technical Report 359, Mathematical Institute, University of Utrecht.
- [11] Cohen, J.W., Boxma, O.J. (1983). *Boundary Value Problems in Queueing System Analysis* (North-Holland Publ. Cy., Amsterdam).
- [12] Delcoigne, F., Proutière, A., Régnié, G. (2004). Modelling integration of streaming and data traffic. *Performance Evaluation* **55** (3-4), 185–209.
- [13] Fayolle, G., Iasnogorodski, R. (1979). Two coupled processors: the reduction to a Riemann-Hilbert problem. *Z. Wahr. verw. Geb.* **47**, 325–351.
- [14] Jalali, A., Padovani, R., Pankaj, R. (2000). Data throughput of CDMA-HDR a high efficiency-high data rate personal communication wireless system. In: *Proc. IEEE VTC 2000 Spring Conf.*, 1854–1858.
- [15] Liu, X., Chong, E.K.P., Shroff, N.B. (2003). A framework for opportunistic scheduling in wireless networks. *Comp. Netw.* **41**, 451–474.
- [16] Shakkottai, S., Stolyar, A.L. (2001). Scheduling algorithms for a mixture of real-time and non-real time data in HDR. In: *Teletraffic Engineering in the Internet Era, Proc. ITC-17*, Salvador da Bahia, eds. J.M. de Souza, N.L.S. da Fonseca, E.A. de Souza e Silva (North-Holland, Amsterdam), 793–804.
- [17] Tsibonis, V., Georgiadis, L., Tassiulas, L. (2003). Exploiting wireless channel state information for throughput maximization. In: *Proc. Infocom 2003*.

[1] 3GPP (2002). Spatial Channel Model Text Description. [Combined ad-hoc from 3GPP and