Distributive flow control in multihop ad hoc communication networks

Diplomarbeit
zur Erlangung des akademischen Grades
Diplomphysiker

vorgelegt von

Ingmar Glauche

geboren in Karl-Marx-Stadt, jetzt Chemnitz, am 16. August 1977

Institut für Theoretische Physik
Fachrichtung Physik
Fakultät Mathematik und Naturwissenschaften
der Technischen Universität Dresden

Oktober 2003

1. Gutachter: Prof. Dr. Gerhard Soff (Technische Universität Dresden)

2. Gutachter: Priv. Doz. Dr. Martin Greiner (Siemens AG, München)

Die Arbeit wurde am 23. Oktober 2003 eingereicht.

In memory
To my grandfather

Abstract
Wireless multihop ad hoc communication networks represent an infrastructureless and self-organised generalisation of todays wireless cellular networks. Lacking a central control authority, the ad hoc devices (nodes) have to organise data traffic and information flow by use of distributive, local control actions such that the overall, global system performs in an optimal way. One important issue is the connectivity of a network which is a prerequisite for multihop connections between participating nodes. A simple distributed protocol is constructed by referring to continuum percolation, guaranteeing strong connectivity almost surely and independently of various spatial arrangements of ad hoc nodes. For the statistical analysis of the dynamical properties of data traffic a generic simulation model is developed that includes the relevant mechanisms of a real-world scenario. By introducing graph theoretical measures an analytical understanding is provided on a global as well as on a local, single-node scale. The overall network performance is limited by heavily used bottelneck nodes. These limitations are removed with a newly developed distributed routing algorithm based on the idea of adaptive distance vector routing. The algorithm is able to independently learn about the congestion state of the network and to adapt the routing decisions to local changes in the network load.

Zusammenfassung
Drahtlose multihop ad hoc Kommunikationsnetze stellen eine selbstorganisierende Verallgemeinerung heutiger zellulärer Netze dar. Da diese ad hoc Netze ohne zentrale Kontrolleinheiten arbeiten, ist es Aufgabe der beteiligten Geräte (Knoten) selbständig ein Netzwerk aufzubauen, das durch lokale Regeln den Austausch von Daten ermöglicht und dabei zu einem optimalen Gesamtzustand des Systems führt. Eine erste wichtige Frage ist die nach der Verknüpftheit des Netzwerkes, die eine Grundvoraussetzung für multihop Verbindungen zwischen einzelnen Knoten darstellt. In Anlehnung an die Kontinuumsperkolation wird ein lokales Protokoll vorgestellt, das Verknüpftheit weitgehend sichert und sich unabhängig auf verschiedene räumliche Anordnungen der ad hoc Knoten einstellt. Für die statistische Analyse der dynamischen Eigenschaften des Datenverkehrs auf speziellen ad hoc Netzen wird ein realitätsnahes Simulationsmodell aufgestellt. Durch die Einführung graphentheoretischer Maße können sowohl globale wie auch lokale Observable adäquat beschrieben werden. Die Gesamtleistung des Systems wird durch besonders häufig genutzte Knoten beschränkt. Durch die Einführung eines neuartigen, lokalen Algorithmus zur Steuerung des Datenverkehrs können diese Beschränkungen umgangen werden. Der Algorithmus ist in der Lage, selbständig den Lastzustand des Netzwerkes zu erlernen und darüber hinaus den Datenverkehrsfluss der lokalen Lastverteilung anzupassen.

Contents

Der Glaube an die Wahrheit beginnt mit dem Zweifel
an allen bis dahin geglaubten Wahrheiten.

Friedrich Nietzsche
"Menschliches, Allzumenschliches"

Introduction

Todays wireless communication is almost exclusively based on the concept of cellular networks [1, 2, 3] like in the existing GSM[1] or the upcoming UMTS[2] networks. A communication cell has to be understood as the spatial area in which all mobile devices are wirelessly connected to the same base station. The base stations representing their communications cells are themselves linked by a high bandwidth backbone structure. Each sending mobile device first contacts its nearest base station, which in turn provides a route either intracellularly by connecting to another device within the same communication cell or intercellularly by employing the backbone structure. In the latter case the base station in which the intended receiver is registered is contacted and provides the enquired route. The left sketch in Figure 0.1 visualises a simple cellular network. As part of the centralised backbone infrastructure each base station acts as a route manager, possesses the network information, controls the single-hop communications within its cell and assigns different channels to its various mobile clients. In order to obtain an optimal coverage for a certain area the base stations need to be intelligently placed. In almost all cases the location of the base stations remains fixed which leads to a static infrastructure, hard to change and adapt to new, revised needs. In order to provide extended communication services in remote areas or in emergency situations this concept has considerable disadvantages. This inflexibility motivates an adaptive and infrastructure-less concept that does not require a supervising control authority: self-organising wireless mobile ad hoc networks [4, 5, 6].

Since a central control authority is missing in a wireless ad hoc network sending devices use inbetween mobile devices to communicate with the intended receiver via so called multihop connections. The inbetween devices just transfer the data and so act as routers. These basic principles are shown in the right sketch of Figure 0.1. As a control authority is missing, the participating devices need coordination amongst themselves to ensure network connectivity, efficient discovery and execution of end-to-end routes and avoidance of data packet collisions on shared radio channels. A further challenge for the devices is their own mobility effecting the persistence and reliability of multihop connections. All these features are of a global nature and can only be established by a complex interaction between a multitude of different mobile devices. The idea of a wireless ad hoc communication is to base the organisation and maintenance of the whole network on local coor-

[1] Global System for Mobile communication
[2] Universal Mobile Telecommunications System

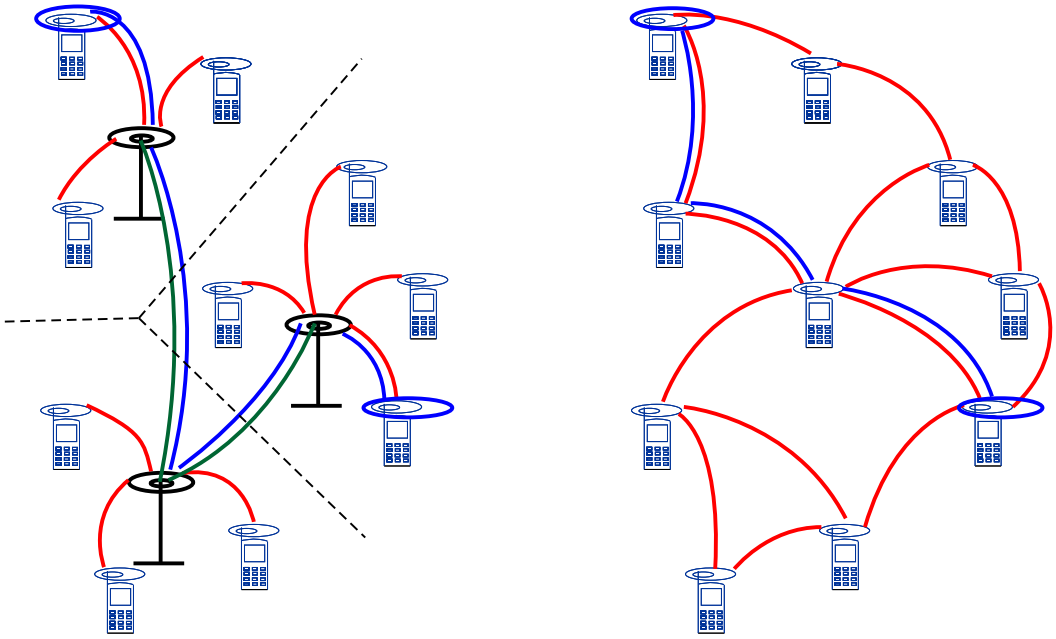

Figure 0.1: Basic layout of a cellular and a multihop ad hoc network
The left sketch illustrates a cellular network consisting of ten mobile devices (depicted as mobile phones) assigned to three communication cells identified by the base stations (black). The base stations are linked by a backbone structure (green). A communication between the two highlighted devices (marked with blue) is set up by use of the device's link to the base stations and the underlying backbone structure. In an ad hoc network, shown right, the same communication would employ a multihop route via inbetween devices (also marked in blue). In such a network no central authorities like base stations exist.

dination rules for each mobile device. The mobile device communicates directly only with its close by current spatial neighbours. Hence, it can only gain direct information about the state of its local surrounding. Since this is the complete input into the coordination rule, the latter is by definition local. Based on these coordination rules the device can for example readjust its transmission power to its new surrounding or perform special operations to access routing information.

The future of wireless ad hoc communication networks lies in areas where a fast adaption to different local surroundings without any predefined infrastructure is desired. Here these kind of networks are superior in the way that they completely self-organise a communication infrastructure and manage data exchange within the network. The idea of multihop ad hoc networks might also be employed to enhance the coverage of classical cellular wireless communication for example for

mobile phones or wireless LAN[3]. Within these so called hybrid networks only mobile devices in the immediate surrounding of a base station establish direct connections to the latter. Other more distant devices employ a multihop connection to connect to the base station. Shorter transmission ranges as they occur in multihop connections can lead to a reduction of transmission power. Besides the aspects of a sensible energy management reduced emission has less influence on the physical environment. Detailed studies about the impact of electro-magnetic radiation used for communication services on the human body ("electro-smog") still need to be prepared.

The structure and properties of communication networks other than wireless have also been heavily discussed within the last years. Intensive research was carried out especially focusing on the internet and data transfer within this network structure. Key issues have been for example phase transition like behaviour from an uncongested to a congested traffic regime [7, 8, 9, 10], self-similar data traffic [11, 12, 13] and network structure [14, 15, 16, 17]. It has to be pointed out that this work almost exclusively concentrates on communication networks based on wired connections. This implies that there are no topological restrictions to the structure of the network. In contrast wireless communication networks do not have dedicated links between participating devices. Communication between two devices is based on a shared medium. This causes a mutual interaction and influence of different communication actions and thus limits communication to spatially confined areas. These topological limitations have clear effects on the structural properties of wireless communication networks and mark a border to their wired counterparts. Additionally the mutual interactions of different communications force the system towards a far more sensible management of data transfer in order to not destruct sending and receiving events. For this reason a medium access control protocol MAC is employed. This protocol is a set of rules controlling the access of the sending device to the shared wireless medium. The protocol is also in charge of the temporary blocking of neighbouring devices that might interfere with a permitted transmission. It is just this effect that works against the tendency to assign higher transmission powers in order to reach a larger neighbourhood. Such a latter action might be in favour of more routing efficiency as more direct and longer-ranged connections are established, but medium access control then counteracts as now more neighbours are being blocked during one-hop transmissions.

The mentioned facts indicate that a wireless ad hoc communication network has to by studied as a complex system where the interplay of the constituents governs the topology as well as the dynamical behaviour of the system in a non-trivial way. Probabilistic arguments, as they are common in the concepts of modern statistical physics, are necessary in order to describe such a non-deterministic system [18].

[3]Local Area Network

The ability of a wireless ad hoc network to self-organise based on the discovery and usage of local interactions will be reviewed under this aspects.

This thesis focuses on three major issues of wireless ad hoc communication: connectivity of the networks, statistical aspects of data traffic and self-organising routing control. In Chapter 1 a precise definition of random geometric graphs in the context of wireless ad hoc communication networks is given. A closer look on various spatial point patterns, the propagation-receiver model and different rules for the assignment of transmission power is included. Chapter 2 focuses on the question how the power assignment has to be managed in order to provide a highly connected network structure. From the viewpoint of continuum percolation [19, 20] the dependence of the probability for the occurance of a highly connected network is studied and a critical range is given above which an ad hoc network graph is almost surely connected. Within the research activities for this thesis large parts of this chapter have already been published [21]. Based on the previous results Chapter 3 studies the global as well as local statistical properties of simple generic data traffic on ad hoc network graphs. By introducing graph theoretical measures an analytical understanding of a number of dynamical features can be provided. Of particular interest is the queueing behaviour of data packets at the nodes of the network [22], which is employed to explain many phenomena of the generic data traffic. In Chapter 4 the so far used approach for packet routing based on a shortest path metric is completely reorganised. A distributive routing scheme is provided that independently learns about the congestion state of the network and additionally adapts the routing decisions to local changes in the network load. Chapter 5 concludes with a summary and a brief outlook to future research.

1 The structure of wireless ad hoc communication networks

The structural analysis of complex networks has recently received much attention [14, 15, 16]. Complex networks can be found in many disciplines of science including technology [23, 24, 25], sociology [26], biology [27, 28, 29, 30] just to name a few. Much of this research was dedicated to the investigation of statistical properties in real-world networks [17, 31, 32] and their modelling by suitable algorithms representing behaviours found in the corresponding real-world systems. The approach taken for the construction of wireless ad hoc networks is slightly different. Due to the lack of an existing real-world system that could be studied the necessary construction of an appropriate communication network has to be performed in order to provide a suitable basis for a possible future implementation.

Section 1.1 introduces the conceptual notion of network graphs in general. The subsequent sections focus on particular properties of wireless ad hoc networks. To account for different spatial distributions of mobile devices in a certain area different point patterns are presented in Section 1.2. The propagation-receiver model is discussed in Section 1.3. Section 1.4 introduces different rules for the transmission power assignment among all mobile devices of a wireless network.

1.1 Random geometric graphs

The wireless communication networks are modeled as graphs $\mathcal{G} = (\mathcal{N}, \mathcal{L})$ [33]. $\mathcal{N}$ refers to the set of mobile devices of a network that from now on will be called nodes. Let $N = |\mathcal{N}|$ denote the number of nodes in $\mathcal{N}$. The set of links that connect the nodes of the network is represented by $\mathcal{L}$. In general these are directed links $i \to j$. In that sense a node i can influence node j. For a vice versa dependence a link $j \to i$ is required. If both links $i \to j$ and $j \to i$ exist for two nodes $i, j \in \mathcal{N}$ the link is called a bidirectional link $i \leftrightarrow j$. The subset $\mathcal{L}^{\text{bidir}} \subset \mathcal{L}$ represents the complete set of all bidirectional links in the graph $\mathcal{G}$. Bidirectional links proved to play a crucial role in wireless communication. The subset $\mathcal{N}_i \subset \mathcal{N}$ is called the communication neighbourhood of node i and represents the complete set of nodes $j \in \mathcal{N}_i$ that all have bidirectional links $j \leftrightarrow i$ in $\mathcal{L}^{\text{bidir}}$ with node i.

The existence of links in $\mathcal{L}$ is a direct result of the location of the nodes in $\mathcal{N}$, their individual transmission power assignment and the propagation-receiver model. Because of the importance of the location of the nodes for the link construction the resulting network graphs are denoted as random geometric graphs

[34]. The particular topics are addressed in detail in the following sections.

A communication route or path is a sequence of nodes such that there are bidirectional links in $\mathcal{L}^{\text{bidir}}$ between all consecutive pairs of nodes. A shortest path $\sigma_{i,f}$ between two nodes $i, f \in \mathcal{N}$ is a route containing fewest possible number of nodes; the number of links of a shortest path $\sigma_{i,f}$ is called the distance d_{if}. The average over all distances $D = \langle d_{if} \rangle$ is referred to as the diameter of a network.

A scaling of the diameter $D \sim \sqrt{N}$ with the network size can be expected. Mapping a network of size N on a lattice, where the nodes are connected by the grid lines, it is intuitive that the average distance of the nodes grows with the lattice side length $\sqrt{N}$. This finding was verified for the network models studied in this thesis [35].

A convenient way to describe the attribution of the links in $\mathcal{L}$ to the set of nodes represented by $\mathcal{N}$ is the so called $N \times N$ adjacency matrix [30]:

$$a_{ij} = \begin{cases} 1 & \text{link } i \to j \text{ exists} \\ 0 & \text{link } i \to j \text{ does not exist or } i = j \end{cases} \tag{1.1}$$

Within this notation bidirectional links are identified by $a_{ij} = a_{ji} = 1$. An adjacency matrix also representing unidirectional links (meaning $a_{ij} \neq a_{ji} = 1$ or vice versa for at least one pair $i, j \in \mathcal{N}$) can be transformed into a corresponding adjacency matrix only representing the existing bidirectional links by a pointwise multiplication $a_{ij}^{\text{bidir}} = a_{ij} a_{ji}$.

The node degree k_i of node i is the number of nodes contained in $\mathcal{N}_i$. It simply counts the number of bidirectional neighbours of i. An alternative definition is possible by use of the bidirectional version of the adjacency matrix: $k_i = \sum_{j \in \mathcal{N}} a_{ij}^{\text{bidir}}$. In an analog manner we can define the outgoing node degree $k_i^{\text{out}} = \sum_{j \in \mathcal{N}} a_{ij}$ and the ingoing node degree $k_i^{\text{in}} = \sum_{j \in \mathcal{N}} a_{ji}$. The outgoing node degree counts all nodes $j \in \mathcal{N}$ if there is a link $(i \to j) \in \mathcal{L}$. The corresponding outgoing neighbourhood of i is denoted $\mathcal{N}_i^{\text{out}}$. In similar fashion $\mathcal{N}_i^{\text{in}}$ is defined.

In good approximation to most of the future applications of wireless ad hoc communication a modelling in two spatial dimensions is reasonable. Thus a node is represented by its spatial coordinates (x, y). In principle the model should be based on spatio-temporal point patterns, where elements of motion do enter [36]. Although mobility modelling is crucial once the velocity of the nodes is large compared to their mutual distances, for reasons of simplicity a quasi-static picture is chosen. An ensemble of successive static point patterns can be considered as a series of snapshots of a slowly changing spatio-temporal environment.

Without lack of generality all nodes $i \in \mathcal{N}$ are assigned positions in the unit square $(x, y)_i \in [0, 1] \times [0, 1]$. A more detailed description of the assignment of the positions for different point patterns is given in the following section.

1.2 Spatial point patterns

Chapter 2 of this thesis asks how the transmission power assignment of the nodes in $\mathcal{N}$ has to be managed in order to provide a highly connected network structure. In order to learn about the dependence of connectivity on the spatial distribution of the set of nodes $\mathcal{N}$ three different generic random spatial point patterns are used: a homogeneous, a multifractal and a Manhattan point pattern. The point patterns are inspired by possible simplified real-world scenarios. The generation of these different patterns is briefly explained in the following lines.

Homogeneous point pattern In a random homogeneous point pattern each of the N nodes of $\mathcal{N}$ is given a random position $(x, y) \in [0, 1] \times [0, 1]$. A typical realisation is illustrated in Figure 1.1. By definition it does not show generic clustering.

Multifractal point pattern One way to construct simple clustered point patterns is to employ a binary multiplicative branching process. The non-uniform probability measure supported on the unit square is constructed by iteration: at first the parent square is divided into four offspring squares with area 1/4. Two randomly chosen offsprings get a fraction $(1+\omega)/4$ of the parent probability mass $m = 1$, whereas the remaining two get a fraction $(1-\omega)/4$. In the next iteration step each offspring square follows the same probabilistic branching rule and non-uniformly redistributes its probability mass onto its own four offsprings. After j iteration steps the probability mass $m = 1$ has been non-uniformly subdivided onto 4^j subsquares with area $1/4^j$, where $\binom{j}{i}2^j$ of these subsquares $(0 \leq i \leq j)$ come with probability mass $[(1+\omega)/4]^i[(1-\omega)/4]^{j-i}$. One after the other each of the N points to be distributed in $\mathcal{N}$ is given an independent and uniform random number between 0 and 1, which, given some probability-mass-weighted ordering of the 4^j subsquares, corresponds to exactly one subsquare, onto which the particle is deposited and randomly placed inside. One such realisation of a point pattern is shown in Figure 1.1. The observed hierarchical clustering of nodes is due to the hierarchical branching structure of the iteration process. The probability measure constructed with such a multiplicative branching process is a multifractal [37]. The construction of multifractal fields has some importance in such diverse fields as turbulence [38, 39], finance [40, 41], Internet traffic [42], high-energetic multiparticle dynamics [43] and deterministic chaos [44], just to name a few.

Manhattan point pattern As a third generic class of spatial point patterns a Manhattan street pattern is used. N_x and N_y streets are equidistantly placed parallel to the x- and y-axis, respectively. One after the other nodes of $\mathcal{N}$ are randomly placed onto one randomly chosen street. Figure 1.1 gives an illustration

of one realisation.

1.3 Radio propagation-receiver model

The existence of a communication link between two nodes is determined by the transmission power of the nodes, the characteristics of the radio propagation medium as well as the receiving efficiency. According to a simplified propagation-receiver model a node i can listen to a transmitting node j, if relative to a NOISE the power received at i is larger than the signal-to-noise ratio SNR:

$$\frac{P_j / R_{ji}^{\alpha}}{\text{NOISE}} \geq \text{SNR} \,, \tag{1.2}$$

P_j denotes the transmission power of node j and R_{ji} represents the Euclidean distance between j and i. The path-loss exponent α is assumed to be constant; depending on the characteristics of propagation its value typically falls into the regime $2 \leq \alpha \leq 6$. A specific fixation of the value of α is not required for the generic view within this thesis. Without any loss of generality the variables NOISE and SNR are set equal to one, implying a rescaling of the transmission power P_j.

The condition (1.2) guarantees that node i is able to hear node j. This alone would define a unidirectional link $j \to i$. If (1.2) also holds for a communication $i \to j$ then the nodes i and j are connected by a bidirectional link $i \leftrightarrow j$. Although not strictly required, bidirectional links are preferred for the operation of wireless ad hoc networks because many communication protocols require instant feedback. Furthermore, the absence of unidirectional links reduces interference effects. Throughout this thesis a communication link as well as the terms neighbour and neighbourhood of a node will always refer to bidirectional connections. Otherwise it will be explicitly stated.

The link construction can be given a simple geometric interpretation as visualised in Figure 1.2. For a given transmission power P_j condition (1.2) translates into a maximum range $R_{ji} = P_j^{1/\alpha}$. All nodes $i \in \mathcal{N}$ that lie inside this circle with radius R_{ji} around node j are able to hear this node. For a communication link to exist between nodes j and i, i has to lie inside j's circle with radius R_{ji} and j has to lie inside i's circle with radius R_{ij}. For the case that all $R_{ji} = R$ are identical, this link construction matches the standard link construction of continuum percolation [45].

The values for the transmission power P_i are in general normalised by P^{norm}. The normalisation $P^{\text{norm}} = (R^{\text{norm}})^{\alpha}$ comes by setting $\pi(R^{\text{norm}})^2 = 1/N$. The 1 in the numerator comes from the area of the unit square. Thus $1/N$ can be interpreted as an inverse density of nodes. $\pi(R^{\text{norm}})^2$ has to be understood as the area in which on average only one node is placed in a realisation of a random homogeneous point pattern; P^{norm} is the transmission power to cover exactly this area.

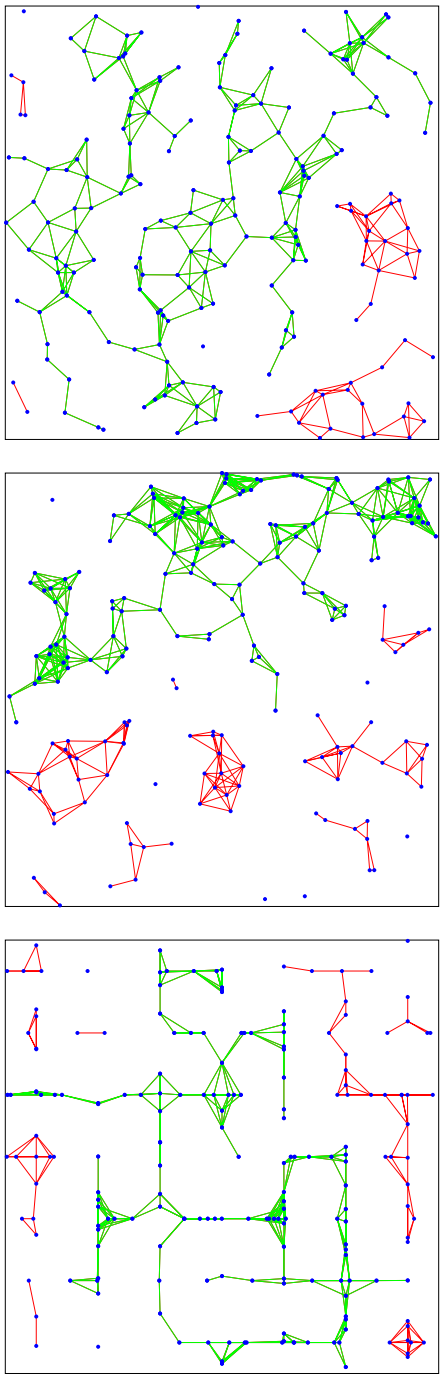

Figure 1.1: Different spatial point patterns
The upper representation corresponds to the random homogeneous point pattern, the middle to a multifractal point pattern with $j = 5$ branchings and weight $\omega = 0.4$, and the lower to a Manhattan pattern with $N_x = N_y = 7$. All graphs are shown for a set of 200 nodes. Each point is given a transmission power to reach on average 5 neighbours given a homogeneous distribution of nodes. The largest connected cluster is marked with green links.

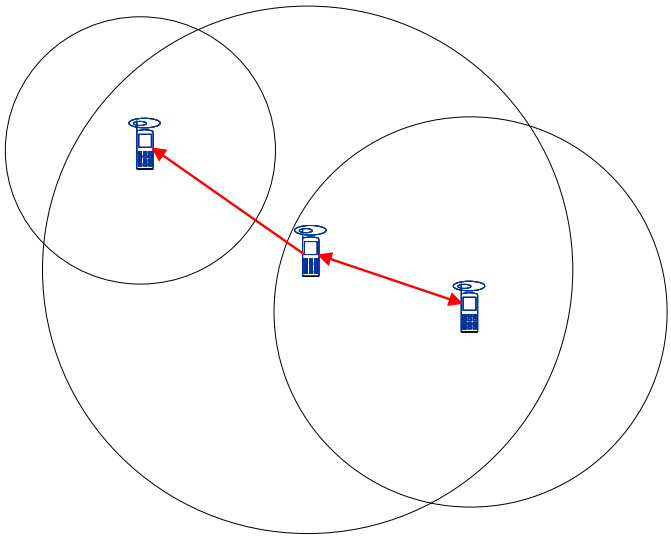

Figure 1.2: Geometric interpretation of the link construction
The left and the right node are both inside the transmission range of the central node and can listen to its transmissions. Additionally the central node is within the transmission range of the right node but not within the range of the left one. Thus a bidirectional link only exists between the central and the right node. Between the left and the central node only a unidirectional link exists. The left node does not hear the right node and vice versa. Thus no direct link between them exists at all.

1.4 Power assignment

In general it can be distinguished between two kinds of transmission power assignment. Employing global power assignment each node $i \in \mathcal{N}$ is assigned its individual transmission power P_i according to requirements based on global observables. The specific transmission power of a node might be chosen according to an independently and identically distributed (iid) random variable [46, 47] or governed by a global optimisation function for the whole network [35]. In contrast the local power assignment leaves it to a node to adjust its own transmission power based on local observables the node itself can acquire. The local rules should be set up in a way that an optimal network state for all nodes $\mathcal{N}$ is reached. A couple of different power assignments have been employed and will be introduced in detail in the following lines.

Constant transmission power In this simple global rule the same transmission power $P_i = P$ is assigned to all nodes $i \in \mathcal{N}$. All existing links are then bidirectional. This is the most widely studied rule for power assignment in wireless

multihop communication networks [46]. It will be used as reference throughout the thesis and is referred to as constant-P rule. The question how the specific choice for the constant transmission power is made will be addressed in Section 2.2.

iid transmission power Due to the homogeneous power assignment of the constant-P rule all nodes have an identical transmission range. In comparison to the small world networks [48] where a couple of long-range links lead to a better connectivity and a smaller diameter D a modification of the constant-P rule including a number of long-range links seems worth studying. These long-range links have to be represented by nodes with higher transmission power. The simplest heterogeneous rule in this context is the iid-P link rule [47]. It treats the transmission power values assigned to the nodes in $\mathcal{N}$ as independently and identically distributed (iid) random variables. Independently from the other nodes each node chooses its transmission power according to the same global probability distribution $p(P)$ with mean $\langle P \rangle$.

As a flexible representative a bimodal distribution

$$p(P) = \frac{\beta_2}{\beta_1 + \beta_2}\delta\Big(P - \langle P\rangle(1-\beta_1)\Big) + \frac{\beta_1}{\beta_1 + \beta_2}\delta\Big(P - \langle P\rangle(1+\beta_2)\Big) \tag{1.3}$$

is chosen. It comes with two parameters β_1 and β_2 determining the variance and skewness of this distribution, but leaving the mean $\langle P \rangle$ untouched. The previously used constant-P rule is reproduced once $\beta_1 = 0$ or $\beta_2 = 0$.

Minimum node degree The widely used constant-P assignment has notorious drawbacks. First of all, it is a global rule, where nodes need to access information about the complete network. Second, the used constant transmission power strongly depends on the nature of the underlying spatial point patterns [21]. It is obvious that a more flexible power assignment is desired that only collects local information and acts in a decentralised manner. The minimum node degree rule introduced in [21] is a first and promising approach in this direction.

By exchanging so-called hello and hello-reply messages each node $i \in \mathcal{N}$ is able to access direct information only from its immediate neighbours $\mathcal{N}_i$. A simple local observable for a node is the number of its communication neighbours, referred to as node degree k_i. Based on this observable alone, the simple minimum node degree strategy for a node i is to adjust its transmission power to have at least $k_i^{\min}$ bidirectional neighbours in $\mathcal{N}_i$.

Upon setting up the communication links to the other nodes, a node i attaches to its hello message information about its current neighbourhood $\mathcal{N}_i$ and its current transmission power P_i. Starting with $P_i = P^{\min}$, the node increases its transmission power by a small amount once it has not reached a minimum node

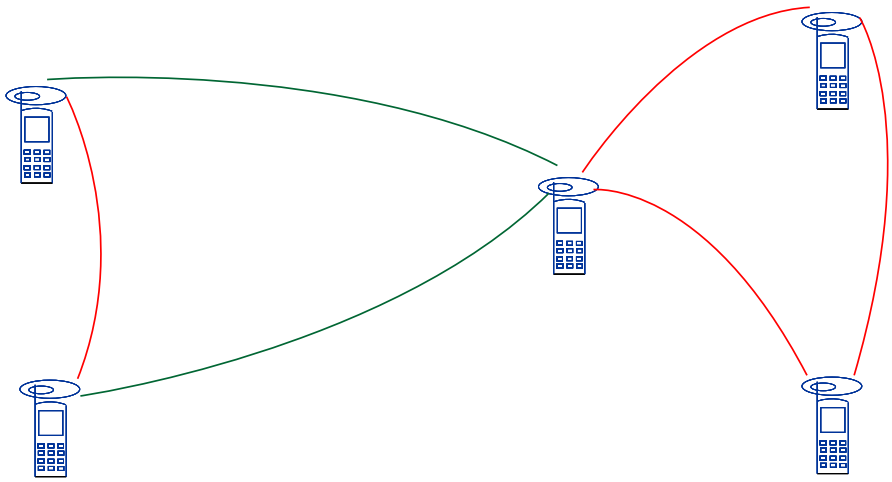

Figure 1.3: Minimum node degree rule
As an example suppose the required minimum degree to be $k_i^{\min} = 2$. The three nodes on the right fulfil this requirement simply by use of the red links. If the left nodes only have a rather small transmission power they can only hear the other left partner - marked by their red link. But they have not yet reached $k_i^{\min} = 2$. According to the minimum node degree rule they start increasing their transmission powers in order to reach more communication neighbours. At a certain point the central node can hear their hello-messages and does itself increase its transmission power such that the node on the left can hear it, too. Now these nodes have established additional bidirectional links, marked with green. The two left nodes now fulfil the minimum degree requirement $k_i^{\min} = 2$. In contrast the central node now has four neighbours. This is more than originally required. The increased transmission power was necessary to serve the needs of the more isolated nodes on the left side.

degree $k_i^{\min}$. Whenever another node j, which so far does not belong to the set of neighbours $\mathcal{N}_i$, hears the hello message of the original node i for the first time, it realises that the latter has too few neighbours, either sets its transmission power equal to the transmission power of the hello-sending node i or leaves it as before, whichever is larger, and answers the hello message. Now the original and new node are able to communicate back and forth and have established a new communication link $i \leftrightarrow j$. The original node i adds the new node j to its neighbourhood $\mathcal{N}_i$ as does j the other way around. Only once the required minimum node degree $k_i^{\min}$ is reached, the original node i stops increasing its transmission power for its hello transmissions. While some of the nodes will end up having exactly $k_i^{\min}$ neighbours, some others are forced to have more than $k_i^{\min}$ neighbours in order to give an additional bidirectional link to a node which so far has not accumu-

lated enough neighbours. The transmission power of these nodes is larger than necessary to obtain only $k_i^{\min}$ neighbours for themselves. As a simple example see Figure 1.3.

The enforcement of already satisfied nodes to reply to hello sending nodes is a crucial part of the minimum node degree rule. It guarantees that nodes do not get trapped in frustrated states. In the absence of an enforcement to reply to a node not having the required number of $k^{\min}$ neighbours, this node would continue to increase its transmission power as long as it is ignored by the already saturated nodes. This could lead to unnecessary high transmission powers that interfere with a multitude of possible communications.

A detailed insight in the algorithmic implementation of the minimum degree rule is presented in [21].

Whatever approach is chosen, according to (1.2) the actual power assignment to each of the nodes in $\mathcal{N}$ precisely defines the set of links $\mathcal{L}$. Here the special properties of a wireless communication network become apparent: the set of links $\mathcal{L}$ is a direct result of the spatial position $(x, y)_i$ and the transmission power P_i of each node $i \in \mathcal{N}$. In that sense not every possible set of links $\mathcal{L}$ can be realised by an appropriate power assignment. This is in clear contrast to any wired network where such restrictions do not influence the existence of connections between different nodes.

2 Continuum percolation approach to network connectivity

For the operation of a multihop ad hoc communication network it is desired that as many as possible nodes within a certain area can communicate with each other. The ability of two nodes to communicate is mapped onto the question if communication routes in the sense of Section 1.1 exist. This issue is referred to as connectivity of a network [49, 50, 51, 21]. This chapter focuses on the question how the transmission power of the nodes within the three introduced power assignment rules from Section 1.4 has to be adjusted in order to provide a highly connected network structure. As already pointed out it is the special property of the MAC protocol to counteract higher transmission powers by an increased number of blocked nodes during one communication. It is in that sense desired to achieve connectivity of a network by possibly low or at least well distributed transmission power.

For a fixed set of nodes $\mathcal{N}$ the set of links $\mathcal{L}$ directly depends on the transmission power of each node. In that context the probability for the connectivity of a network depends on the transmission power assignment of the nodes in $\mathcal{N}$. According to equation (1.2) the transmission power P_i can be transfered into transmission ranges R_{ij} that allow a geometric interpretation of the link construction as visualised in Figure 1.2. Similar issues are approached by continuum percolation theory [19]. Here one tries to determine a critical range of transmission such that the resulting links form a highly connected graph. Such a critical transmission range can be defined for ad hoc network graphs as well. Furthermore it can be translated into a critical node degree k^{crit}. The node degree k_i itself is a local observable of each node $i \in \mathcal{N}$. This finding gives again reason for the construction of the minimum node degree rule in Section 1.4.

Within Section 2.1 of this Chapter suitable observables for the connectivity will be introduced. Section 2.2 investigates the connectivity issue for the different power assignments introduced in Section 1.4. Section 2.3 addresses the issues of the transmission power and node degree distributions for different power assignments and different point patterns. Section 2.4 concludes with a brief summary.

2.1 Observables

For the investigation on the connectivity of a network graph $\mathcal{G}$ two observables are introduced.

Giant component The giant component of a network is given as the set of nodes $\mathcal{N}^{\text{giant}} \in \mathcal{N}$ being part of the largest bidirectionally connected cluster appearing in a graph realisation $\mathcal{G} = (\mathcal{N}, \mathcal{L})$. Normalising the number of nodes in $\mathcal{N}^{\text{giant}}$ to the total number of nodes N the relative giant component $G \in [0, 1]$ can be defined [33, 34]. A simple flooding algorithm is used to determine the giant component of a graph realisation: a random node i is tagged in first place, then its neighbours $j \in \mathcal{N}_i$ are tagged, which then continue to tag their untagged neighbours, and so on, until the corresponding cluster is saturated. This procedure is repeated for all untagged nodes, until all nodes of the graph are tagged. By definition, the nodes in the largest found cluster are referred to as the giant component $\mathcal{N}^{\text{giant}}$. See also Figure 1.1 for reference.

One-connectivity One-connectivity is defined as the probability that all nodes $i \in \mathcal{N}$ can communicate with each other. In other words it is the probability to find a network realisation $\mathcal{G}$ with $G = 1$ which is equivalent to $\mathcal{N}^{\text{giant}} = \mathcal{N}$.

One-connectivity has to be seen as a specific version of the k-connectivity. A graph $\mathcal{G}$ is called k-connected if between every pair of nodes there exist at least k independent paths, which implies, that once $k - 1$ nodes are removed at random, the graph remains at least one-connected. A k-connected ad hoc communication network is more flexible and robust to routing failure. In a numerical implementation (k>1)-connectivity requires to check for each of the $N(N-1)/2$ pair of nodes belonging to one graph whether at least k independent paths exist. This is algorithmically a very extensive procedure, that suggest a switch from costly k- to cheap pseudo-k-connectivity. Given one-connectivity, the latter only requires each node i to have at least k neighbours in $\mathcal{N}_i$. In fact, a theorem exists [52], which, translated into the language of ad hoc network graphs, guarantees for a geometric graph ensemble based on random homogeneous point patterns and the constant-P rule of Section 1.4 that in the limit of large N the probabilities for k- and pseudo-k-connectivity converge as they approach one.

2.2 Continuum percolation for different power assignments

2.2.1 Constant transmission power

The average relative giant component $\langle G \rangle$ as a function of $\langle P \rangle$ obtained from a sample of 500 geometric graph realisations of the random homogeneous spatial point patterns and employing the constant-P rule is shown in Figure 2.1. The brackets $\langle ... \rangle$ refer to the sample average over the graph realisations. For the constant-P rule the average transmission power sampled over the nodes of a graph realisation $\langle P \rangle_{\mathcal{N}}$ equals the transmission power P_i of an individual node i since

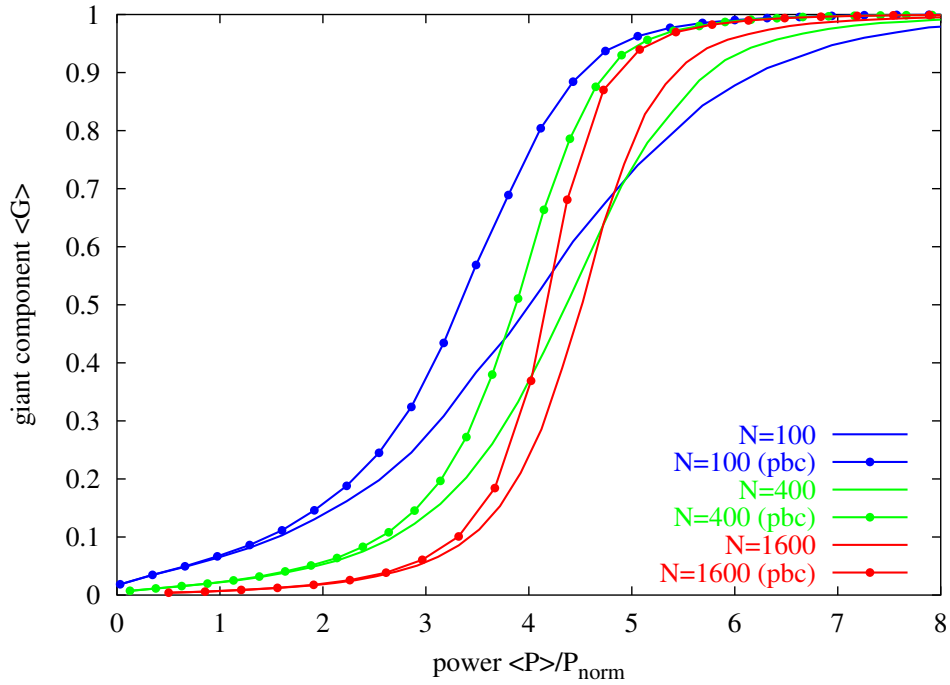

Figure 2.1: Giant component for networks of different size
Average relative giant component $\langle G \rangle$ as a function of transmission power $\langle P \rangle / P^{\text{norm}}$. A sample of 500 geometric graphs generated with random homogeneous spatial point patterns and the constant-P rule has been used. The path-loss exponent of equation (1.2) has been set to $\alpha = 2$. Different curves correspond to different number of nodes: $N = 100$ (blue), 400 (green), 1600 (red); curves marked without/with circles correspond to exclusion/inclusion of periodic boundary conditions (pbc).

for all $i \in \mathcal{N}$ it holds that $P_i = P = \langle P \rangle_{\mathcal{N}}$. The chosen rather small sample size produces already more than sufficient statistical convergence and keeps statistical error bars to such small values, that those will not be shown for this and the following figures. For a path-loss exponent set equal to $\alpha = 2$ a percolation threshold behaviour around $P^{\text{crit}} \approx 4.5 P^{\text{norm}}$ is observed: for $P \ll P^{\text{crit}}$ the average relative giant component is close to zero, whereas for $P \gg P^{\text{crit}}$ it is almost equal to one. The sharpness of the threshold depends on the number of nodes in $\mathcal{N}$; with increasing N the transition becomes sharper.

For the determination of the exact threshold position the limit $N \to \infty$ would be needed. To partially account for this, the relative giant component has also been determined from small and medium N simulations with periodic boundary conditions. Results are also shown in Figure 2.1. The percolation transition becomes

sharper and moves a little to the left once compared with the previous results, but still no full convergence for the employed increasing N values is obtained. Since for realistic ad hoc networks the limit $N \to \infty$ is unrealistic, this matter is not pursued further. Within this chapter the focus is on network realisations with $N = 1600$ nodes. For wireless ad hoc networks this is already a high bound.

The critical power $P^{\mathrm{crit}} \approx 4.5 P^{\mathrm{norm}}$, obtained by setting the path-loss exponent equal to $\alpha = 2$, can be given a more illustrative interpretation: the factor 4.5 reflects the average number of neighbouring nodes, which is also denoted as the average node degree $\langle k \rangle$. In the constant-P case the transmission ranges of all nodes of $\mathcal{N}$ are equal. The nodes contained in the circular disc around node i are thus neighbours of i. One comes up with $\langle k \rangle = (\pi (R^{\mathrm{crit}})^2) N$, where N represents the density of nodes in the unit square. In analogy to $P^{\mathrm{crit}} \approx 4.5 P^{\mathrm{norm}}$ it follows that $(R^{\mathrm{crit}})^2 \approx 4.5 (R^{\mathrm{norm}})^2$. This allows to formulate the average node degree in terms of the normalised radius $\langle k \rangle \approx 4.5 \pi (R^{\mathrm{norm}})^2 N = 4.5$ employing the normalisation condition given in Section 1.4. Note, that due to finite-size effects and the usage of no periodic boundary conditions the actually sampled critical node degree is a little smaller than the asymptotic value $k^{\mathrm{crit}} = 4.53$, which is stated for example in [34, 53].

This short derivation demonstrates that other than in terms of P^{crit} it is also convenient to characterise the percolation phase transition in terms of R^{crit} or k^{crit}. Once choosing other path-loss exponents $\alpha \neq 2$, the former will change according to $P^{\mathrm{crit}}(\alpha \neq 2) = (R^{\mathrm{crit}})^{\alpha - 2} P^{\mathrm{crit}}(\alpha = 2)$, whereas R^{crit} or k^{crit} remain as before. Within this chapter the transmission power of a node is given in units of P^{norm} because of the nice correspondence with the average node degree $\langle k \rangle$ given a path-loss exponent $\alpha = 2$. For the following chapters of this thesis it will be switched to a characterisation of a node's transmission power via the expected number of neighbours k^{constP}. For any α the condition $k_i^{\mathrm{constP}} = k^{\mathrm{constP}} = \pi (R_{ij})^2 N = \pi (P_i)^{2/\alpha} N$ fixes the transmission power of node i such that on average k^{constP} neighbouring nodes can be reached via direct bidirectional links from i.

Next, spatial point patterns other than random homogeneous are discussed. Figure 2.2 compares the relative average giant component $\langle G \rangle$ as a function of transmission power obtained for random multifractal and Manhattan spatial point patterns with the random homogeneous case; consult again Figure 1.1. A sample of 500 geometric graphs generated with random homogeneous spatial point patterns of 1600 nodes has been used. Evidently for small $\langle P \rangle$ the relative giant component is larger for the multifractal than for the homogeneous patterns, but for the convergence of the relative giant component towards one a much larger $\langle P \rangle$ is needed. Due to the hierarchical clustering, subclusters of points are easily formed at small $\langle P \rangle$ since only a small transmission range is needed to connect the corresponding nodes. However, in order to connect the various subclusters either directly to each other or via isolated nodes lying inbetween it needs a rather large transmission power. Another consequence of the pronounced clustering is that the

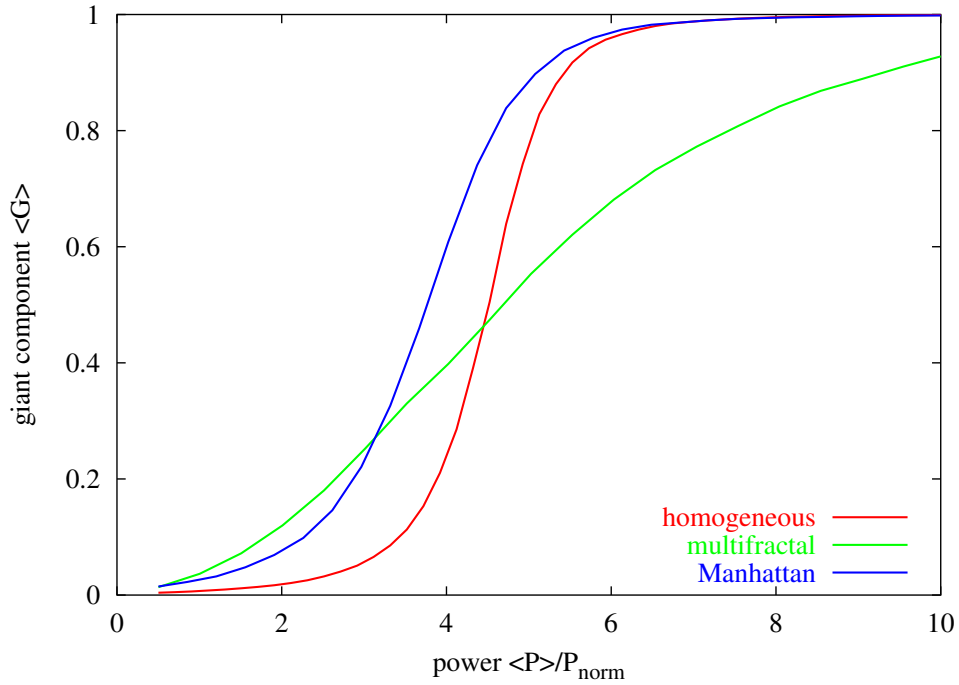

Figure 2.2: Giant component for different point patterns

Average relative giant component $\langle G \rangle$ as a function of transmission power $\langle P \rangle / P^{\text{norm}}$ upon using the constant-P rule. The number of nodes has been fixed to $N = 1600$ and the path-loss exponent has been set to $\alpha = 2$. Different curves correspond to different random spatial point patterns: homogeneous (red), multifractal (green) with parameters $\omega = 0.4$ and $j = 5$, and Manhattan (blue) with parameters $N_x = N_y = 7$. A sample of 500 geometric graphs has been used for each case.

rather sharp percolation threshold observed for homogeneous point patterns blurs more the larger the splitting parameter ω characterising the multifractal point patterns is chosen.

Also for the Manhattan point patterns the relative giant component increases faster for small transmission powers than for the homogeneous point patterns. The reason is that the average nearest-point distance is smaller for the random points confined to the one-dimensional Manhattan streets. The Manhattan threshold for the relative giant component is relatively sharp and comes at a transmission power value, which is somewhat smaller than for random homogeneous point patterns.

It is instructive to map the continuous percolation based on Manhattan point patterns onto the well-known square-lattice bond percolation [19]. Two nodes are assumed to lie on a one-dimensional straight line and to have a distance

$l = 1/N_x = 1/N_y$ corresponding to the distance of two successive Manhattan street crossings. They can only communicate with each other, if inbetween nodes i come with successive distances r smaller than their transmission range $R = R_{ij}$; otherwise the occurring void is too large to be bridged. For a one-dimensional Poissonian point pattern with density $\lambda = N/(N_x + N_y)$ the distance r of two consecutive nodes is exponentially distributed according to $p(r) = \lambda \exp(-\lambda r)$. This allows to estimate the probability for the occurrence of at least one too large void between the two picked nodes with distance l. The expression is

$$p(\text{void} > R) = \sum_{j=1}^{m} \frac{(-\lambda)^{j-1}}{(j-1)!}(l - jR)^{j-1} \left[1 + \frac{\lambda}{j}(l - jR)\right] e^{-j\lambda R}\,, \tag{2.1}$$

where $m = \lfloor l/R \rfloor$ denotes the largest integer smaller than or equal to l/R. A derivation of this formula is given for example in [51]. Setting the path-loss exponent to $\alpha = 2$ and according to equation (1.2) converting the threshold power $P \approx 4P^{\text{norm}}$ into R, one arrives at a value $p(\text{void}{>}R) = 0.47$, which almost matches the critical bond probability $p_{\text{bond}} = 0.50$ of bond percolation on a square lattice [19].

In addition to the average relative giant component the curve for the one-connectivity is shown in Figure 2.3. The quantities are given as a functions of $\langle P \rangle$ obtained from a sample of 500 geometric graphs generated with random homogeneous spatial point patterns of 1600 nodes and the constant-P rule. The one-connectivity also reveals a threshold behaviour, which sets in once the relative giant component has approached one. It is intuitive that at this point the probability for $\mathcal{N}^{\text{giant}} = \mathcal{N}$ starts to deviate from zero. The relative factor of the respective threshold positions is about 2.3 compared to the average relative giant component.

A brief comment is in order here. The percolation threshold for the giant component does not depend on the size of a network since it is defined as the point where the probability to find an infinite cluster within an infinite quantity of nodes $\mathcal{N}^{\infty}$ starts to deviate from zero. It does not imply that all nodes in $\mathcal{N}^{\infty}$ have to be part of that giant component. But this is what one-connectivity requires. As shown in [50] this quantity explicitly depends on the size of the network, in particular it scales with $(\log N)$. Thus a threshold in the above mentioned context does only make sense for a given network size N. As already stated the investigations presented here are limited to networks of size up to 1600 nodes which is a reasonable bound for ad hoc communication networks.

For completeness the probabilities for pseudo-k connectivity with $k = 2$ and 3 are also depicted in Figure 2.3. Of course, their threshold is shifted to even larger P when compared to the pseudo-one threshold. A part of Figure 2.6 shows the probability for one-connectivity for random multifractal and Manhattan point patterns. Whereas the Manhattan curve is close to the homogeneous curve, the

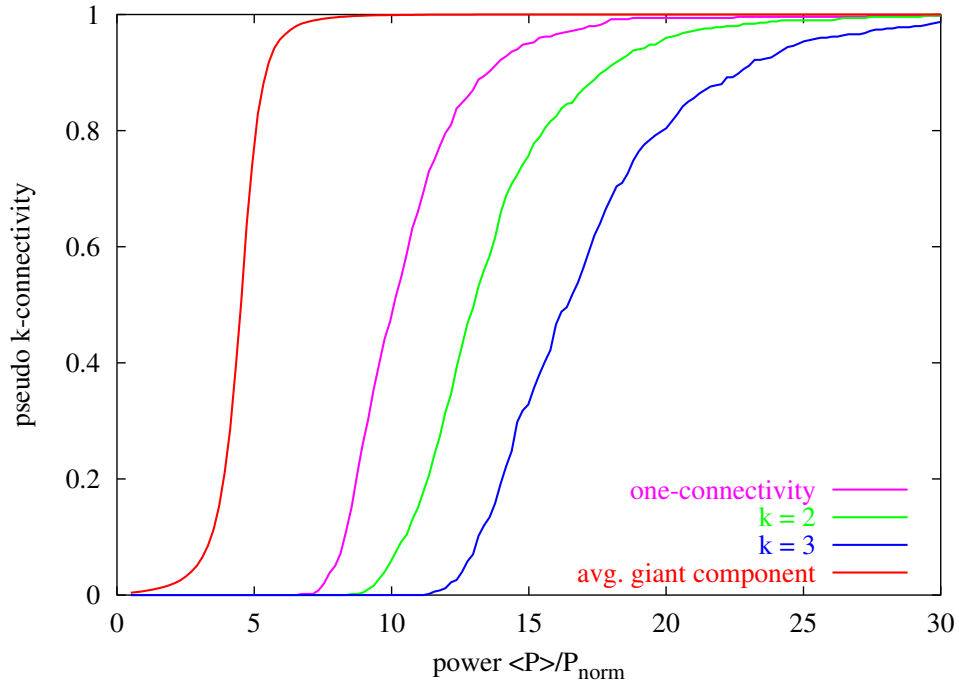

Figure 2.3: One-connectivity and pseudo k-connectivity
One-connectivity and pseudo k-connectivity as a function of transmission power $\langle P \rangle / P^{\text{norm}}$ upon using the constant-P rule. The number of nodes for the simulated 500 random homogeneous point patterns has been fixed to $N = 1600$ and the path-loss exponent has been set to $\alpha = 2$. Different curves correspond to the one-connectivity (pink), $k = 2$ (green), $k = 3$ (blue); for comparison the average relative giant component is shown as the red curve.

onset for a non-vanishing probability in case of the multifractal point patterns is shifted to extremely large values of the transmission power. This is due to the inhomogeneous spatial clustering and demonstrates that the constant-P rule has serious drawbacks in such an environment.

In addition to these disadvantages within more complex surroundings it has to be mentioned again that the constant-P rule is by construction an artificial as well as unrealistic rule. It would require all nodes either to be designed for only the same single-valued transmission power operation or, given already network connectivity, to carry out fast synchronisation; of course, also an outside provider could adjust all node transmission powers to one single value, but this would give up the advocated philosophy of self-organising ad hoc networks. Nevertheless, the constant-P rule is a valuable and well-studied reference [46].

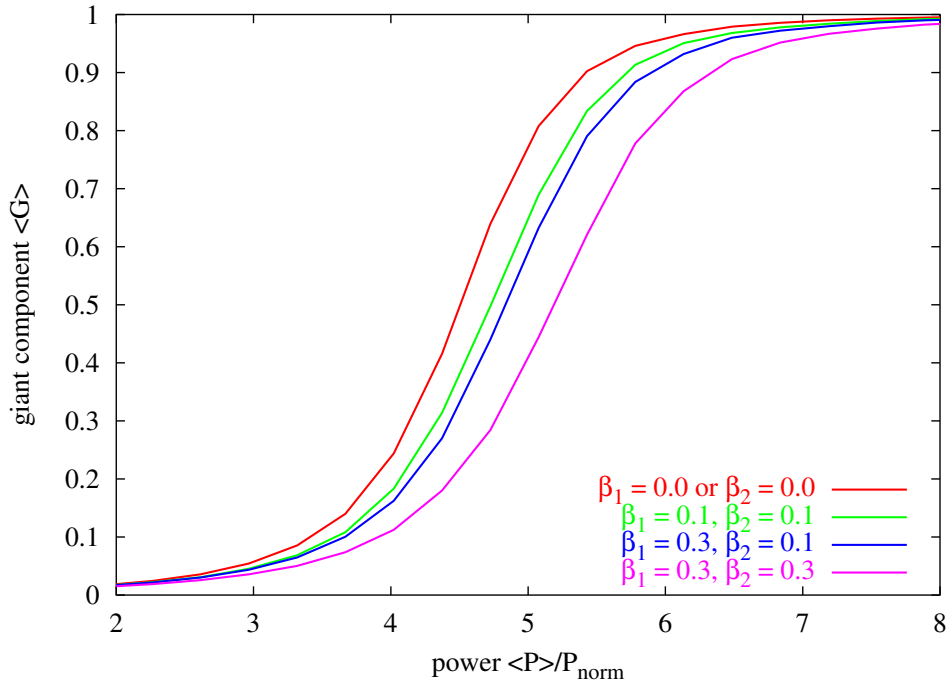

Figure 2.4: Giant component for the iid transmission power rule
Average relative giant $\langle G \rangle$ component as a function of the average transmission power $\langle P \rangle / P^{\text{norm}}$ upon using the iid transmission power rule for the bimodal distribution of equation (1.3). The number of nodes for each of the used 500 random homogeneous point patterns has been fixed to $N = 1600$ and the path-loss exponent has been set to $\alpha = 2$. Different curves correspond to different parameter choices: β_1 or $\beta_2 = 0.0$ (red, constant-P rule), β_1,$\beta_2 = 0.1$,0.1 (green), 0.3,0.1 (blue), 0.3,0.3 (pink).

2.2.2 iid transmission power

The idea behind an iid transmission power rule like the one presented in Section 1.4 is to add a number of long-range connections to the network that might increase connectivity and lead to shorter network diameters D.

Figure 2.4 illustrates simulation results obtained with random homogeneous point patterns. The average relative giant component $\langle G \rangle$ as a function of the average transmission power $\langle P \rangle$ is shown for some combinations of β_1 and β_2 of the bimodal distribution in (1.3). All settings result in a shift of the percolation threshold to larger $\langle P \rangle$ values when compared to the outcome of the constant-P rule. The randomness and spatial decorrelation of the long-range links introduced by the iid-P link rule does not allow for a shift of the percolation threshold towards smaller $\langle P \rangle$ values. A mean-field approach for the average node degree $\langle k \rangle$ confirms

this reasoning. According to equation (1.3) two values for the transmission power are possible: $P^{\text{low}} = \langle P \rangle (1 - \beta_1)$ with probability $p(P^{\text{low}}) = \beta_2/(\beta_1 + \beta_2)$ and $P^{\text{high}} = \langle P \rangle (1 + \beta_2)$ with probability $p(P^{\text{high}}) = \beta_1/(\beta_1 + \beta_2)$. A node i with low transmission power $P_i = P^{\text{low}}$ has bidirectional links to all other nodes placed within the circular disc characterised by its transmission range $R_{ij}^{\text{low}} = (P^{\text{low}})^{1/2}$ given $\alpha = 2$. Its node degree can then be described by the product of its disc area and the total density of nodes in the unit square $\langle k \rangle^{\text{low}} = \pi (R_{ij}^{\text{low}})^2 N = \pi N \langle P \rangle (1 - \beta_1)$. A node i with high transmission power $P_i = P^{\text{high}}$ has bidirectional links to all nodes within an inner circular disc characterised by the lower transmission range R_{ij}^{low} of the replying node. A node degree for this inner surrounding can be given in a similar manner than above: $\langle k^{\text{inner}} \rangle^{\text{high}} = \pi (R_{ij}^{\text{low}})^2 N = \pi N \langle P \rangle (1 - \beta_1)$. Any node j placed in an annulus around i characterised by an inner radius R_{ij}^{low} and a outer radius $R_{ij}^{\text{high}} = (P^{\text{high}})^{1/2}$ can only set up bidirectional links with i if $P_j = P^{\text{high}}$, too. An estimate for the number of long-range neighbours is thus given by $\langle k^{\text{outer}} \rangle^{\text{high}} = p(P^{\text{high}}) \pi [(R_{ij}^{\text{high}})^2 - (R_{ij}^{\text{low}})^2] N = p(P^{\text{high}}) \pi N \langle P \rangle [(1 + \beta_2) - (1 - \beta_1)]$. The average node degree $\langle k \rangle$ can then be written as:

$$\begin{aligned} \langle k \rangle^{\text{bimodal}} &= p(P^{\text{low}}) \langle k \rangle^{\text{low}} + p(P^{\text{high}}) [\langle k^{\text{inner}} \rangle^{\text{high}} + \langle k^{\text{outer}} \rangle^{\text{high}}] \\ &= \pi N \langle P \rangle \left(1 - \frac{\beta_1 \beta_2}{\beta_1 + \beta_2} \right) \end{aligned} \tag{2.2}$$

For all $\beta_1 > 0$ and $\beta_2 > 0$ the node degree $\langle k \rangle^{\text{bimodal}}$ for the iid transmission power assignment according to the bimodal distribution (1.3) is smaller than in the constant-P case with $\langle k \rangle^{\text{constP}} = \pi N \langle P \rangle$. In terms of continuum percolation it is this reduction of the average number of links that causes the shift of the percolation threshold. The limited number of long-range connections can not compensate this shortcoming.

Without showing it shall be remarked that similar results of the simulations are obtained for other point patterns of multifractal or Manhattan type. Also other iid transmission power distributions like a lognormal distribution did not show better performance.

As already mentioned the uncorrelated introduction of long-range links is not sufficient for a better network connectivity. They rather have to be established in areas with a lower density of nodes and to connect separated clusters. This will be achieved by the minimum node degree rule as presented in the next section.

2.2.3 Minimum node degree

The simulation results obtained with the local minimum node degree rule are illustrated in Figures 2.5 and 2.6 and compared to the respective outcomes of the constant-P rule. For geometric graphs based on random homogeneous spatial point patterns the threshold of the average giant component $\langle G \rangle$ is reduced by

about a factor of 1.45; for the probability for one-connectivity the threshold reduction factor around 2.05 is even slightly larger. Note that for $k^{\min} = 3$ the relative giant component is already very close to one; once $k^{\min} \geq 6$ the probability for one-connectivity becomes one almost surely for networks of up to at least 1600 nodes. It is worth mentioning that for a wide spectrum of given values for the minimum degree $k^{\min}$ it holds $\langle k \rangle = a\, k^{\min}$ with $a \approx 1.2 \ldots 1.4$.

For geometric graphs based on random multifractal spatial point patterns the local minimum node degree rule beats the constant-P rule even more impressively. Whereas for the constant-P rule the giant component threshold is very blurred, the local rule transforms it into a sharp threshold, which almost exactly coincides with the respective threshold obtained for the previously discussed random homogeneous point patterns. Due to the strong spatial clustering the artificial constant-P rule leads to a highly suppressed one-connectivity yield; for the parameters used for Figure 2.6 the onset for non-vanishing one-connectivity is already at a rather large transmission power around $\langle P \rangle / P^{\text{norm}} \approx 13$, but it needs an even much larger $\langle P \rangle$ for this probability to come close to one. The local minimum degree rule, on the other side, perfectly adapts to the new environment, compensating the strong spatial clustering introduced by the heterogeneous spatial point patterns and pushing the one-connectivity threshold down to $\langle P \rangle / P^{\text{norm}} \approx 6$, nearly matching the respective homogeneous point pattern threshold; with $k^{\min} \approx 7$, which is equivalent to $\langle P \rangle / P^{\text{norm}} \approx 10$, the probability for one-connectivity is practically one.

In case of the geometric graphs based on random Manhattan spatial point patterns and in comparison with the artificial constant-P rule, the local minimum node degree rule reduces the threshold of the average giant component as well as of the one-connectivity by about a factor of 1.7. The magnitude of reduction is comparable to the values stated for random homogeneous point patterns. For the one-connectivity probability to become almost one, a value of at least $k^{\min} = 10$ is needed for the model parameters stated in Figure 2.6.

2.3 Transmission power and node degree distribution

The results for the node degree and transmission power distributions from the minimum node degree rule compared to their constant-P counterparts will be presented. The particular focus is on random homogeneous spatial point patterns. The minimum degree is chosen to be $k^{\min} = 6$ in order to guarantee one-connectivity almost surely; consult again Figure 2.6. The node degree distribution $p(k)$, which reflects the probability for a node to have k neighbours, is shown in the upper graph of Figure 2.7. The number of nodes has been fixed to $N = 1600$ and the path-loss exponent has been set to $\alpha = 2$. The values are sampled over 500 network realisations. By construction, $p(k) = 0$ for $k < k^{\min}$. It comes with

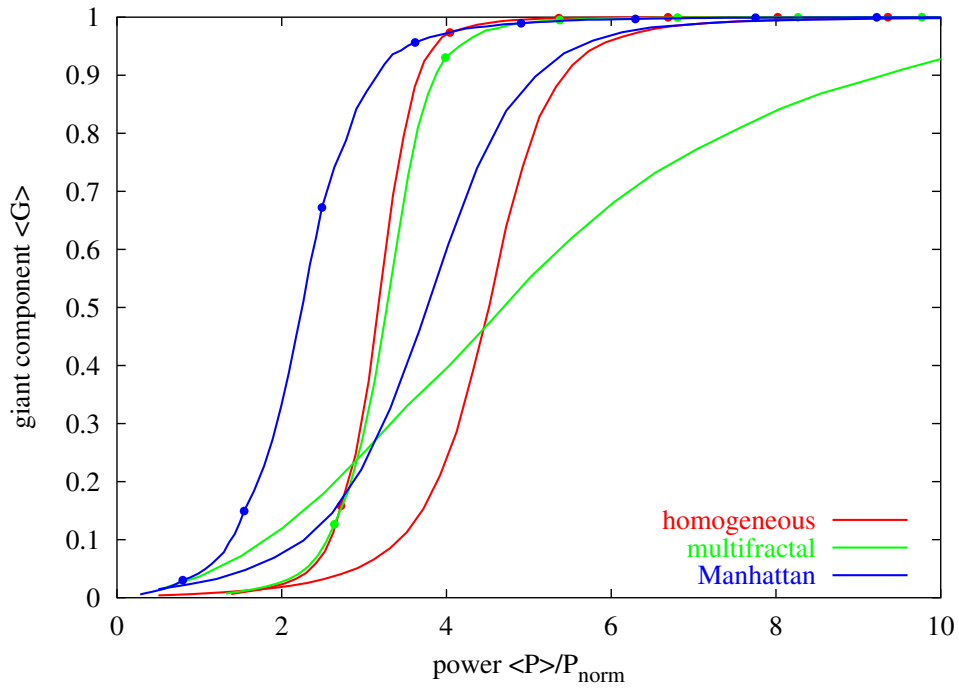

Figure 2.5: Giant component for the minimum node degree rule
Average relative giant component $\langle G \rangle$ as a function of average transmission power $\langle P \rangle / P^{\mathrm{norm}}$ upon using the local minimum node degree rule (curves with circles) and the constant-P rule (curves without circles, identical to curves of Figure 2.2). The number of nodes has been fixed to $N = 1600$ and the path-loss exponent to $\alpha = 2$. Different line colours correspond to different random spatial point patterns: homogeneous (red), multifractal (green) with $\omega = 0.4$ and $j = 5$, Manhattan (blue) with $N_x = N_y = 7$; sampled over 500 graph realisations. From the left to the right the circles on each minimum node degree rule curve stand for $k^{\min} = 2{-}7$ (homogeneous, multifractal) and $2{-}9$ (Manhattan).

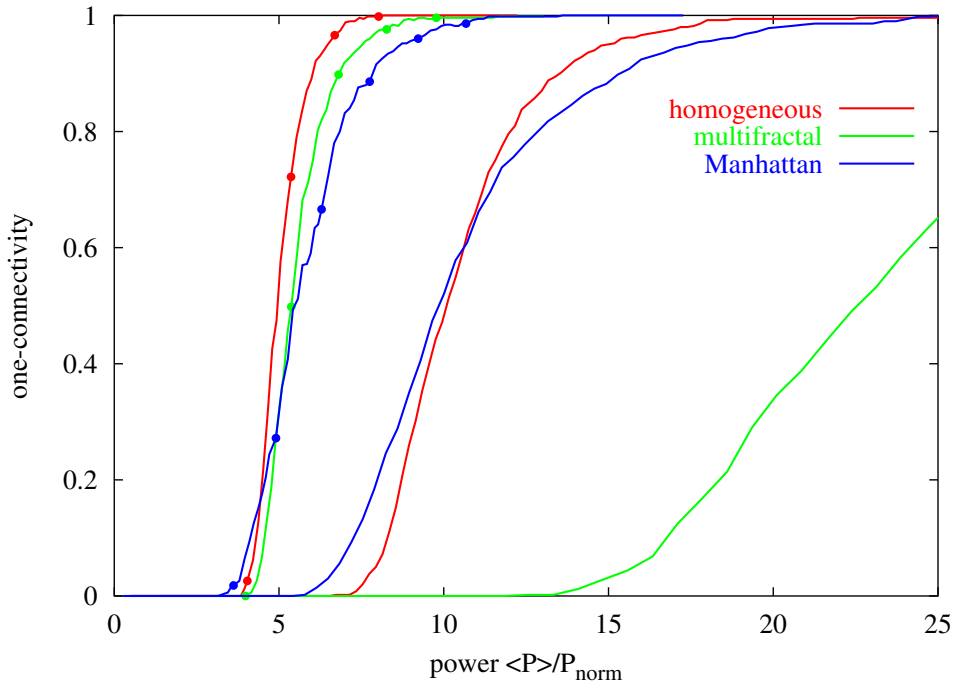

Figure 2.6: One-connectivity for the minimum node degree rule
One-connectivity as a function of average transmission power $\langle P \rangle / P^{\mathrm{norm}}$ for the minimum node degree rule (curves with circles) and the constant-P rule (curves without circles). Parameters are $N = 1600$ and $\alpha = 2$. Different line colours correspond to different random spatial point patterns: homogeneous (red), multifractal (green) with $\omega = 0.4$ and $j = 5$, Manhattan (blue) with $N_x = N_y = 7$; each sampled over 500 graph realisations. From the left to the right the circles on each minimum node degree rule curve stand for $k^{\mathrm{min}} = 3{-}6$ (homogeneous), $3{-}7$ (multifractal) and $5{-}10$ (Manhattan).

a peak at $k = k^{\min}$ and falls off sharply for $k > k^{\min}$. Upon changing from discrete to continuous $k \geq k^{\min}$, the distribution can be fitted as the superposition $p(k) = a\delta(k-k^{\min})+bN(k;\mu^{\text{Gauss}},\sigma)$ of a δ-function, placed at the minimum node degree, and a normalised Gaussian with shift μ^{Gauss} and width σ. The parameters used for the fit in Figure 2.7 are $a = 0.14$, $b = 1.49$, $\mu^{\text{Gauss}} = 5.93$ and $\sigma = 2.39$.

For comparison the much broader node degree distribution resulting from the constant-P rule is shown in the upper graph of Figure 2.7. The setting $P_i/P^{\text{norm}} = \langle P\rangle/P^{\text{norm}} = 20$ is necessary to guarantee one-connectivity almost surely as stated in Figure 2.6. A geometric argumentation is helpful in order to understand the distribution. For simplicity the path-loss exponent is set to $\alpha = 2$. Picking one node $i \in \mathcal{N}$, the probability to find k_i other nodes inside its circular disc with radius $R_{ij} = P_i^{1/2} = P^{1/2}$ is equal to the binomial distribution [34]

$$p(k) = \binom{N-1}{k} q^k (1-q)^{N-1-k} \approx (\lambda^k/k!)e^{-\lambda} \tag{2.3}$$

where $q = \pi R_{ij}^2$ is the fraction of the disc area to the unit square. For N large and q small, $p(k)$ becomes a Poissonian with mean $\langle k\rangle = \lambda = q(N-1)$. The small deviations of the generic distribution in Figure 2.7 to the Poissonian are due to finite-size effects, that on average nodes close to the boundary experience a degree less than λ. The generic distribution can be well fitted with a normalised Gaussian; parameters used in the upper graph of Figure 2.7 are $\mu = 18.98$ and $\sigma = 4.92$.

The transmission power distribution in case of the constant-P rule is simply a δ-function at $P_i/P^{\text{norm}} = 20$ as illustrated in the lower graph of Figure 2.7. Its position naturally represents an approximate upper bound for transmission power values obtained from the minimum node degree rule. For the distribution resulting from the minimum node degree rule an analytic estimate can be given. The path-loss exponent is again set to $\alpha = 2$. As mentioned above the degree distribution of a node, not too close to a boundary, is given by the Poissonian $p(k) = (qN)^k e^{-qN}/k!$, with $q = (P/P^{\text{norm}})\pi(R^{\text{norm}})^2 = (P/P^{\text{norm}})/N$ representing the area covered by the node's transmission radius relative to the unit square. Here, the transmission power P is kept fixed and the degree k is the discrete random variable. Another look on the upper graph of Figure 2.7 reveals, that the very narrow degree distribution resulting from the minimum node degree rule can be crudely approximated as $p(k) \approx \delta(k-k_0)$ with $k_0 = \langle k\rangle = 7.36$. From that perspective the node degree k is now kept fixed within the above Poissonian and the transmission power is considered as the continuous random variable. This leads to the transmission power distribution

$$p(P/P^{\text{norm}}) \sim (P/P^{\text{norm}})^{k_0} \exp\left(-P/P^{\text{norm}}\right) . \tag{2.4}$$

It corresponds to a Gamma distribution $p(x;a,b) = x^{a-1}e^{-x/b}/(b^a\Gamma(a))$ with $x = P/P^{\text{norm}}$ and $a = k_0 + 1 \approx 8.36$, $b = 1$. The actual best fit to the sampled

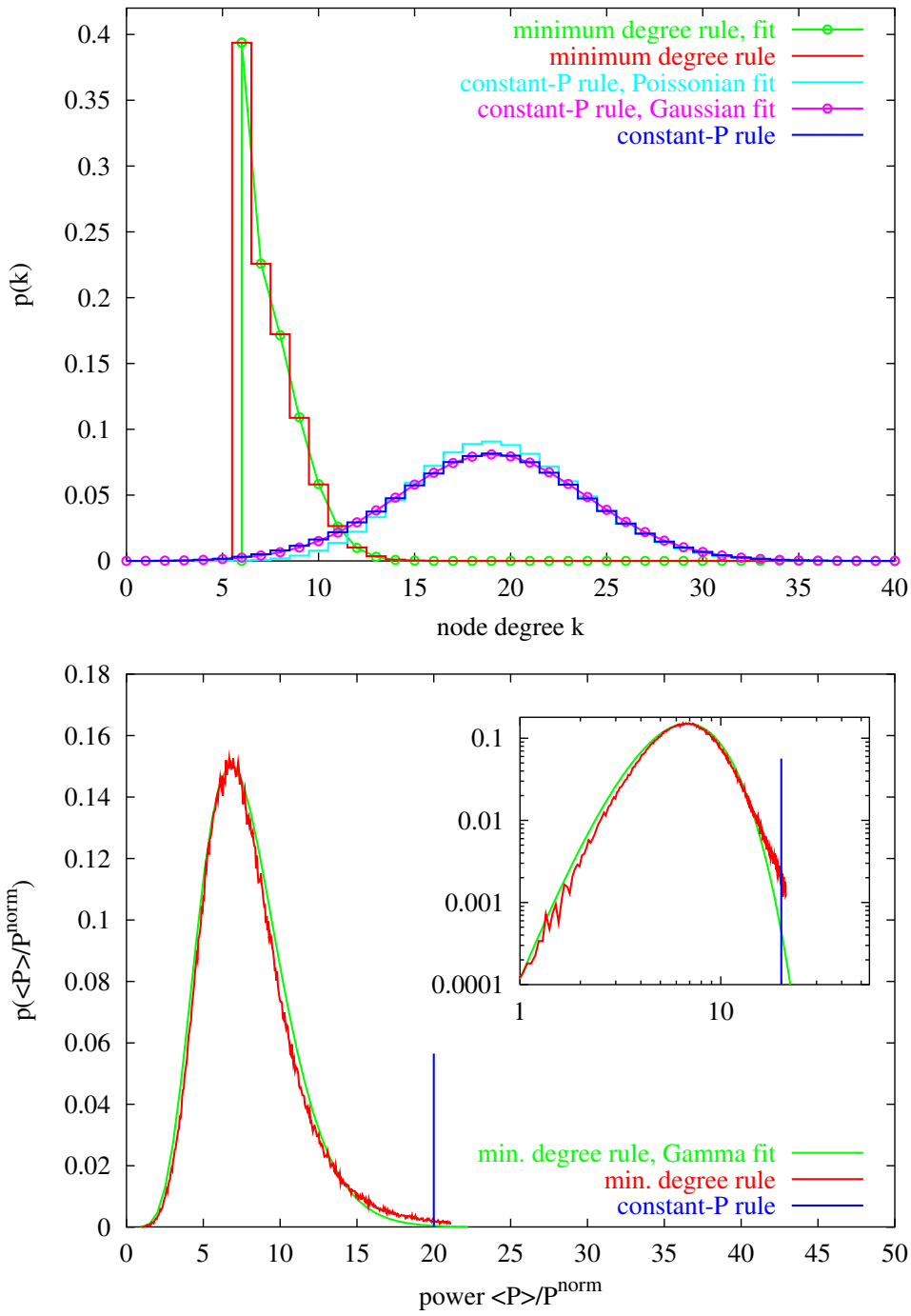

Figure 2.7: Node degree and transmission power distributions
Node degree and transmission power distributions obtained from the minimum degree and the constant-P rule for random homogeneous point patterns; sampled over 500 graph realisations. The number of nodes has been fixed to $N = 1600$ and the path-loss exponent has been set to $\alpha = 2$. Details about the fitted distributions are given in the main text.

distribution, shown in the lower graph of Figure 2.7 , yields the parameters $a = 7.74$, $b = 1.01$ and more or less confirms the given estimate.

A similar discussion for point patterns other than random homogeneous is given in detail in [21].

2.4 Summary

Connectivity represents an important issue for wireless mobile ad hoc communication networks. In a self-organising manner, the participating ad hoc nodes have to tune their transmission powers to establish direct one-hop communication links to their spatial neighbours. Furthermore all other nodes should be reachable via multihop routes. In analogy to two-dimensional continuum percolation a percolation threshold for the transmission power exists in the constant-P case above which the connectivity of a network can be guaranteed. However, the percolation threshold does show a sensitive dependence on the specific spatial patterning of the ad hoc nodes. Different classes of uncorrelated and correlated random point patterns, like homogeneous, multifractal or Manhattan-like distributions, make a significant difference. The constant-P rule as well as the iid transmission power rule are not flexible enough to adapt to local spatial inhomogeneities. A local generalisation of these rules leads to the minimum node degree rule that can be viewed as a first important step towards a self-organising connected network structure. It requires each ad hoc node to be connected to a minimum number of closest neighbours. This distributed rule is able to counterbalance local spatial inhomogeneities occurring in the different point patterns. As a function of average transmission power the percolation thresholds associated to the various considered classes of random point patterns almost collapse onto each other and are tremendously reduced, when compared to the outcomes with the global rules.

For the networks used in the following chapters on data traffic connectivity is a prerequisite. To be on the safe side all networks of the constant-P type are generated with $P = 24P^{\mathrm{norm}}$ for a path-loss exponent $\alpha = 2$ which translates into $k^{\mathrm{constP}} = 24$ in the general case. For the networks of minimum degree type $k^{\mathrm{min}} = 8$ is required.

3 Generic data traffic

For many networks, natural or artificial, the understanding of the functional properties requires the study of dynamical processes on the network graph. The principles of the dynamical rules, the properties of the network constituents as well as the topology of the network graph have an impact on the dynamical behaviour of the system [54, 55]. It is for that reason why dynamical processes in the form of exchange of data packets have to be studied for wireless ad hoc communication networks. In the previous chapter possible topologies of ad hoc network graphs have been proposed, fulfilling the prerequisite of connectivity. Based on these topologies the question of this chapter is how these networks behave under conditions of data traffic. In particular it will be asked how much data can be transmitted in a given network, what are suitable measures for this critical load and which dynamical properties arise on a global as well as a local scale. One can also test the influences of different setups of data generation as well as the impact of the topology of the network graph on the dynamical behaviour.

These and other phenomena can be best studied by setting up an appropriate model that includes the basic key points of a real world communication network with data traffic. Such a model allows the prediction of particular features of network data traffic and the verification of analytically obtained results. As a suitable reference throughout this chapter the constant-P topology with $k^{\mathrm{constP}} = 24$ is employed to study relevant dynamical properties of network traffic. Since all nodes have the same transmission power no unidirectional links exist in the realisations of the network graph. All networks are exclusively generated on random homogeneous point patterns. An introduction to the principles of network traffic and the specific setup of the model will be given in Section 3.1. Section 3.2 looks on global observables of the generic data traffic in such networks whereas Section 3.3 studies properties of selected nodes. A detailed analytic approach to describe the outflux of packets from a node is given in Section 3.4. A brief discussion of the behaviour of networks with minimum node degree rule for connectivity under generic data traffic follows in Section 3.5 whereas Section 3.6 focuses on the impact of special user behaviour on the dynamical properties of network traffic. A summary in Section 3.7 concludes the chapter.

3.1 Principles of network traffic

3.1.1 A model of generic data traffic

Retaining the conventions from Chapter 1 mobility of the nodes is not taken into account, instead an ensemble of static point patterns is considered. In that respect it is consistent not to allow the nodes to change their transmission power with time. In this way one ensures that for a given static point pattern the network graph stays the same for the whole time of data traffic execution.

To completely reproduce a real-world data traffic one would have to implement engineering-like protocols which by use of control messages could precisely manage all traffic actions. Since the research interest of this thesis is more focused on the overall behaviour and statistical properties of the actual data transfer a simplified generic model will be employed. The model is based on discrete time steps. All nodes within the network are assumed to start each time step simultaneously. At the very beginning of a time step new packets can be generated. After a short contention phase in which the data network organises the following packet transmissions the remaining period is reserved for the actual data transfer.

Packet creation In the simplest version of the model one single new data packet of fixed size is created at each node $i \in \mathcal{N}$ with a probability $\mu_i = \mu < 1$. This probability is also referred to as the packet creation rate whereas the product $\mu \cdot N$ corresponds to the mean of the total number of newly created packets in one time step within the network. For the simulation of specific user behaviour the generation of more than one packet at a certain node is employed. This will be discussed in more detail in Section 3.6. If a packet is created at a certain node i a destination is randomly chosen among all other nodes $\{\mathcal{N} \setminus i\}$. Labeled with the time of generation and the final destination node the packet is put at the end of the buffer queue of the generating node. All nodes are assumed to have infinite capacity. Nodes, for which a new packet has been created, are blocked for the remainder of this time step.

Routing For all data traffic simulations employed in this chapter data packets are forwarded along the shortest path $\sigma_{i,f}$ in the sense of Section 1.1 from the generating node i to the final receiver f [8, 56].

The shortest paths $\sigma_{i,f}$ within a given network realisation $\mathcal{G}$ are obtained using Dijkstra's algorithm [57]. The algorithm solves the shortest path problem for a connected graph $\mathcal{G} = (\mathcal{N}, \mathcal{L}^{\text{bidir}})$ which has non-negative link weights. The cost of a path between two nodes is the sum of costs of the link weights in that path. To find shortest paths in the hop metric all weights are initialised to $w_{i,j} = 1$. The cost of a path is thus equal to the distance. The algorithm works by constructing a subgraph $\mathcal{H}$ of $\mathcal{G}$ starting from an initial node i such that the distance of any

node $j \in \mathcal{H}$ from i is known to be a minimum within $\mathcal{G}$. Initially $\mathcal{H}$ is simply the single node i, and the distance of i from itself is known to be zero. Links (and corresponding nodes) are added to $\mathcal{H}$ at each stage by (a) identifying all the links $\mathcal{L}^{\mathcal{H}-\mathcal{G}} = j_1 \leftrightarrow j_2$ such that j_1 is in $\mathcal{H}$ and j_2 is in $\{\mathcal{G} \setminus \mathcal{H}\}$, and then (b) choosing the link $j_1^* \leftrightarrow j_2^*$ which gives the minimum distance from i to any j_2^* among all links $\mathcal{L}^{H-G}$. For each newly added node j_2^* its predecessor node j_1^* is stored which latter allows to backtrace each shortest path from any j towards i. The algorithm terminates once $\mathcal{H} = \mathcal{G}$. Since the network graph $\mathcal{G}$ is static the information about shortest paths obtained once remains valid during the data traffic simulation.

Each node of $i \in \mathcal{N}$ has full access to the information about the shortest paths to all other nodes $f \in \mathcal{N} \setminus i$. In that sense each node knows to which of its neighbouring nodes it has to send a packet from its buffer queue in order to reach a given final recipient. Thus for every node i a routing table

$$F_f^i = \mathcal{S} \tag{3.1}$$

is supplied that for every final receiver $f \in \{\mathcal{N} \setminus i\}$ provides a list $\mathcal{S}$ of bidirectional neighbours of i that minimises the hop-distance $d_{i,f}$ towards f. This routing scheme is referred to as full table routing [8].

If more than one shortest path to the final receiver f exist that split at i then more than one possible next hop neighbours are listed in $\mathcal{S}$. For all routing decision within this chapter path degeneracy between two nodes i and f will be completely suppressed by randomly choosing one of the shortest paths between i and f. This choice is then once and forever kept fixed, not only for one particular execution of the end-to-end route from i to j but for all future communications between the same original sender and the same final receiver. All routes in the network between any two nodes are thus completely deterministic.

Medium access control (MAC) During the short contention phase at the beginning of each time step, the non-blocked nodes (that have not generated data packets in that particular time step) compete for gaining sender status. A node i out of this set of nodes with a non-zero buffer is randomly picked first and obtains the permission to transmit its first packet in the buffer queue to the respective neighbouring node defined by $F_f^i = \mathcal{S} = j$. Both, the sending node i as well as receiving node $j \in \mathcal{N}_i$ MAC-block their respective outgoing neighbours $n \in \mathcal{N}_i^{\text{out}} \cup \mathcal{N}_j^{\text{out}}$ for the remainder of this time step. Then another node with non-zero buffer that has not been blocked so far is chosen at random and tries to attempt the permission for the transmission of its first-in-line packet. If the intended receiver has already been blocked before, the node tests its second-in-line packet and so on, until either the first idle recipient is found or the end of its buffer queue is reached. This strategy is denoted as first-in-first-possible-out (fifpo). If this node succeeds to gain sending permission for a packet, it then MAC-blocks

again its remaining outgoing neighbours as well as those of the receiving node. This iteration is repeated until no free one-hop transmission is left.

In case of existing unidirectional links, which occur for a heterogeneous distribution of transmission power, it might occur that an unidirected outgoing link associated to the latest MAC-operation blocks a node, which within the contention phase of this time step has already gained sender or receiver status in a previously approved one-hop transmission. For such cases, the previously assigned sender and receiver are blocked again. Furthermore, their outgoing neighbourhoods remain blocked and are not freed.

Packet transmission After the contention phase at the beginning of the time step all nodes with sending permission now submit their selected packet and remove it from their respective buffer queue. The receiving nodes either add the incoming packet to the end of their buffer queue or, if they are the final recipient, destroy the packet. This concludes the actions taken for one time step and the whole process is repeated for the next time step. From a single-node perspective, a node can only perform a single operation per time step: either it creates a new packet, it sends a packet, it receives a packet from a neighbour, it is blocked by a neighbour or it remains idle.

It should be pointed out that the simple blocking of outgoing neighbours via medium access control (MAC) only reduces interference, but does not eliminate it completely. Interference might occur due to the cumulative strength of many, although distant one-hop transmissions taking place at the same time [58]. Such effects are not taken into account in the present data traffic simulation.

3.1.2 Inbetweenness measures

Inbetweenness centrality is a classical graph theoretical measure that counts the number of shortest paths via a given node [59, 33]. Since data traffic within the above stated model is based on shortest routes this measure becomes an important tool for the understanding of different phenomena within this chapter.

The so-called node inbetweenness employed in this thesis is given by

$$B_i^{\text{node}} = \frac{1}{(N-1)^2}\left(\sum_{k\neq f\in\mathcal{N}\setminus i}\frac{|\sigma_{k,f}(i)|}{|\sigma_{k,f}|}\right) \tag{3.2}$$

where $|\sigma_{k,f}|$ represents the total number of shortest paths between k and f and $|\sigma_{k,f}(i)|$ the number of shortest paths between k and f that passes i. The fraction $1/(N-1)^2$ is due to a proper normalisation of the node inbetweenness $B_i^{\text{node}} \in [0,1]$. A total of $N(N-1)$ different end-to-end communications is possible. In order to be counted in (3.2) node i can not act as receiver which reduces the

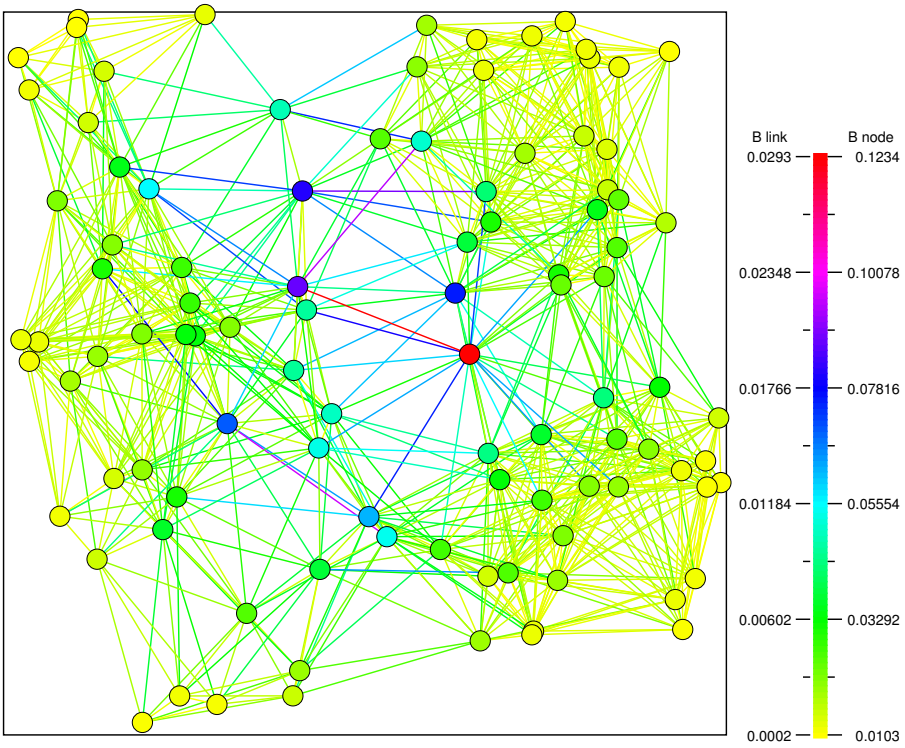

Figure 3.1: Visualisation of the inbetweenness
The colour of the nodes/links correspond to the node/link inbetweenness. This is a realisation of a network graph of 100 nodes, power assignment with constant-P rule ($k^{\mathrm{constP}} = 24$). For reasons of reference this particular spatial distribution of the nodes for a random homogeneous point pattern will be called constant-P reference network.

summation by $(N-1)$. Thus $N(N-1)-(N-1) = (N-1)^2$ is the number of possible end-to-end communications summed over in equation (3.2).

So far the node inbetweenness defined in (3.2) is a purely graph theoretical measure. But the generic data traffic model uses shortest paths for the packet forwarding as explained in the previous subsection. In that respect the node inbetweenness is also a measure of how frequently a node will be used for packet transfer. For a certain route execution from node j to node f only nodes shall raise their inbetweenness counter that do really store the particular data packets in their respective buffer queue. This is not the case for the final receiving nodes since it destroys the arriving packet. This is the reason for the exclusion of the final node in the counting in (3.2).

For the calculation of the node inbetweenness an all-pairs shortest path algorithm is used similar to that described in [60]. The algorithm starts out from node

i with distance assignment $d = 0$, all neighbours $j \in \mathcal{N}_i$ get distance $d = 1$. For each node j with distance d all of its neighbours j_2 are updated as follows. If j_2 has not already been assigned a distance, it now gets distance $d+1$ and j is identified as j_2's predecessor. If j_2 has already distance $d+1$ then no further assignment is necessary. Simply j is added to the list of predecessors. If j_2's distance is smaller then $d+1$ nothing happens at all. Finally the distance assignment is started again with the neighbours of j_2 with distance $d \leftarrow d+1$ and so on until the highest distance is reached.

The sum rule

$$\begin{aligned} N[(N-1)^2\langle B_i^{\text{node}}\rangle] &= N(N-1)D \\ (N-1)\langle B_i^{\text{node}}\rangle &= D \end{aligned} \tag{3.3}$$

relates the average node inbetweenness to the diameter of the network D. The square bracket on the left side of (3.3) represents the average number of shortest paths via one node of a network with size N.

Since the definition of the node inbetweenness B_i^{node} in (3.2) counts degenerate shortest paths in a weighted fashion it is not fully consistent with the actual data traffic. In the traffic model one predefined shortest path $\sigma_{i,f}^*$ is always superior to other degenerate versions. This causes the nodes along $\sigma_{i,f}^*$ to be frequented more often. Due to the randomness in the selection of the $\sigma_{i,f}^*$ the resulting differences are expected to be negligible.

In complete analogy to the definition of the node inbetweenness B_i^{node} in (3.2) a link inbetweenness $B_{i,j}^{\text{link}}$ can be defined.

$$B_{i,j}^{\text{link}} = \frac{1}{N(N-1)} \left(\sum_{k \neq f \in \mathcal{N}} \frac{|\sigma_{k,f}(i,k)|}{|\sigma_{k,f}|} \right) \tag{3.4}$$

It is a normalised measure of the number of shortest paths passing a link between the two neighbouring nodes i and j. Since there are no restrictions on the receiving node, this measure is normalised by the total number of end-to-end communications $N(N-1)$.

For visualisation, the nodes of a geometric graph with 100 nodes and the constant-P rule ($k^{\text{constP}} = 24$) are plotted in Figure 3.1. The colour of the nodes and the links correspond to their inbetweenness. The nodes and links marked with red have highest values of the corresponding inbetweenness measures. Due to the congruence with the generic data traffic this illustration gives an impression of how frequent various nodes are used to forward data packets along end-to-end communications. For reasons of reference this particular spatial distribution of the 100 nodes for a random homogeneous point pattern will be called constant-P reference network. It is the preferably studied spatial arrangement of nodes within this and the next chapter.

3.2 Characterisation of data traffic I: global

Observables like the travelling time of a data packet and the maximum number of packets to be served in one time step are typical for the characterisation of data networks under traffic conditions. They allow for an understanding of the systems behaviour as a whole. In contrast there are also observables characterising the behaviour of a single node within a complex network. Although this section focuses on the global observables of network traffic close connections to the single node picture become obvious.

3.2.1 Throughput

In general one can distinguish between two regimes of data traffic: the sub- and the supercritical phase. The characteristic feature of the supercritical phase is non-ability of the system to serve all data packets that are generated in the network. Above a critical packet creation rate μ^{crit} that separates the two regimes certain nodes receive more packets that they can process. This causes an accumulation of packets in these buffer queues. The number of active packets in a network $M(t)$ counts all the data packets presently contained in the system:

$$M(t) = \sum_{i \in \mathcal{N}} n_i(t) \tag{3.5}$$

n_i is the number of packets in the buffer queue of node i. In the supercritical data traffic regime the average number of active packets within the network $\langle M(t) \rangle_{\Delta t}$, with Δt referring to a temporal average, increases as a linear function of the elapsed time. Figure 3.2 documents this linear behaviour, where generic data traffic in a constant-P reference network ($k^{\text{constP}} = 24$) with 100 nodes was looked at.

In contrast, in the subcritical phase with $\mu < \mu^{\text{crit}}$ the average number of active packets in the network is finite and independent of time. A respective simulation result is also contained in Figure 3.2. All data packets generated in the subcritical regime will by transfered to their final receivers in finite time. The ratio of finally transmitted packets to the number of generated packets is consequentially equal to 1. In the transition to the supercritical regime this ratio decreases as packets are accumulating the buffer queues of overloaded nodes. Figure 3.3 is a typical representation. Data traffic was measured in 200 network realisations each containing 100 nodes with the constant-P power assignment ($k^{\text{constP}} = 24$). Error bars correspond to the variance of the sampling over the 200 network realisations. The critical packet creation rate $\langle \mu^{\text{crit}} \rangle \approx 0.0103$ is characterised as the point where the ratio of generated to finally received packets deviates from 1.

The end-to-end throughput T_{e2e} is a measure of the capacity of a network. It is defined as

$$T_{e2e} = \mu^{\text{crit}} N \qquad . \tag{3.6}$$

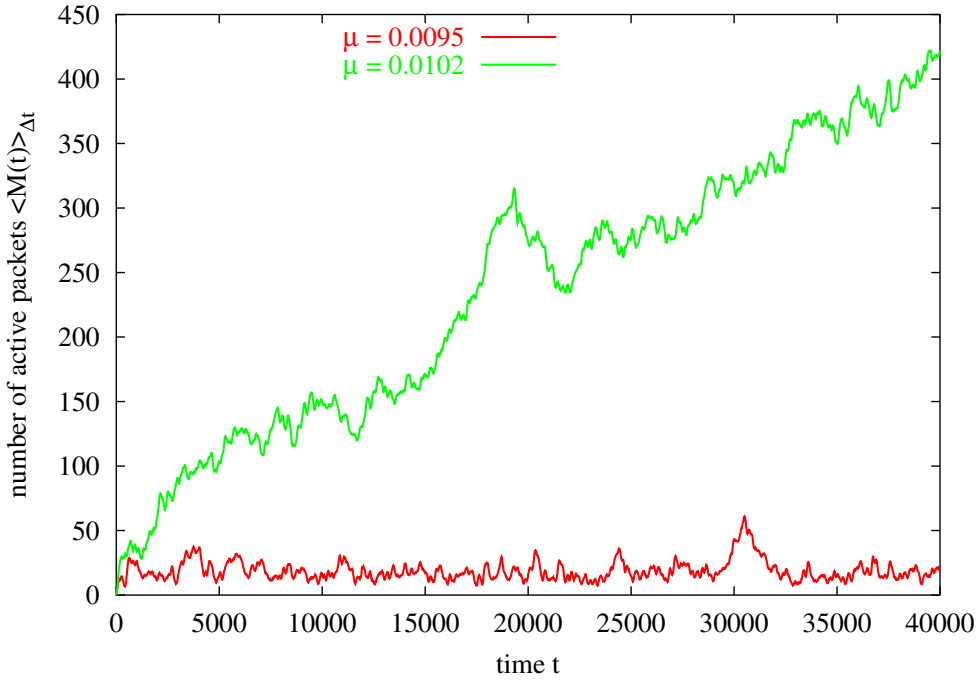

Figure 3.2: Sub- and supercritical regime
The number of active packets $\langle M(t)\rangle_{\Delta t}$ is shown for two different packet creation rates $\mu = 0.0095$ and $\mu = 0.0102$ depending on the duration of the simulation run. The data traffic simulation was performed on a constant-P reference network of 100 nodes with $k^{\text{constP}} = 24$. The linear dependence on time is a clear indication of supercriticality for $\mu = 0.0102$.

and represents the maximum number of packets that on average can be transmitted to their final receiver in one time step without network overloading. In that respect μ^{crit} can be understood as a specific capacity of that network.

For a conceptual understanding of the end-to-end throughput T_{e2e} a short experiment of thought is set up. One might imagine a network in which each node has sufficiently high transmission power to establish a direct bidirectional link to every other node in the network. Such a network is referred to as fully connected. Because any route i to j has distance $d_{ij} = 1$ the diameter D is consequently equal to 1. The set of MAC-blocked nodes during one one-hop transmission between i and j is given by $\mathcal{N}_{ij}^{\text{MAC}} = \{(\mathcal{N}_i^{\text{out}} \cup \mathcal{N}_j^{\text{out}}) \setminus i, j\} = \{\mathcal{N} \setminus i, j\}$. Since this set of MAC-blocked nodes spans all the remaining nodes of the network only one transmission per time step is possible. On the other hand the transmitted packet reaches its final receiver in one time step since $D = 1$. Thus for every generated packet precisely one time step is needed in order to transmit and destroy it. This

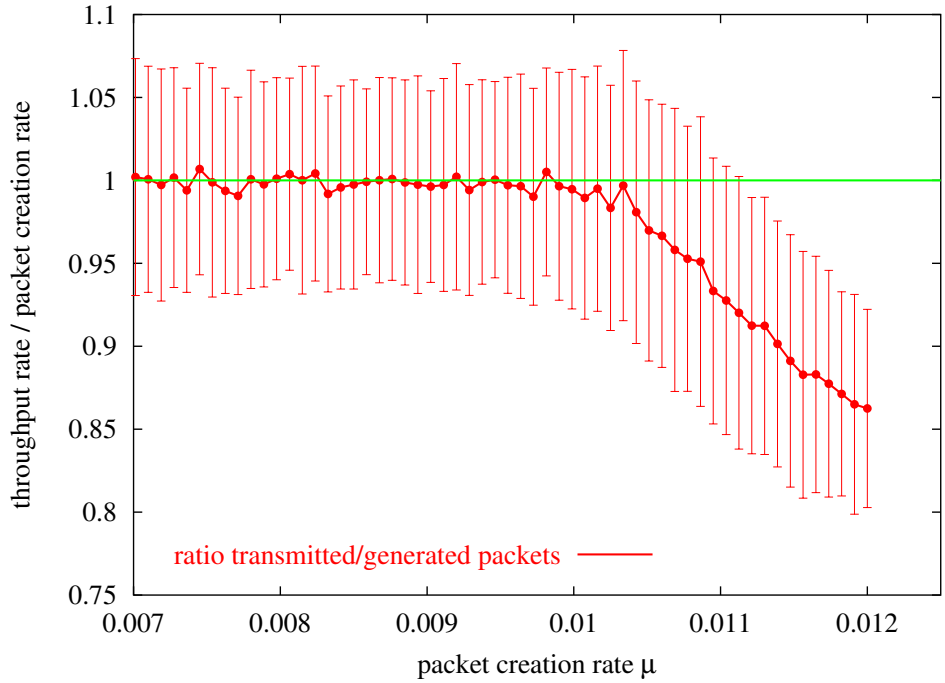

Figure 3.3: Transition from sub- to supercritical regime
The ratio of transmitted to generated packets is shown depending on the packet creation rate μ for networks of 100 nodes. A critical packet creation rate $\langle \mu^{\text{crit}} \rangle \approx 0.0103$ can be obtained from the graph above, from which on more packets are generated in the network than finally transmitted. The particular ratios were obtained by checking 200 different network realisations with constant-P rule ($k^{\text{constP}} = 24$) for $50,000$ time steps. Error bars correspond to the variance of the sampling process.

simple restriction is only satisfied for $\mu^{\text{crit}} = 1/N$ which yields for the end-to-end throughput $T_{e2e} = 1$. Note that this result is independent of the network size N.

As Figure 3.4 indicates, this is not the case for different network topologies. The graphs are obtained by sampling the critical packet creation rate μ^{crit} for networks of different size over 100 realisations. The critical packet creation rate μ^{crit} was determined with a binary search algorithm using nested intervals for simulation times of 10,000 time steps. At a critical system size $N^{\text{crit}} \approx 110$ the end-to-end throughput for the constant-P topologies ($k^{\text{constP}} = 24$) exceeds the value $T_{\text{e2e}} = 1$ for fully connected networks.

Focusing on the conflicting interests of an efficient routing with high throughput and the impact of the MAC protocol, Figure 3.4 provides another encouraging result. The ambition of the nodes in $\mathcal{N}$ to reach a possibly large neighbourhood

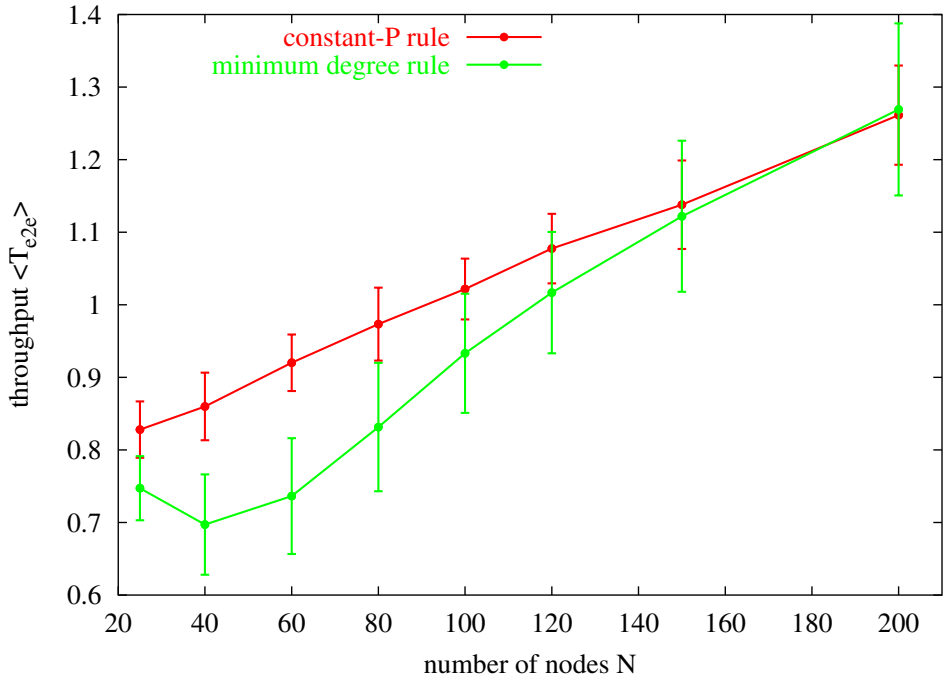

Figure 3.4: Throughput
The throughput for networks constructed with the constant-P rule ($k^{\mathrm{constP}} = 24$) and the minimum node degree rule ($k^{\min} = 8$) depending on the network size N is shown. The results are obtained by sampling over 100 network realisations. Each simulation spanned 10,000 time steps. The determination of the critical creation rate μ^{crit} was done by employing a binary search algorithm. The error bars correspond to the variance of the sampling process.

in order to minimise the number of one-hop transmission for upcoming data packet is only beneficial for networks of size up to N^{crit}, which depends on the specific topology of the network graph. For networks larger than N^{crit} the wide-ranging MAC-blocking of substantial parts of the network suppresses the possibility of having more than one transmission in the same time step. It is precisely this multitude of simultaneous but spatially separated transmissions on the shared communication medium that favours networks that are not based on long ranging links with their intensive MAC-blocking.

A simple estimate allows to get an impression of the dependence of the end-to-end throughput T_{e2e} on the system size N for a connected ($\mathcal{N}^{\mathrm{giant}} = \mathcal{N}$) but not fully connected topology. With $\langle|\mathcal{N}_{ij}^{\mathrm{MAC}}|\rangle$ being the average number of blocked nodes while transmitting one data packet from node i to its neighbouring node j the ratio $N/(2 + \langle|\mathcal{N}_{ij}^{\mathrm{MAC}}|\rangle)$ refers to the average number of maximum possible

one packet transmissions in one time step. For a network with diameter D the end-to-end throughput T_{e2e} can be approximated as

$$T_{e2e} = \frac{1}{D} \frac{N}{(2 + \langle |\mathcal{N}_{ij}^{\text{MAC}}| \rangle)} \quad . \tag{3.7}$$

This expression is consistent with the finding $T_{e2e} = 1$ for fully connected networks, since $D = 1$ and $\langle |\mathcal{N}_{ij}^{\text{MAC}}| \rangle = N - 2$. Employing the scaling behaviour $D \sim \sqrt{N}$ from Section 1.1 and assuming the average number of blocked nodes $\langle |\mathcal{N}^{\text{MAC}}| \rangle$ to be independent of the system size N the end-to-end throughput T_{e2e} should also scales like $T_{e2e} \sim N^{\kappa}$ with $\kappa = 0.5$ [46]. This scaling behaviour could not be verified for generic data traffic on topologies with constant-P as well as minimum node degree power assignment. In [35] the throughput T_{e2e} has been studied for networks of up to 2000 nodes. Scaling exponents of $\kappa = 0.38$ for the constant-P rule with $k^{\text{constP}} = 24$ and $\kappa = 0.24$ for the minimum node degree rule with $k^{\min} = 8$ have been measured. The inherent difficulty of the estimate in (3.7) is the sole utilisation of average network properties. It does not take the local character of criticality in network traffic into account. This will be resolved in Section 3.4 which allows a better reproduction of the actually observed scaling properties.

The crucial relation for every node $i \in \mathcal{N}$ to determine whether it is a sub- or supercritical state can be reduced to the simple relation

$$\mu_i^{\text{in}} \leqq \mu_i^{\text{out}} \tag{3.8}$$

The left hand side of this equation represents the rate of incoming packets to node i whereas the right hand side represents the rate of outgoing packets, given that node i has at leats one packet in its buffer queue. It is obvious that the buffer queue of a node i does not overflow and accumulate more and more data packets as long as this equation holds. As long as all nodes fulfil this relation the system is in the subcritical regime. The critical node is characterised as the first node for which $\mu_i^{\text{in}}/\mu_i^{\text{out}} \geq 1$ with increasing μ. This node will also be referred to as the most sensible node. A sufficient graph theoretical description of μ_i^{in} and μ_i^{out} would be a valuable tool in the analysis of different network topologies, since it allows the graph theoretical determination of the critical load.

By use of the inbetweenness measure B_i^{node} as introduced in (3.2) a graph theoretical description of the rate of incoming packets μ_i^{in} is possible. Per time step and in a statistical uniform manner, μN new packets are introduced into the network. Out of these new packets, on average the fraction $(N-1)^2 B_i^{\text{node}}/N(N-1) = (N-1)B_i^{\text{node}}/N$ will hop via node i along their respective shortest end-to-end communication path. Thus

$$\mu_i^{\text{in}} = \mu N \frac{(N-1)B_i^{\text{node}}}{N} = \mu(N-1)B_i^{\text{node}} = \frac{1}{\langle t_i^{\text{arrive}} \rangle} \tag{3.9}$$

describes the average number of incoming packets to node i. The upper part of Figure 3.14 illustrates this relation for a node out of a constant-P network with $k^{\mathrm{constP}} = 24$. The average interarrival time $\langle t_i^{\mathrm{arrive}} \rangle$ of node i is defined as the reciprocal of the influx μ_i^{in}. The interarrival time measures the time between two successive arrivals of packets at node i. It will be studied in more detail in Section 3.3.1.

The outflux μ_i^{out} is in a similar fashion linked to the so called sending time $\langle t_i^{\mathrm{send}} \rangle$ defined as the average time interval between subsequent sending events, given that node i has a non-zero buffer queue:

$$\mu_i^{\mathrm{out}} = \frac{1}{\langle t_i^{\mathrm{send}} \rangle} \tag{3.10}$$

A graph theoretical estimate of the outflux μ_i^{out} of a node proved to be a serious challenge. Since an analytic approach requires a deeper understanding of the single-node phenomena of network traffic the corresponding explanations will be given later in the course of Section 3.3.

As function of the packet creation rate μ the curves for $\mu_i^{\mathrm{in}}(\mu)$ and $\mu_i^{\mathrm{out}}(\mu)$ of every individual node $i \in \mathcal{N}$ intersect at $\mu = \mu_i^{\mathrm{crit}}$. The critical packet creation rate of the network is given as

$$\mu^{\mathrm{crit}} = \min_{i \in \mathcal{N}} \mu_i^{\mathrm{crit}} \tag{3.11}$$

The estimate of this critical network load allows the determination of the throughput according to equation (3.6).

3.2.2 Active packets and end-to-end time delay

The number of active packets $M(t)$ within a network has already been defined in equation (3.5) as the total number of data packets stored within all of the network's buffer queues at time step t. It is intuitive that this number is closely connected to the network size, the packet creation rate, the routing decisions and the topology of the network graph. As pointed out in Section 3.2.1 a clear sign of a network being in a supercritical state is a linear dependence of the average number of active packets $\langle M(t) \rangle_{\Delta t}$ on time, for reference consult again Figure 3.2.

The end-to-end time delay t_{e2e} of a packet is defined as the number of time steps between the generation at the originating node and the destruction at the final receiver. To reduce the high volatility of the end-to-end time delay measure a temporal average $\langle t_{\mathrm{e2e}}(t) \rangle_{\Delta t}$ over a short time interval Δt prior to time step t is used as a suitable observable of the network performance during a data traffic simulation. Being in the subcritical regime with $\mu < \mu^{\mathrm{crit}}$ the system reaches a steady state, where the number of packets created per time step given by μN must be equal to the number of packets delivered per unit time. Since the average time a

packet spends in the network is $\langle t_{\mathrm{e2e}}(t)\rangle_{\Delta t}$, we can assume that $\langle M(t)\rangle_{\Delta t}/\langle t_{\mathrm{e2e}}(t)\rangle_{\Delta t}$ packets are delivered to their final destination per unit time yielding

$$\frac{\langle M(t)\rangle_{\Delta t}}{\langle t_{\mathrm{e2e}}(t)\rangle_{\Delta t}} = \mu N \tag{3.12}$$

This relation is known in queueing theory as *Little's Law* [8, 22, 61]. Little's Law could be verified for the above stated model of wireless ad hoc communication networks. A visualisation is given in Figure 3.5. For a network of 100 nodes with the constant-P rule ($k^{\mathrm{constP}} = 24$) the number of active packets $\langle M(t)\rangle_{\Delta t}$ averaged over $\Delta t = 100$ time steps as well as the local average of the end-to-end time delay $\langle t_{\mathrm{e2e}}(t)\rangle_{\Delta t}$ over the same interval is shown for two different subcritical packet creation rates. The correspondence of the curves is striking. Due to the fact that with a system size of 100 nodes the critical packet creation rate is in the order of $\mu_{crit} = 0.01$, the product μN in equation (3.12) evaluates to approximately 1 which leads to the observed correspondence in the order of magnitude of the two curves.

The probability density function of the end-to-end time delay $p(t_{\mathrm{e2e}})$ for all packets generated in a generic data traffic simulation with $5{\cdot}10^5$ time steps is shown in Figure 3.6 for different subcritical packet creation rates μ. The simulation was performed on the so called constant-P reference network. This particular network is shown in Figure 3.1 and consists of 100 nodes with constant-P power assignment ($k^{\mathrm{constP}} = 24$). For higher packet creation rates μ the probability distribution is shifted such that there is a higher probability of having higher end-to-end time delays. This is obvious because of the larger number of packets being present in the network that might interact or obstruct each other.

In the log-lin plot of Figure 3.6 it seems promising to suspect a generalised exponential probability density function

$$p_{(a_E,b_E)}(t_{\mathrm{e2e}}) = \frac{1}{b_E} e^{-(t_{\mathrm{e2e}}-a_E)/b_E} \tag{3.13}$$

for lower packet creation rates. A fit of the probability distribution of the end-to-end time delay at a packet creation rate of $\mu = 0.005$ is provided by the green line in in the upper graph of Figure 3.7, that matches the numerically obtained values quite well. Fit parameters are $a_E = 1.509$ and $b_E = 3.066$. A fit with a log-normal probability density function

$$p_{(\mu_L,\sigma_L)}(t_{\mathrm{e2e}}) = \frac{1}{\sigma_L t_{\mathrm{e2e}}\sqrt{2\pi}}\, e^{-(\ln t_{\mathrm{e2e}}-\mu_L)^2/(2\sigma_L^2)} \tag{3.14}$$

as suggested in [56] is less suited. It is also included in the upper graph of Figure 3.7; Parameters are $\mu_L = 1.305$ and $\sigma_L = 0.677$.

The same investigation was repeated for a higher packet creation rate $\mu = 0.0095$ that is much closer to the critical packet creation rate $\mu^{\mathrm{crit}} = 0.0101$ of the

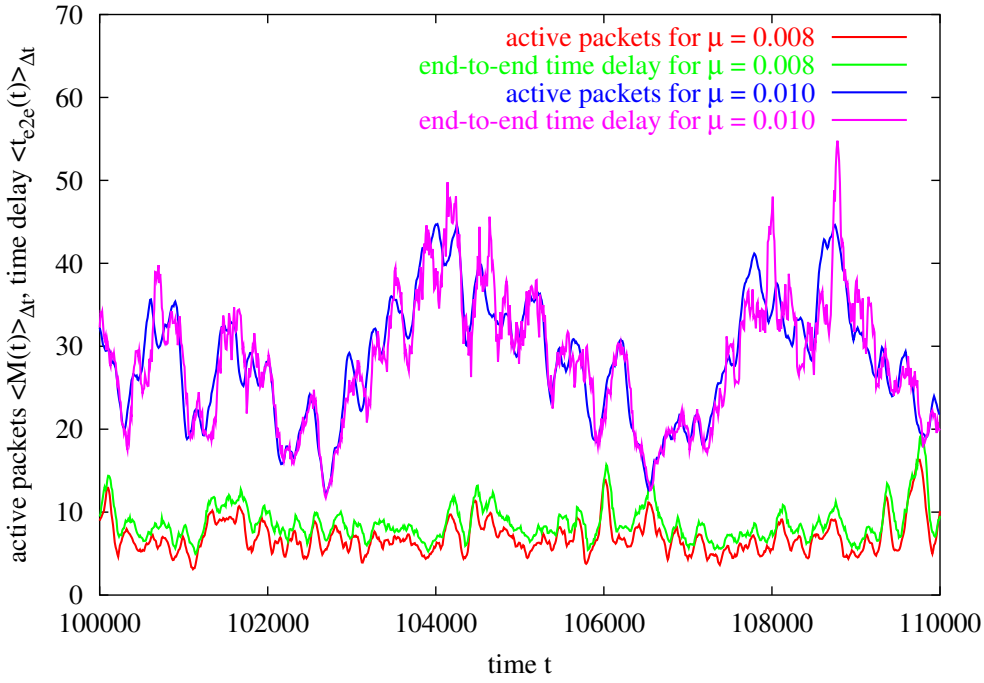

Figure 3.5: Little's Law

For a time interval of 10,000 time steps out of a longer simulation run on a constant-P reference network with 100 nodes, constant-P rule ($k^{\mathrm{constP}} = 24$), the number of active packets $\langle M(t)\rangle_{\Delta t}$ and the end-to-end time delay $\langle t_{\mathrm{e2e}}(t)\rangle_{\Delta t}$ is shown. The averaging interval was set to $\Delta t = 100$. Two different subcritical packet creation rates $\mu = 0.008$ and $\mu = 0.01$ have been employed.

constant-P reference network. The result is shown in the lower graph of Figure 3.7. The fit with an exponential probability density function like in (3.13) fails to provide an adequate description of the numerically obtained data. Fitting parameters are $a_E = -0.669$ and $b_E = 15.360$. Here the fit with the log-normal probability density function of 3.14 outperforms the exponential one, although in the regime of shorter buffer queues the fit is not perfect. Parameters are $\mu_L = 2.504$ and $\sigma_L = 1.060$.

A clear understanding of the log-normal behaviour of the probability distribution of the end-to-end time delay could not be given so far. Although Huberman et al. in [62] suggests an explanation based on a sociological analysis of individual decisions by using the internet, this is not applicable here since generation events of data packets are independent of the actual system state. The log-normal behaviour of the probability distribution of the end-to-end time delay seems to be an inherent feature of data traffic on communication networks [56, 62, 63]. The

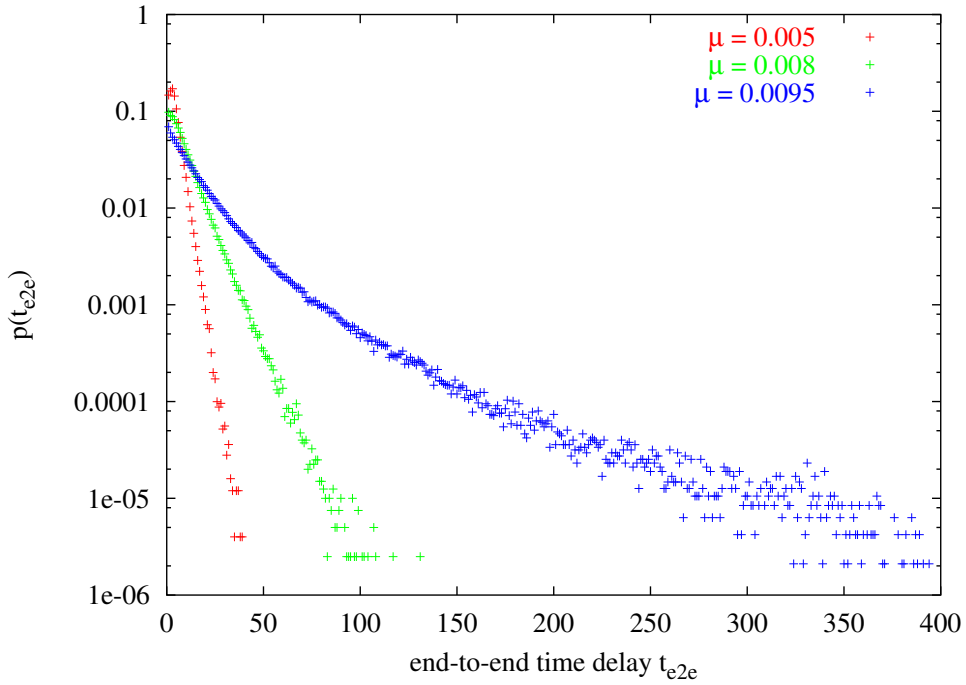

Figure 3.6: Probability distribution of the end-to-end time delay
The probability distribution of the end-to-end time delay $p(t_{\mathrm{e2e}})$ for three different subcritical packet creation rates $\mu = 0.005\,,\ 0.008\,,\ 0.0095$ is shown. The generic data traffic simulation of $5 \cdot 10^5$ time steps was performed on the constant-P reference network that consists of 100 nodes with constant-P power assignment ($k^{\mathrm{constP}} = 24$).

above presented findings are persistent also for significant changes in the packet generation of the generic data traffic model previously introduced. Such aspects of particular user behaviour are discussed in Section 3.6.

3.3 Characterisation of data traffic II: single-node

So far only observables that describe the behaviour of the network as a whole have been studied. But from the graph theoretical discussion of the end-to-end throughput in Section 3.2.1 one already knows about the influence of certain nodes on the overall behaviour of the network. In particular it is one single node that marks the transition from a sub- to a supercritical data traffic regime. A better understanding of a single-node behaviour may thus provide more insight into the dynamics of the whole system.

As presented in the previous section wireless ad hoc communication networks

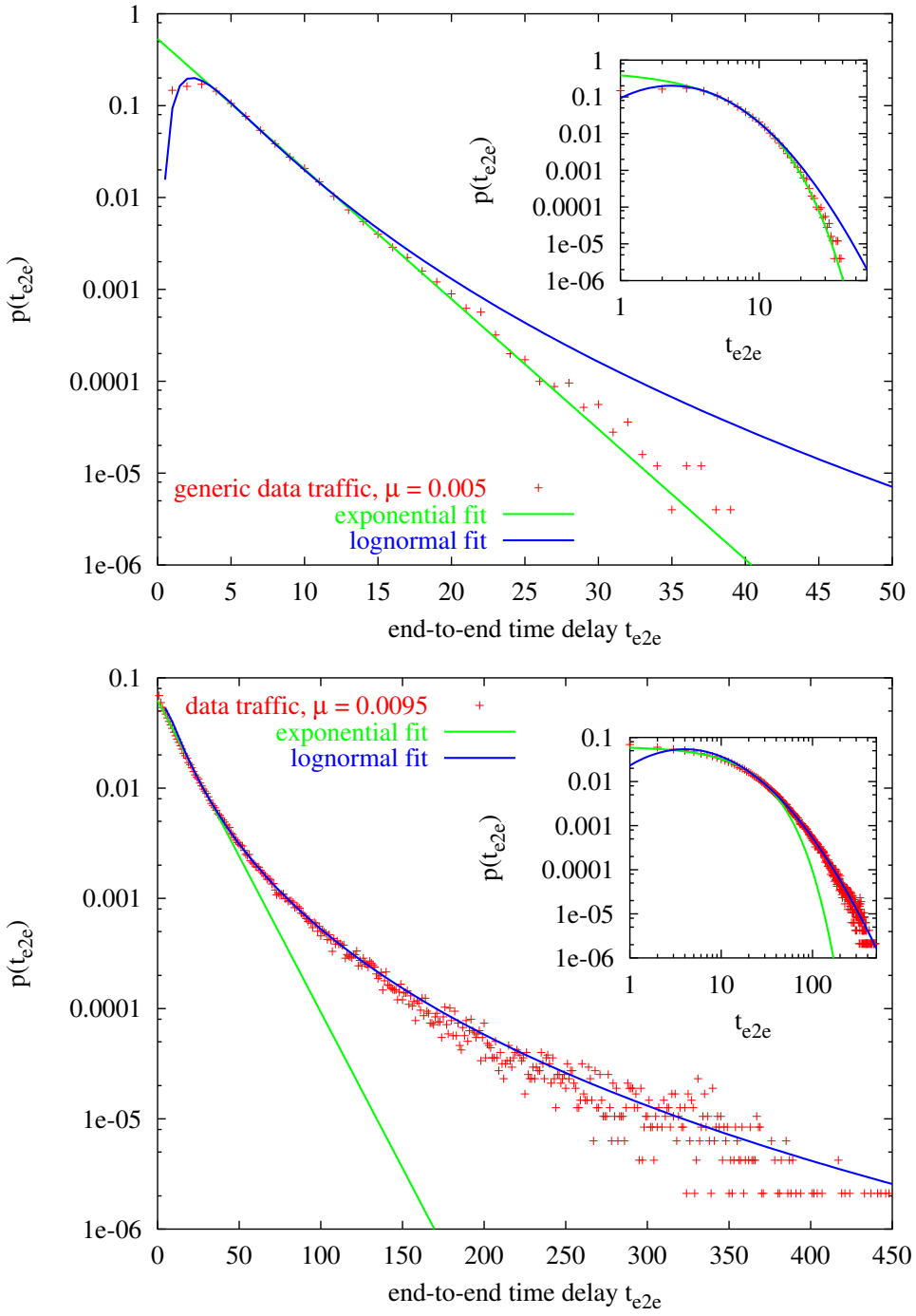

Figure 3.7: Probability distribution of the end-to-end time delay (fitted) The upper graph shows the probability distribution of the end-to-end time delay t_{e2e} for a packet creation rate $\mu = 0.005$ as well as a fit with an exponential and a log-normal probability density function. The lower graph illustrates the same for $\mu = 0.0095$. A detailed explanation of the fits is provided in the text. Parameters are the same as in Figure 3.7. The inset represent the corresponding graphs in a log-log plot.

are subject to congestion if the packet creation rate exceeds a certain value. Congestion does not immediately take over the complete systems but starts out at one single node. Knowing about the behaviour of nodes around the critical state may help to develop strategies to reduce traffic via certain nodes that are objective to congestion and thus keeping the system in a subcritical regime.

From the perspective of queueing theory [22, 61] it is interesting to find out if classical concepts can be applied to wireless communication networks. A possible adaption may allow an analytic description and support a more detailed understanding.

3.3.1 Interarrival and sending time

Among others the routing decisions and the global packet creation rate μ have an impact on the number and time intervals of incoming and outgoing packets to and from a given node. To make further statements about the behaviour of the buffer queue one has to know more about these so called interarrival and sending times, in particular if they follow certain probability distributions. For this thesis the interarrival time t_i^{arrive} of node i is defined as the time between two successive arrivals of data packets that are put at the end of the buffer queue of i. Two things should be remarked here: First, it does not matter if a packet is received via transmission from a neighbouring node $j \in \mathcal{N}_i$ or if they are generated right at node i as long as they are added to the buffer queue of i. Second, arriving packets to node i with i being the final receiver are not counted because they are not handed over to the buffer queue but are deleted from the network.

Figure 3.8 shows the probability distribution of the interarrival times for the most sensible node of the constant-P reference network, introduced in Section 3.1.2, consisting of 100 nodes with constant-P power assignment ($k^{\text{constP}} = 24$) for a subcritical packet creation rate $\mu = 0.008 < \mu_{\text{ref}}^{\text{crit}} = 0.0101$. The generic data traffic simulation covered $5 \cdot 10^5$ time steps. In the log-lin plot the exponential behaviour of the probability distribution can easily be verified which is an indication that the arrival event possibly follows a Poissonian statistic. Because of the discretised time a discrete representation is preferred. This suggests to describe the probability density function in terms of a binomial probability distribution.

Regarding the arrival/non-arrival of packets as a Bernoulli trail a geometric probability density function should be sufficient to describe the statistics [61]. Identifying the inverse of the mean interarrival time $1/\langle t_i^{\text{arrive}} \rangle = \mu_i^{\text{in}}$ as the probability for a new data packet to arrive at i the probabiltity density function for the arrivaltimes $p(t_i^{\text{arrive}})$ is given as

$$p(t_i^{\text{arrive}} = t) = (1 - \mu_i^{\text{in}})^{t-1} \mu_i^{\text{in}} \tag{3.15}$$

This is a representation of a geometric probability density function. It supports the idea that subsequent packet arrival events are completely independent.

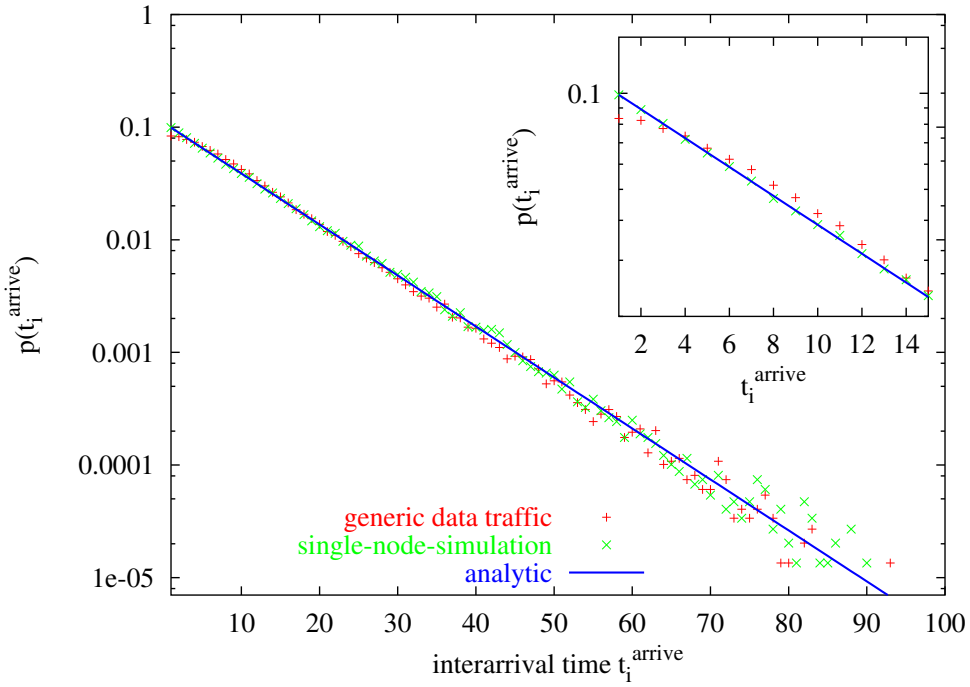

Figure 3.8: Probability distribution of the interarrival time
The probability distribution $p(t_i^{\text{arrive}})$ (red) for the most sensible node of the constant-P reference network is shown in a log-lin plot. The network consists of 100 nodes with constant-P power assignment ($k^{\text{constP}} = 24$). The generic traffic simulation covered $5 \cdot 10^5$ time steps for a subcritical packet creation rate $\mu = 0.008$. Additionally a graphical representation of the geometric probability distribution function (3.15) with $\mu_i^{\text{in}} = 0.099$ is provided (blue) and a probability distribution of the interarrival times reproduced by the single-node simulation (green).

The corresponding graphical representation of equation (3.15) is also included in Figure 3.8 (blue line) and verifies the shown approach. Note that the only input for this analytical description is the rate of incoming packets $\mu_i^{\text{in}} = 1/\langle t_i^{\text{arrive}} \rangle$ to node i. For the shown case $\mu_i^{\text{in}} = 0.099$ was measured in the generic data traffic.

Without showing it should be remarked that higher packet creation rates cause a shift in the probability distribution towards lower interarrival times. The exponential behaviour is maintained.

A numerical investigation reveals that equation (3.15) is applicable for all nodes $i \in \mathcal{N}$. It is not restricted to the most sensible node.

The sending time t_i^{send} of node i is defined as the time between two successive sending events from node i to any of its neighbouring nodes $j \in \mathcal{N}_i$ given that

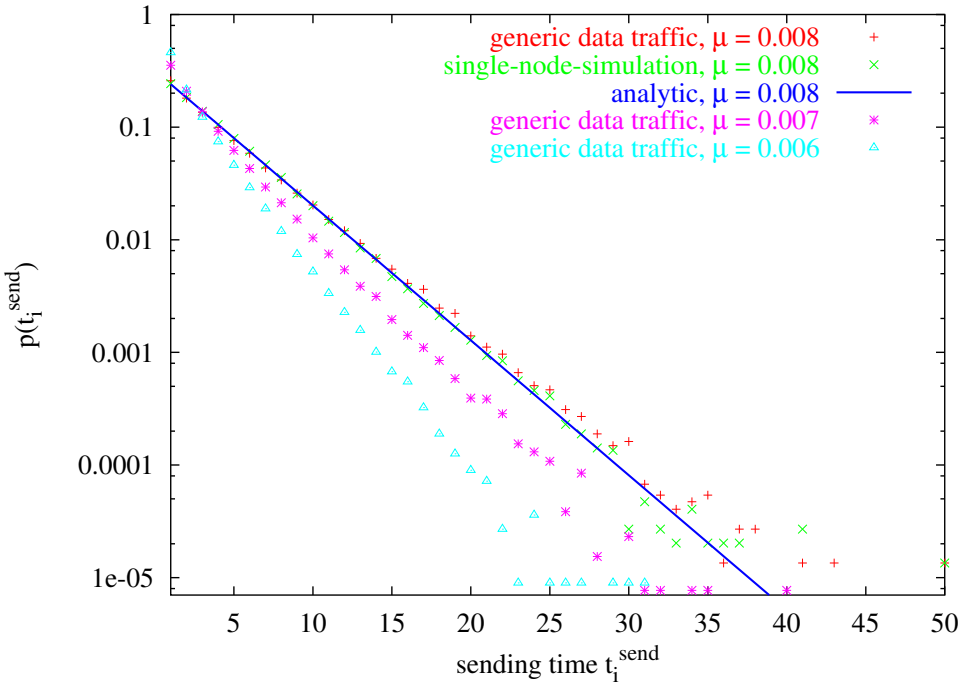

Figure 3.9: Probability distribution of the sending time $p(t_i^{\text{send}})$
Parameter settings are the same as in Figure 3.8. In addition to the probability distribution of the sending time $p(t_i^{\text{send}})$ (red for $\mu = 0.008$, pink for $\mu = 0.007$, light blue for $\mu = 0.006$) the corresponding analytical curve in analogy to equation (3.15) with $\mu_i^{\text{out}} = 0.241$ (blue) and a probability distribution reproduced by the single-node simulation (green) is included for the higher packet creation rate.

the buffer queue is not empty. This restriction incorporates the fact that only nodes with non-zero buffer have a desire to send. In that respect the counter for the sending time is reset either with the sending of a packet or with the arrival of a packet to an empty buffer queue. Typical probability distributions for the sending time $p(t_i^{\text{send}})$ of the most sensible node of the constant-P reference network at $\mu = 0.006, 0.007, 0.008 < \mu_{\text{ref}}^{\text{crit}} = 0.0101$ are shown in Figure 3.9. The network consists of 100 nodes with constant-P power assignment ($P = 24P^{\text{norm}}$). The generic data traffic simulation covered $5 \cdot 10^5$ time steps. A higher packet creation rate μ causes a shift of the probability distribution towards larger sending times.

The same argumentation as for $p(t_i^{\text{arrive}})$ with the use of the outflux rate $\mu_i^{\text{out}} = 1/\langle t_i^{\text{send}} \rangle$ instead of μ_i^{in} in equation (3.15) holds for the analytical description of this graph. The corresponding fit (blue) with $\mu_i^{\text{out}} = 0.241$ is included in Figure 3.9 for the higher packet creation rate $\mu = 0.008$.

An extended discussion of the above stated results will be given in Section 3.3.3.

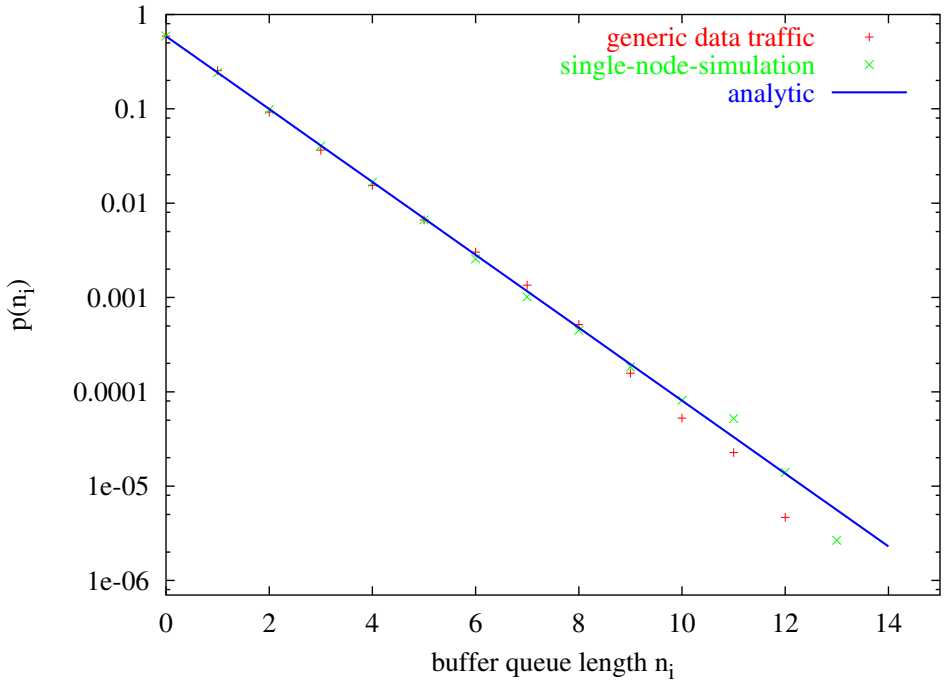

Figure 3.10: Probability distribution of the buffer queue length
In addition to the probability distribution of the buffer queue length $p(n_i)$ (red) at $\mu = 0.008$ the corresponding analytical curve given by equation (3.20) with $\mu_i^{\text{in}} = 0.099$ and $\mu_i^{\text{out}} = 0.241$ (blue) and a probability distribution reproduced by the single-node simulation (green) is included. Parameter settings are the same as in Figure 3.8.

3.3.2 Buffer queue length

The in- and outflux of data packets to a node i is directly connected to its buffer content described by the number of packets n_i stored in the particular buffer queue, also referred to as buffer queue length. The probability distribution $p(n_i)$ of the number of packets in i's buffer queue is shown in Figure 3.10 for the most sensible node of the constant-P reference network. 100 nodes with constant-P power assignment ($k^{\text{constP}} = 24$) constitute the particular network graph. The simulation run spans $5 \cdot 10^5$ time steps. The packet creation rate is set to $\mu = 0.008 < \mu_{\text{ref}}^{\text{crit}} = 0.0101$.

In the log-lin plot in Figure 3.10 one observes a similar exponential behaviour for the probability distribution of the buffer queue length $p(n_i)$ as for the distribution of interarrival and sending times shown in Figures 3.8 and 3.9. For completeness it should be mentioned that higher packet creation rates μ cause a shift in the corresponding probability distribution towards larger buffer queue lengths. A

graph theoretical understanding of the probability distribution can be aimed for by using the above stated results.

One can express the static probability of a node i to have n_i packets in its buffer queue as a function of its possible previous states and the corresponding transition probabilities, represented by the interarrival and sending rate μ_i^{in} and μ_i^{out},

$$p(n_i, t+1) = \mu_i^{\text{in}} p(n_i - 1, t) + \mu_i^{\text{out}} p(n_i + 1, t) + (1 - \mu_i^{\text{in}} - \mu_i^{\text{out}}) p(n_i, t) \quad (3.16)$$

which holds for $n_i > 0$. Such an equation is known as a rate equation. Reducing to the stationary limit $p(n_i, t+1) = p(n_i, t)$ it can be rewritten as

$$(\mu_i^{\text{in}} + \mu_i^{\text{out}}) p(n_i) = \mu_i^{\text{in}} p(n_i - 1) + \mu_i^{\text{out}} p(n_i + 1) \quad (3.17)$$

In the special case $n_i = 0$ the equation above reduces to

$$\mu_i^{\text{in}} p(0) = \mu_i^{\text{out}} p(1) \quad (3.18)$$

due to the fact, that no packets can be removed from an empty queue. Employing this the recursive formulation in equation (3.17) can be transformed into an expression for $p(n_i)$

$$p(n_i) = \left(\frac{\mu_i^{\text{in}}}{\mu_i^{\text{out}}} \right)^{n_i} p(0) \quad (3.19)$$

Normalising this to 1 one comes up with a full description of the probability density function for the length of the buffer queue

$$p(n_i) = \left(\frac{\mu_i^{\text{in}}}{\mu_i^{\text{out}}} \right)^{n_i} \left(1 - \frac{\mu_i^{\text{in}}}{\mu_i^{\text{out}}} \right) \quad (3.20)$$

This probability density function is also shown in Figure 3.10 (blue line) only taking $\mu_i^{\text{in}} = \frac{1}{\langle t_i^{\text{arrive}} \rangle} = 0.099$ and $\mu_i^{\text{out}} = \frac{1}{\langle t_i^{\text{send}} \rangle} = 0.241$ for the most sensible node at $\mu = 0.008$ as its inputs.

The understanding of the probability density function of the buffer queue length $p(n_i)$ allows an analytical treatment of the number of active packets of a network M and the end-to-end time delay t_{e2e}. Using equation (3.20) one can calculate the mean buffer queue length of node i:

$$\langle n_i \rangle = \sum_{n_i=0}^{\infty} n_i p(n_i) = \frac{\frac{\mu_i^{\text{in}}}{\mu_i^{\text{out}}}}{1 - \frac{\mu_i^{\text{in}}}{\mu_i^{\text{out}}}} \quad (3.21)$$

The mean number of active packets is now given in analogy to (3.5) as

$$\langle M(t) \rangle = \sum_{i \in \mathcal{N}} \langle n_i(t) \rangle = \sum_{i \in \mathcal{N}} \frac{\frac{\mu_i^{\text{in}}}{\mu_i^{\text{out}}}}{1 - \frac{\mu_i^{\text{in}}}{\mu_i^{\text{out}}}} \quad . \quad (3.22)$$

Employing equations (3.9) and (3.10) the ratio of in- and outflux can be expressed as $\mu_i^{\text{in}}/\mu_i^{\text{out}} = \mu(N-1)B_i^{\text{node}}\langle t_i^{\text{send}}\rangle$.

One can now study two limiting cases of (3.22). For very small packet creation rates $\mu \to 0$ the sending time $t_i^{\text{send}} \to 1$ since almost no other packets exist in the network that could cause other transmissions and MAC-block node i. The average number of active packets thus becomes

$$\lim_{\mu\to 0}\langle M(t)\rangle = \sum_{i\in\mathcal{N}} \frac{\mu_i^{\text{in}}}{\mu_i^{\text{out}}} = \mu(N-1)\sum_{i\in\mathcal{N}} B_i^{\text{node}} = \mu\, N\, D \quad . \tag{3.23}$$

For the last step the sum rule (3.3) was used. Applying Little's Law (3.12) this can be transformed into an expression for the end-to-end time delay

$$\lim_{\mu\to 0}\langle t_{\text{e2e}}\rangle = D \quad . \tag{3.24}$$

This finding is intuitive since for small traffic load almost no interactions between data packets occur. It is the diameter D of the network that determines the average end-to-end time delay.

In the other case where $\mu \to \mu^{\text{crit}}$ the ratio $\mu_i^{\text{in}}/\mu_i^{\text{out}} \to 1^-$ for the most sensible node; see again Section 3.2.1 for reference. It is precisely this node that governs the number of active packets since the denominator in the sum of equation (3.22) goes to 0. Thus the sum breaks down and it follows

$$\begin{aligned}\lim_{\mu\to\mu^{\text{crit}}}\langle M(t)\rangle &= \frac{1}{1-\frac{\mu_i^{\text{in}}}{\mu_i^{\text{out}}}} \\ &= \frac{1}{(N-1)B_i^{\text{node}}}[\mu^{\text{crit}}\,\langle t_i^{\text{send}}(\mu^{\text{crit}})\rangle - \mu\,\langle t_i^{\text{send}}(\mu)\rangle]^{-1} \quad .\end{aligned} \tag{3.25}$$

Since $\langle t_i^{\text{send}}(\mu)\rangle$ is limited by the number of neighbours of i that compete with it for sending permission it holds

$$\lim_{\mu\to\mu^{\text{crit}}}\langle M(t)\rangle \sim [\mu^{\text{crit}} - \mu]^{-1} \quad . \tag{3.26}$$

The number of active packets goes to infinity for $\mu \to \mu^{\text{crit}}$. The critical exponent is $\xi = -1$ [64]. The same critical exponents $\xi = -1$ holds for the end-to-end time delay which becomes obvious by applying Little's Law (3.12).

3.3.3 Single-node-simulation

In order to verify the above drawn picture of a single-node behaviour within a wireless ad hoc communication network a further simple simulation was set up that allows to independently check the previously obtained results. The so called

single-node-simulation is based on a Poissonian input and output of data packets to a solitary existing node. Corresponding probabilities are $p^{\text{in}} = \mu_i^{\text{in}}$ and $p^{\text{out}} = \mu_i^{\text{out}}$, respectively. An output can only be realised if at least one packet is present in the buffer queue. Running this model in discretised time analogous to the generic data traffic where only a packet arrives, departs or nothing happens at each time step, it allows to gain equivalent data about the probability distributions of the buffer queue length, the interarrival and sending times.

Interactions and correlations between parts and subparts of network play a significant role in its specific dynamics. These correlations are completely neglected in the single-node-simulation since only one solitary node exists in the artificial environment. A study of similarities and differences with the generic data traffic simulation allows to find out the limits of the simplified approach and to learn about the possible reasons for occurring differences.

The results of the single-node-simulation are included in Figures 3.8 - 3.10 where the probability distributions of the interarrival time, sending time and of the buffer queue length is given in green. As in the generic data traffic $5 \cdot 10^5$ time steps were covered. As input parameter to the simulation only $p^{\text{in}} = \mu_i^{\text{in}} = \frac{1}{\langle t_i^{\text{arrive}} \rangle} = 0.099$ and $p^{\text{out}} = \mu_i^{\text{out}} = \frac{1}{\langle t_i^{\text{send}} \rangle} = 0.241$ were used as measured in the generic data traffic simulation for the most sensible node at $\mu = 0.008$.

The perfect correspondence of the single-node-simulation with the analytically obtained result in the previous sections is expected since both approaches are based on the assumptions of a strictly Poissonian in- and output of packets to and from the node. The advantage of the single-node-simulation compared to the analytical results is the accessibility of further, harder-to-describe observables.

On a very first look the results from the generic data traffic fit quite well to the data obtained from the single-node-simulation and from the analytic approach neglecting the statistical fluctuations for higher interarrival/sending times and buffer queue lengths. But a closer look is in order here. Small deviations occur for example for the probability distributions of the interarrival time. For very small values of $\langle t_i^{\text{arrive}} \rangle$ the generic data traffic curve lies below the reference, at around $t_i^{\text{arrive}} = 10$ it is above. See also the inset in Figure 3.8 that emphasises that part of the plot.

In the probability distribution of the buffer queue length slight differences are visible, too. But looking at a larger number of similar graphs for other nodes within this particular network realisation and other network realisations as well no unitary trend is observable.

To be more precise, the simplified picture of a node with strictly Poissonian input and output to a node does only hold up to a certain point of precision. The described deviations are obviously caused by effects not taken into account in the single-node-simulation nor in the analytical approach. For a principle understanding of the most important features of a single node's behaviour the approaches

shown in the previous sections are sufficient. But one has to keep in mind that effects based on complex interactions in the network structure complicate this picture.

3.3.4 Single-node temporal correlations

So far only one-point statistics were used in the investigation of simulated data traffic in wireless ad hoc communication networks. In order to gain more insight in the dynamical behaviour of nodes single-node temporal correlations were analysed.

The observable studied in more detail is the single-node temporal correlation of the buffer queue length $n_i(t)$ of node i, named $r_i(\Delta t)$, which is defined as

$$r_i(\Delta t) = \frac{\langle n_i(t+\Delta t) n_i(t) \rangle}{\langle n_i(t) \rangle^2} \tag{3.27}$$

This observable may provide results about the status of the buffer queue n_i at different points in time separated by an interval Δt. Figure 3.11 shows a characteristic plot of $r_i(\Delta t)$ for the buffer queue of the most sensible node of the constant-P reference network for different packet creation rates μ. The simulation run spans $5 \cdot 10^5$ time steps. The critical packet creation rate of this particular network was determined to be $\mu^{\text{crit}} = 0.0101$. The convergence of the single-node temporal correlation $r_i(\Delta t)$ to 1 indicates that no correlations exist between the buffer queues $n_i(t)$ and $n_i(t+\Delta t)$ of this node after a certain interval Δt. This convergence to 1 is shifted to higher correlation times Δt as the packet creation rate grows. Such long range correlations indicate that a buffer queue state of a node is not independent of its former states. Especially higher packet creation rates that result in a higher input to a node and a more competitive output from a node cause longer buffer queues. Packets once inserted in the buffer queue of a node are kept there for much longer time scales which explains the shift of the decorrelation time to higher Δt.

To be able to compare the shape of the curves and to find possible similarities one may assume them to be of a more fundamental form:

$$r_i(\Delta t) = \left(f\left(\frac{\Delta t}{T} \right) \right)^{\zeta} \tag{3.28}$$

where the exponent ζ accounts for a scaling of the y-axis and T for a scaling of the x-axis. Following equation (3.27) one can impose the condition

$$r_i(\Delta t = 0) = \frac{\langle n_i^2(t) \rangle}{\langle n_i(t) \rangle^2} \tag{3.29}$$

and comes up with an expression for the scaling exponent ζ.

$$\zeta = \frac{\ln(\frac{\langle n_i^2(t) \rangle}{\langle n_i(t) \rangle^2})}{\ln f(0)} \tag{3.30}$$

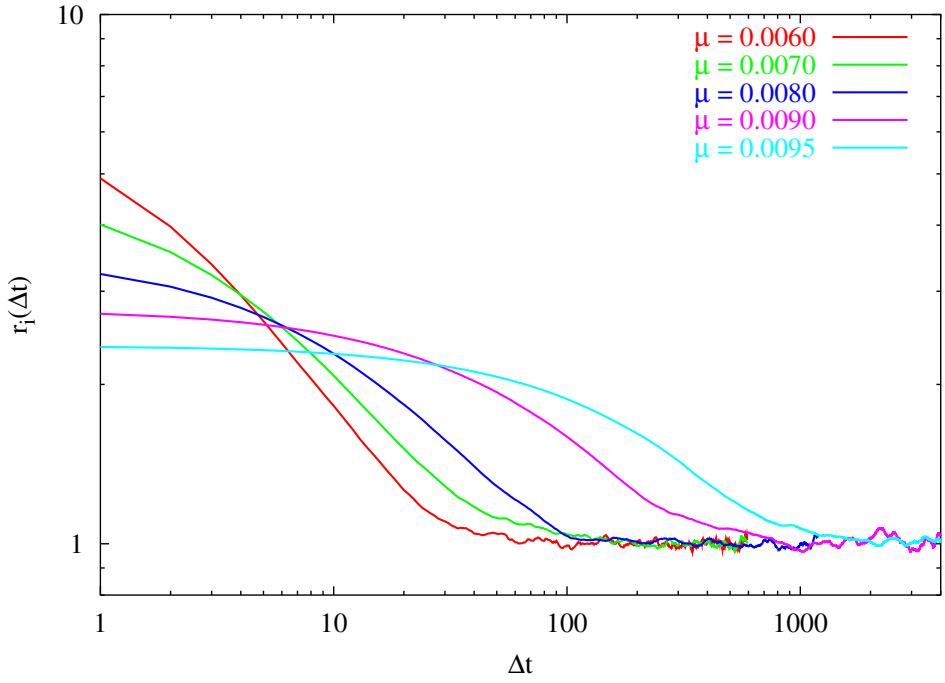

Figure 3.11: Single-node temporal correlations
The single-node temporal correlations $r_i(\Delta t)$ for the most sensible node of the constant-P reference network are shown. The network graphs consists of 100 nodes with constant-P rule ($k^{\text{constP}} = 24$). The simulation covered $5 \cdot 10^5$ time steps. Different colours correspond to different packets creation rates $\mu = 0.006$ (red), $\mu = 0.007$ (green), $\mu = 0.008$ (blue), $\mu = 0.009$ (pink), $\mu = 0.0095$ (light blue).

Setting the free parameter $f(0) = 2$ one fixes the transition through the ordinate. This causes a scaling of the curves of Figure 3.11 as shown in the upper graph of Figure 3.12 simply by plotting $(r_i(\Delta t))^{1/\zeta}$ against Δt. The figure clearly indicates the direct relation between the normalised second moment $\frac{\langle n_i^2(t)\rangle}{\langle n_i(t)\rangle^2}$ of the buffer queue and the respective y-scaling of the single-node temporal correlations.

When changing the values for T in equation (3.28) the corresponding curves are shifted along the x-axis. The similar shape of the curves for different packet creation rates μ almost pleads for a curve collapse onto one. The lower graph in Figure 3.12 shows the result, where a best fit was obtained "by eye". The almost perfect correspondence of the curves indicates similar mechanisms underlying the single-node temporal correlation $r_i(\Delta t)$ for a wide range of packet creation rates.

It is clear that the fitted μ-dependend parameter $T(\mu)$ has to correspond to some characteristic time of a node. No suitable observable that describes this

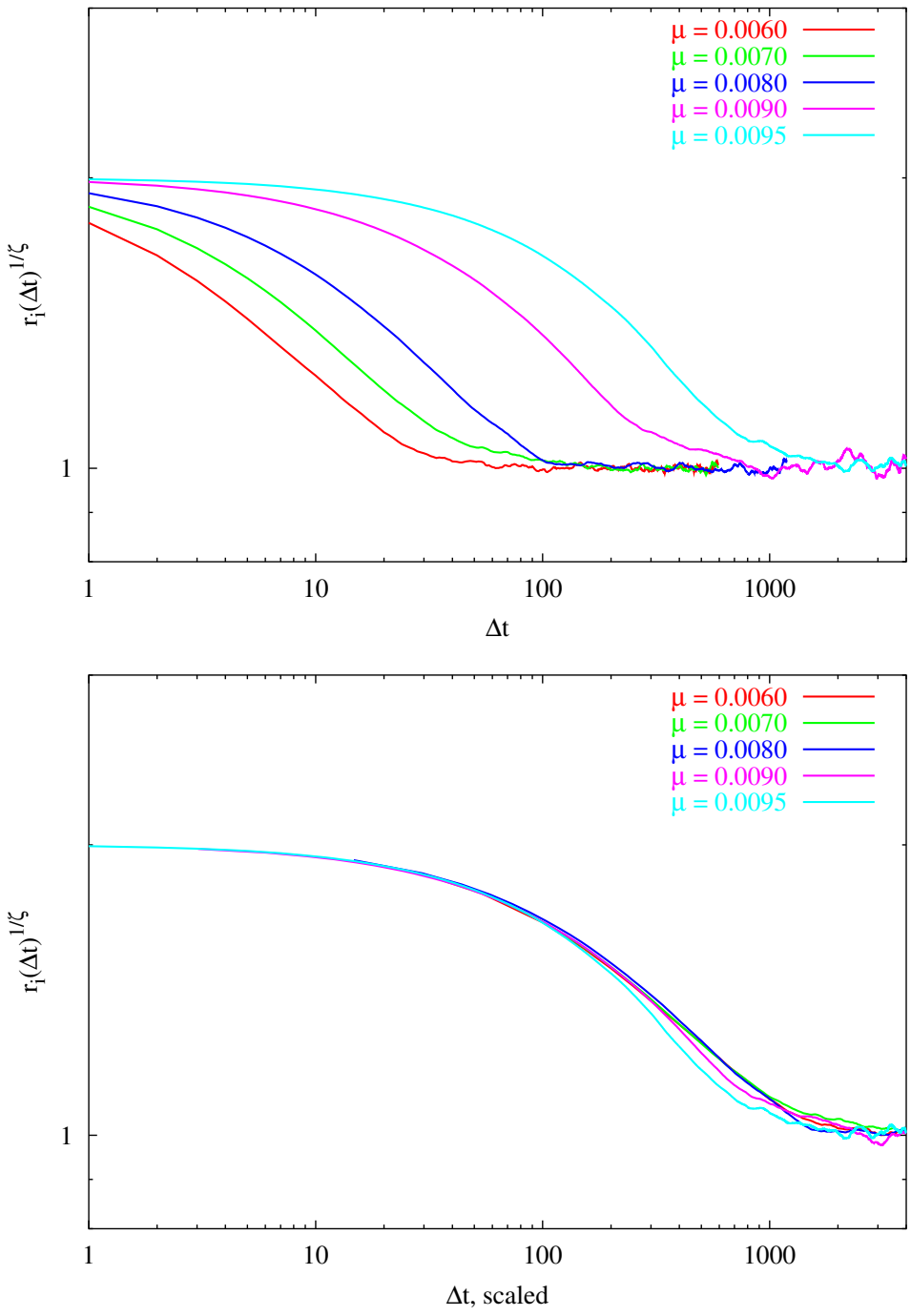

Figure 3.12: Scaling of the single-node temporal correlations
The single-node temporal correlations $r_i(\Delta t)$ for the most sensible node of the constant-P reference network are shown. Parameters are the same as in Figure 3.11. In the upper graph a scaling according to equations (3.28) to (3.30) was done. In the lower graph the curves were collapsed by rescaling the parameter T in equation (3.28) "by eye".

scaling sufficiently well has been identified so far. Knowing about this problem and the presented similarity of shape for the single-node temporal correlations $r_i(\Delta t)$ one desires a more fundamental understanding of this issue.

Describing the state of a queue at a time step $t+1$ analogous to (3.16) as

$$p(n_i, t+1) = \mu_i^{\text{in}} p(n_i - 1, t) + \mu_i^{\text{out}} p(n_i + 1, t) + (1 - \mu_i^{\text{in}} - \mu_i^{\text{out}}) p(n_i, t) \quad (3.31)$$

with μ_i^{in} and μ_i^{out} being the in- and outflux to node i assuming Poissonian distributed events. Suppressing the indices i this equation can be rewritten as

$$\begin{aligned} p(n, t+1) - p(n, t) = {} & \frac{(\mu^{\text{out}} - \mu^{\text{in}})}{2} (p(n+1, t) - p(n-1, t)) \\ & + \frac{(\mu^{\text{out}} + \mu^{\text{in}})}{2} (p(n+1, t) - 2p(n, t) + p(n-1, t)) \quad . \end{aligned} \quad (3.32)$$

In the limit of long buffer queues $n \gg 1$ one introduces $1 = \Delta t = \Delta n$. A switch to a continuous representation of this equation is possible by applying the limit $\Delta t \to 0$ and $\Delta n \to 0$. Replacing n with x one comes up with

$$\frac{\partial p(x,t)}{\partial t} = -(\mu^{\text{in}} - \mu^{\text{out}}) \frac{\partial p(x,t)}{\partial x} + \frac{(\mu^{\text{out}} + \mu^{\text{in}})}{2} \frac{\partial^2 p(x,t)}{\partial x^2} \quad (3.33)$$

which is a representation of a Fokker-Planck equation. $(\mu^{\text{in}} - \mu^{\text{out}}) = -\gamma$, $(\gamma > 0)$ can be identified as the time independent drift coefficient and $(\frac{\mu^{\text{out}} + \mu^{\text{in}}}{2}) = \frac{D}{2}$ is referred to as the diffusion constant. In that respect the Fokker-Planck equation describes the temporal evolution of a probability density function, in this case of the distribution of the buffer length $n = x$ with a constant drift and an additional diffusive element. A restriction to non-negative buffer queues imposes a first boundary condition. For reasons of normalisation the probability of having an infinite buffer queue has to vanish, too, which is a second boundary condition. Writing equation (3.33) as a continuity equation one defines a probability current

$$J(x,t) = -\gamma p(x,t) - \frac{D}{2} \frac{\partial p(x,t)}{\partial x} \quad . \quad (3.34)$$

The boundary conditions can now be expressed in terms of this probability current as $J(x = 0, t) = 0$ and $J(x = \infty, t) = 0$. These conditions fix the free integration constant for the stationary solution

$$p_s(x) = C_s e^{-\frac{2\gamma}{D} x} \quad (3.35)$$

with

$$C_s = \frac{2\gamma}{D} (1 - e^{-\frac{2\gamma}{D} L})^{-1} \quad . \quad (3.36)$$

L is an upper integration constant which has to go to infinity at an appropriate point.

To solve the time dependent Fokker-Planck equation the ansatz of Gardiner [65] is suitable

$$p(x,t) = p_s(x)q(x,t) \tag{3.37}$$

which is expanded in eigenfunctions

$$q(x,t) = q_\lambda(x)\, e^{-\lambda t} \quad . \tag{3.38}$$

Using this ansatz in (3.33) the eigenfunctions have to be of the form

$$q_\lambda(x) = C_1 e^{\alpha_1 x} + C_2 e^{\alpha_2 x} \tag{3.39}$$

with

$$\alpha_{1/2} = \frac{\gamma}{D} \pm \left[\frac{\gamma^2}{D^2} - \frac{2\lambda}{D}\right]^{\frac{1}{2}} \quad . \tag{3.40}$$

The current from equation (3.34) can be expressed in terms of the ansatz in equation (3.37) as

$$J(x,t) = -\frac{D}{2} p_s(x) \frac{\partial q(x,t)}{\partial x} \tag{3.41}$$

which allows the incorporation of the boundary conditions $J(x=0,t)=0$ and $J(x=L,t)_{L\to\infty}=0$. This results in a condition for the relation of α_1 and α_2

$$\alpha_1\alpha_2(e^{\alpha_2 L} - e^{\alpha_1 L}) = 0 \quad . \tag{3.42}$$

Two cases can now be distinguished. For $\alpha_1\alpha_2 = 0$ it follows the eigenvalue $\lambda = 0$. Because of the boundary conditions the corresponding eigenfunction has to be fixed to $q_0(x) = 1$. In the case $e^{\alpha_2 L} = e^{\alpha_1 L}$ the eigenvalue λ is a function of an integer m

$$\lambda_m = \frac{D}{2L^2}\left[\left(\frac{\gamma L}{D}\right)^2 + (m\pi)^2\right] \tag{3.43}$$

which causes for the exponents

$$\alpha_{1/2} = \frac{\gamma}{D} \pm i\,\frac{m\pi}{L} \quad . \tag{3.44}$$

One can now derive the expressions for $q_m(x) = q_{\lambda_m}(x)$. Using the normalisation [65]

$$\int_0^L dx\, p_s(x)\, q_m(x)\, q_{m'}(x) = \delta_{mm'} \tag{3.45}$$

the eigenfunctions and their corresponding eigenvalues are given by

$$\begin{aligned} q_0 &= 1\,, \\ q_m &= C_m e^{\frac{\gamma}{D}x}\left[\frac{\gamma}{D}\sin\left(\frac{m\pi}{L}x\right) - \frac{m\pi}{L}\cos\left(\frac{m\pi}{L}x\right)\right] \end{aligned} \tag{3.46}$$

with

$$C_m^2 = \frac{DL}{\gamma}\left[\frac{1-e^{-\frac{2\gamma L}{D}}}{(\frac{\gamma L}{D})^2+(m\pi)^2}\right] \quad , \quad m = 1,2,3,... \tag{3.47}$$

The solution to the Fokker-Planck equation (3.33) can be written in terms of eigenfunctions:

$$p(x,t) = \sum_m A_m\, p_s(x)\, q_m(x)\, e^{-\lambda_m t} \tag{3.48}$$

with

$$A_m = \int_0^L dx\, q_m(x)\, p(x,0) \quad . \tag{3.49}$$

The conditional probability $p(x,t|x_0,0)$ is given by the initial condition $p(x,0|x_0,0) = \delta(x-x_0)$. It follows

$$A_m = \int_0^L dx\, q_m(x)\, \delta(x-x_0) = q_m(x_0) \tag{3.50}$$

and

$$p(x,t|x_0,0) = \sum_m p_s(x)\, q_m(x)\, q_m(x_0)\, e^{-\lambda_m t} \quad . \tag{3.51}$$

Following the calculations in Gardiner [65] a single-node temporal correlation is given by

$$\begin{aligned}\langle x(t)\, x(0)\rangle &= \int dx \int dx_0\, x\, x_0\, p(x,t|x_0,0)\, p_s(x_0) \\ &= \sum_m \left(\int dx\, x\, p_s(x)\, q_m(x)\right)^2 e^{-\lambda_m t} \quad .\end{aligned} \tag{3.52}$$

In order to let $L \to \infty$ the substitution $x = \frac{Dm\pi}{\gamma L}$ is introduced that replaces the sum with a suitable integration

$$\sum_m (...) \to \int_0^\infty \frac{\gamma L}{D\pi}\, dx... \tag{3.53}$$

In the special case $t = 0$ the single-node temporal correlation evaluates to

$$\langle x^2(0)\rangle = 2\langle x(0)\rangle^2 \tag{3.54}$$

which makes the plot of the single-node temporal correlation to start at a value of 2 on the ordinate. This matches the observations in Figure 3.11. The further the packet creation rate μ approaches its critical value μ^{crit} the better $\frac{\langle x^2(0)\rangle}{\langle x(0)\rangle^2}$ converges to 2. This is consistent with the continuum assumption, the Fokker-Planck equation is based on. To justify a transition from discrete packets to a continuous

queue, the mean buffer queue length has to be high, $\frac{\Delta n}{\langle n \rangle} \ll 1$, respectively. This is the case for $\mu \to \mu^{\text{crit}}$.

In the specific case equation (3.52) translates to

$$\langle x(t)\, x(0) \rangle = \frac{1}{4}\left(\frac{D}{\gamma}\right)^2 + \frac{4}{\pi}\left(\frac{D}{\gamma}\right)^2 e^{-\frac{\gamma^2}{2D}} \int_0^\infty dy\, \frac{y^2}{(1+y^2)^3}\, e^{-\frac{\gamma^2}{2D} t y^2} \quad . \tag{3.55}$$

The later integral is a representation of a confluent hypergeometric function [66] which simplifies the result to

$$\begin{aligned} \langle x(t)\, x(0) \rangle &= \frac{1}{4}\left(\frac{D}{\gamma}\right)^2 + \frac{4}{\pi}\left(\frac{D}{\gamma}\right)^2 e^{-\frac{\gamma^2}{2D}} \left[\frac{1}{2}\,\Gamma\left(\frac{3}{2}\right) U\left(\frac{3}{2}, -\frac{1}{2}, \frac{\gamma^2}{2D} t\right)\right] \\ &= \frac{1}{4}\left(\frac{D}{\gamma}\right)^2 + \frac{4}{\pi}\left(\frac{D}{\gamma}\right)^2 e^{-\frac{\gamma^2}{2D}} \left[\frac{\pi}{16} M\left(\frac{3}{2}, -\frac{1}{2}, \frac{\gamma^2}{2D} t\right)\right. \\ &\quad \left. + \frac{3}{2}\sqrt{\pi}\left(\frac{\gamma^2}{2D} t\right)^2 M\left(3, \frac{5}{2}, \frac{\gamma^2}{2D} t\right)\right] \quad . \end{aligned} \tag{3.56}$$

This expression can be seen as a fully analytical description of the temporal correlation of the buffer queue length at two points in time separated by $\Delta t = t$. For three different packet creation rates μ the y-scaled single-node temporal correlation of the most sensible node of the constant-P reference network is shown in Figure 3.13. Therefore equation (3.56) was normalised analogous to equation (3.27). Parameters for the plot are $\gamma = -(\mu_i^{\text{in}}(\mu) - \mu_i^{\text{out}}(\mu))$ and $D = (\mu_i^{\text{out}}(\mu) + \mu_i^{\text{in}}(\mu))$ with $\mu_i^{\text{in}}(\mu)$ and $\mu_i^{\text{out}}(\mu)$ obtained from the generic data traffic simulations for different packet creation rates μ.

The plots of the single-node temporal correlation obtained from generic data traffic (as already shown in Figure 3.11) and a single-node temporal correlation from the single-node-simulation introduced in Section 3.3.3 are also included in Figure 3.13. The later also using the corresponding $\mu_i^{\text{in}}(\mu)$ and $\mu_i^{\text{out}}(\mu)$ as input parameters that have been measured in the generic data traffic simulation for the particular packet creation rate μ. Both simulations span $5 \cdot 10^5$ time steps.

Two important results can be derived from this plot: First, for packet creation rates $\mu \to \mu^{\text{crit}}$ the decline of the correlation functions is shifted to higher correlation times. This result holds for the data driven single-node temporal correlation as well as for the analytically obtained result and the single-node-simulation. In that respect the analytic approach reveals the same results which indicates that the x-scaling is an inherent feature of the simple assumptions made about the behaviour of the buffer queue.

The reason for differences between the analytical results and the single-node-simulation has to be seen in the transition from the discrete description of the buffer queue to a continuous representation. This transition is only justified for longer buffer queues which is the case for large packet creation rates $\mu \to \mu^{\text{crit}}$.

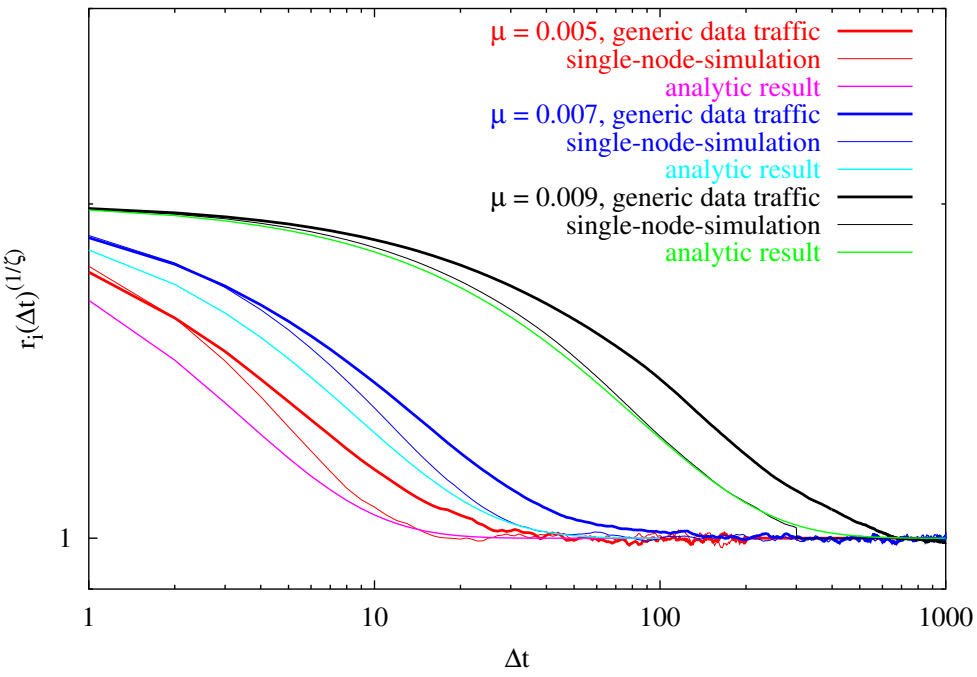

Figure 3.13: Analytic and simulated single-node temporal correlations
The single-node temporal correlations $r_i(\Delta t)^{1/\zeta}$ for the most sensible node of the constant-P reference network are shown. The network graph consists of 100 nodes with constant-P rule ($k^{\mathrm{constP}} = 24$). The simulations covered $5 \cdot 10^5$ time steps. Different colours correspond to different packets creation rates. The thick lines are the results from the generic data traffic, thin lines are obtained from the single-node-simulation. Analytic results are according to equation (3.55).

This is verified in Figure 3.13. For $\mu = 0.009$ the curves of the single-node-simulation and the analytic description almost collaps. The differences are more pronounced for smaller μ due to the discrepancy between the discrete and continuous handling of the buffer queue.

Second, clear differences between the generic result and the analytic solution are visible. It can be assumed that these differences occur because of correlations between neighbouring but also further distant nodes. Such phenomena have to be taken into account especially in the next section where the focus is on a deeper understanding of the mean sending times of the nodes in $\mathcal{N}$.

It has to be pointed out that the theoretical approach shown above as well as the single-node-simulation is based on an ideal picture of just an isolated node assuming the arrival and sending of data packets to follow a pure Poissonian process. No correlations and interactions between nodes have been taken into account. But

several signs like the differences in the probability density functions of the interarrival and sending times mentioned above as well as the just shown results indicate that the observation of an isolated node does not reveal the complete picture. All these approaches in this chapter are valuable tools in the analysis of a complex communication network, and many of the results provide a more then sufficient description of the underlying dynamics. The complete understanding of the influences on the dynamic by the complex interplay of many constituents in different kinds of networks are still a major challange in network research [54, 55].

3.4 Estimation of the sending time

The importance but also the difficulties in the determination of the average sending times $\langle t_i^{\text{send}} \rangle$ or their reciprocal, the outgoing rates μ_i^{out} have already been addressed in Section 3.2. Not only the estimation of the critical load depends very sensitively on the sending time, also for the purpose of optimising network topologies an accurate description is of high value [35]. This graph theoretical description of the sending times is a complex problem reflecting many difficulties inherent in such networks. With the accumulated knowledge especially about the behaviour of a nodes buffer queue as presented in Section 3.3.2 this matter will now be pushed further. The dependence on the packet creation rate μ is explicitly incorporated, although a fully consistent approach can not be obtained.

Starting out from the perspective of a single node i with a non-empty buffer queue $n_i > 0$, as long as there are no other nodes with non-empty buffer queues in its surrounding, it will have no problems forwarding its packets. The sending time of this node will thus be $t_i^{\text{send}} = 1$. In the other case the node will not get immediate sending permission if it is already MAC-blocked by a first order neighbour $j_1 \in \mathcal{N}_i^{\text{in}}$ that has at least an outgoing link towards i and takes part in another transmission. Either node j_1 is itself sender and thus blocks all its outgoing neighbours or it does receive a packet from one of its bidirectional neighbours $j_2 \in \{\mathcal{N}_{j_1} \setminus i\}$. Based upon these considerations the sending time has to be expressed as

$$t_i^{\text{send}} = 1 + \Delta t_1 + \Delta t_2 \quad . \tag{3.57}$$

The contribution Δt_1 represents the competition for sending with the first order neighbours j_1 that have at least an unidirectional link towards i. It can be expressed as the sum over the probability of all these nodes to have at least one data packet in their respective buffer queues

$$\Delta t_1 = \sum_{j_1 \in \mathcal{N}_i^{\text{in}}} p(n_{j_1} \geq 1) \quad . \tag{3.58}$$

Along the same line Δt_2 is defined as

$$\Delta t_2 = \sum_{j_2 \in \{\mathcal{N}_{\mathcal{N}_i^{\text{in}}} \setminus \mathcal{N}_i^{\text{in}}\}} p(n_{j_2} \geq 1) \sum_{j_1 \in \mathcal{N}_i^{\text{in}}} \frac{N}{(N-1)} \frac{B_{j_2 j_1}^{\text{link}}}{2B_{j_2}^{\text{node}}} \quad . \tag{3.59}$$

In addition to the prerequisite that the second order neighbour $j_2 \in \{\mathcal{N}_{\mathcal{N}_i^{\text{in}}} \setminus \mathcal{N}_i^{\text{in}}\}$ has at least one packet it is also required that the packet is forwarded to a first order neighbour $j_1 \in \mathcal{N}_i^{\text{in}}$ of i. This is represented by the sum over the ratio $B_{j_2 j_1}^{\text{link}}/2B_{j_2}^{\text{node}}$ which is appropriately normalised.

Using equation (3.20) the probability for a non-empty buffer queue can be expressed in terms of in- and outflux to node j

$$p(n_j \geqq 1) = 1 - p(n_j = 0) = \frac{\mu_j^{\text{in}}}{\mu_j^{\text{out}}} \quad . \tag{3.60}$$

Employing the identities (3.9) and (3.10) equation (3.57) can be rewritten as

$$t_i^{\text{send}} = 1 + \sum_{j_1 \in \mathcal{N}_i^{\text{in}}} \mu B_{j_1}^{\text{node}} (N-1) t_{j_1}^{\text{send}} + \sum_{j_2 \in \{\mathcal{N}_{\mathcal{N}_i^{\text{in}}} \setminus \mathcal{N}_i^{\text{in}}\}} \mu B_{j_2}^{\text{node}} N t_{j_2}^{\text{send}} \sum_{j_1 \in \mathcal{N}_i^{\text{in}}} \frac{B_{j_2 j_1}^{\text{link}}}{2B_{j_2}^{\text{node}}}. \tag{3.61}$$

This is an inhomogeneous set of N coupled linear equations in t_i^{send}. The solution for the node-dependent sending time, that can be obtained numerically, does explicitly depend on the packet creation rate μ in a non-linear manner.

It has to be pointed out that equations (3.58) and (3.59) overestimate the competitive strength of the first and second order neighbours of i. These neighbouring nodes can already be MAC-blocked by some further distant communication. In such cases these nodes do not compete with i to gain sending permission. In order to correct this shortcoming one would have to include a probability that adequately reflects the chances of such previous MAC-blocking events. Since an appropriate modelling of this quantity can not be provided, two versions of equation (3.61) will be discussed; one excluding and one including the contributions from Δt_2 given in (3.59).

For a given network realisation the graph theoretical quantities node inbetweenness B_i^{node} and link inbetweenness B_i^{link} can be calculated as presented in Section 3.1.2. The coupled system of linear equations (3.61) can be solved numerically depending on the packet creation rate μ. The lower graph in Figure 3.14 gives an illustration of these results for the most sensible node of the constant-P reference network. Again, this network consists of 100 nodes with the constant-P rule ($k^{\text{constP}} = 24$). The simulations of the generic data traffic covered $5 \cdot 10^5$ time steps. They are compared to the results of (3.61). The blue data points refer to the exclusion of the Δt_2 contributions, the green points take them into account.

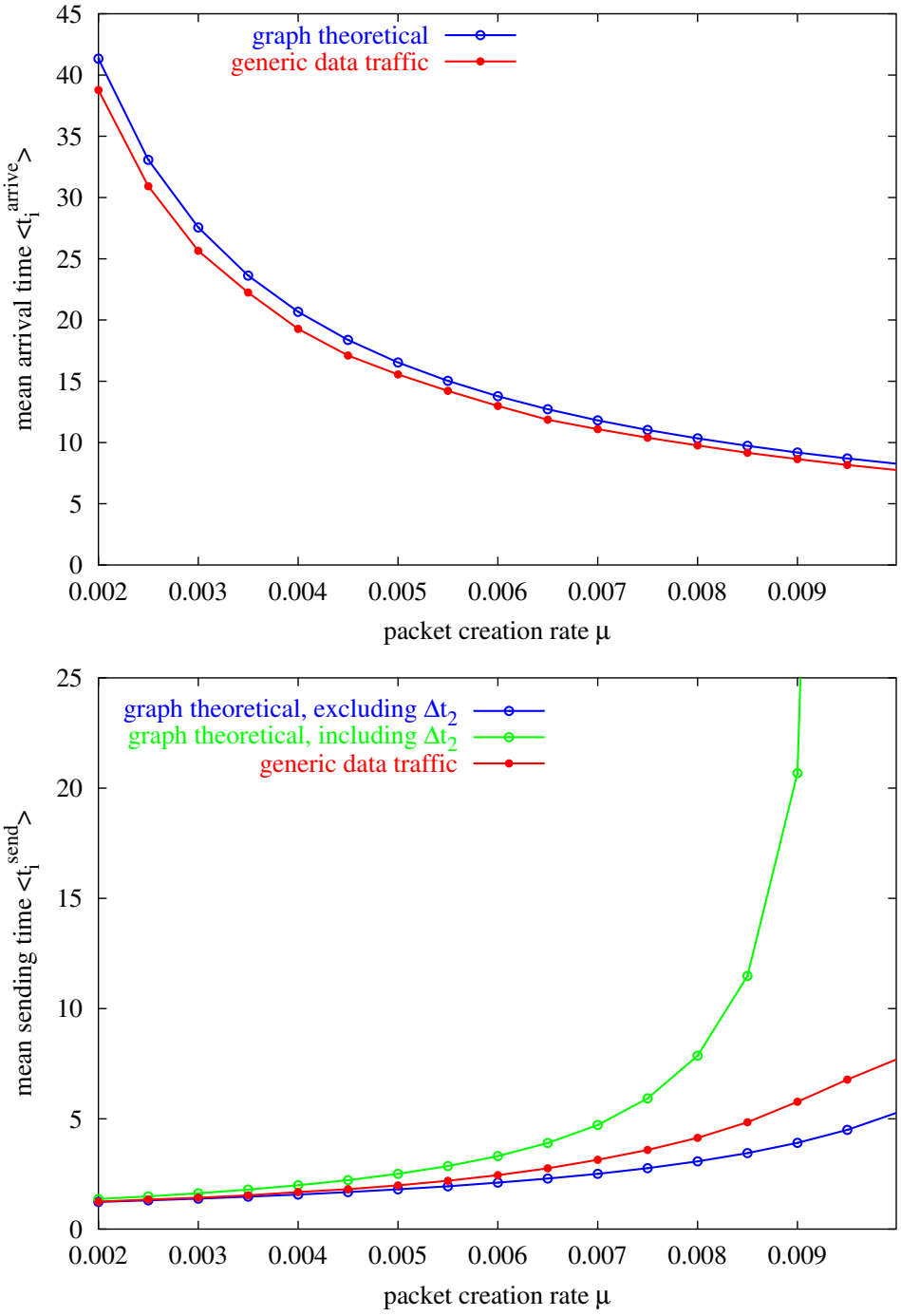

Figure 3.14: Estimation of interarrival and sending time
The upper graph shows the estimate of the mean interarrival time $\langle t_i^{\text{arrive}} \rangle$ according to equation (3.9) in comparison to the generic data traffic. The graphs apply for the most sensible node in the constant-P reference network (100 nodes, $k^{\text{constP}} = 24$, $5 \cdot 10^5$ time steps). The lower graph provides an illustration of the mean sending time $\langle t_i^{\text{send}} \rangle$ according to (3.61). The blue data points correspond to the exclusion of the Δt_2 contributions, the green ones to their inclusion.

As expected the complete description with the Δt_2 contributions overestimates the results from the generic data traffic. In contrast, their exclusion is closer to the real data. However, both cases show a strong non-linear dependence of the nodes sending time on the packet creation rate μ.

The estimated sending times from equation (3.61) can be used to approximate the end-to-end throughput of a given network. It was shown in [35] that the scaling behaviour $T_{e2e} \sim N^{\kappa}$, discussed in Section 3.2.1, with the particular scaling exponents κ for constant-P and minimum node degree networks observed in the data traffic simulations could be reproduced by use of these results.

3.5 Data traffic in networks with minimum node degree rule

Except of the throughput analysis in Section 3.2.1 the studies within this chapter have been focused on network graphs with constant-P power assignment $k^{\mathrm{constP}} = 24$. In particular one realisation of such a network, called the constant-P reference network, of 100 nodes was chosen to exemplarily study dynamical properties inherent to data traffic on this class of random geometric graphs. It should be explicitly pointed out, that the features studied in the previous sections for the constant-P reference network have been checked for a number of similar network realisations constructed out of the constant-P ensemble with $k^{\mathrm{constP}} = 24$. In that respect the obtained results can be seen as general properties for this specific model of data traffic on such network topologies.

In this respect it is a legitimate question to ask now whether and how the observed properties do show a dependence on modified network topologies, like for example those generated with the minimum node degree rule for connectivity introduced Section 1.4. These networks exhibit one significant difference compared to the networks with constant-P rule: Since networks with the minimum node degree rule come with an inhomogeneous transmission power distribution they can contain unidirectional links. As discussed in Section 3.1.1 these links can not be used for data transfer but do still carry the MAC-blocking signal. Considering the fact that the actual transmission event regulating the in- and outflux to a node is still a competitive action, the pure existences of unidirectional links is not expected to show strong impact on the dynamical properties of the generic data traffic. A detailed analysis of network realisations with the minimum node degree rule reproduced the results found for the networks with constant-P power assignment. Without showing it shall be remarked that the probability distributions of the interarrival time, the sending time and the buffer queue length showed exponential behaviour strongly suggesting the Poissonian in- and outflux to the nodes in $\mathcal{N}$ for the networks with minimum node degree rule. Also, the single-node temporal correlations exhibit a qualitative identical picture compared to the constant-P rule

networks. A comparison of the throughput for constant-P networks with networks constructed by the minimum node degree rule has been given in Figure 3.4. The differences especially towards smaller network sizes N have to be attributed to the fact that a constant-P network with $k^{\mathrm{constP}} = 24$ is more likely to be almost fully connected than a minimum degree network with $k^{\min} = 8$. As explained in Section 3.2.1 a fully connected network where each node i has a direct bidirectional link to every other node $j \in \{\mathcal{N} \setminus i\}$ is superior for small network sizes N by always guaranteeing $T_{e2e} = 1$.

This close correspondence suggests that the actual routing decisions are insensitive to topological changes between networks with constant-P and minimum node degree rule. The reasons have to be seen in the exclusive use of shortest paths for packet forwarding that eliminates all further degrees of freedom within the routing decisions. It is precisely this enforcement of one predefined path that supports the Poissonian in- and outflux. Even if a node is almost critical no detour is allowed. As soon as permission is obtained the node will receive further packets. This behaviour is independent from the existence of unidirectional links. The influences of a modified routing that is not based on topological shortest paths $d_{i,f}$ are presented in Chapter 4. The impact of the actual routing decisions on the dynamical properties of data traffic becomes obvious.

3.6 User behaviour

In a real world scenario different amounts of data are send back and forth between to two or more partners. This data is split into packets of definite size, which causes not just one but a multitude of packets to be generated at once and to travel along the same or similar paths. This particular property is one effect that contributes to multiscale phenomena in network traffic. Prominent examples are self-similarity of round-trip-times in Ethernet traffic [13] and in the World Wide Web [67, 68], connection durations [11] and interarrival times [42] in communication networks.

Two specific approaches have been applied to incorporate such phenomena in the generic data traffic simulation. In the first approach not just one but a fixed number of packets was generated at each chosen node, in the second the number of packets was determined according to a truncated Pareto distribution. These modifications do not represent all the features present in real world user behaviour. Nevertheless they allow to study the differences in network performance that occur if not just one but a multitude of packets start from a certain node. A presentation of the simulation results is given.

In the first modified version of data traffic not just one data packet is created at a node with probability μ at each time step but PCS packets are generated there with a probability $\mu^{\mathrm{eff}} = \frac{\mu}{\mathrm{PCS}}$. PCS is referred to as the packet cluster size

which is nothing else then an integer number counting the packets created in one instance at one node. The packet cluster size PCS is fixed within one simulation, so every node i faces the same static probability of a packet generation event $\mu_i = \mu^{\text{eff}} = \frac{\mu}{\text{PCS}}$. Because of this latter definition the average total number of packets generated in one time step $\mu^{\text{eff}} \cdot \text{PCS} \cdot N = \frac{\mu}{\text{PCS}} \text{PCS} \cdot N$ is still identical to the former version μN.

The throughput $T_{e2e} = \mu^{\text{crit}} N$ has been tested exemplarily for two packet cluster sizes PCS $= 4$ and PCS $= 16$. The results are shown in Figure 3.15. The simulations have been performed on networks of different size N with the constant-P power assignment, $k^{\text{constP}} = 24$ and spanned $5 \cdot 10^5$ time steps. Each throughput was sampled over 100 different network realisations. As a reference the curve for PCS $= 1$, already shown in Figure 3.4, is also included in the plot. A massive decline of the throughput is observed for the PCS $= 16$ case at higher system sizes. For networks of size $N \approx 20$ the network is almost fully connected for $k^{\text{constP}} = 24$ with $D \approx 1$ where packets reach their final destination in one hop. As soon as the network size grows multihop connections are necessary in order to transmit packets. For large clusters of data packets the individual packets now start to interact with each other along their shortest paths. This significantly slows down data traffic as indicated in Figure 3.15. The effect is not that pronounced for the PCS $= 4$ case.

The arrival of one first packet to any node is an indication of (PCS $- 1$) packets to arrive soon. This effect becomes visible in the probability distribution of the buffer queue length and the interarrival time. Figure 3.16 shows the probability distribution of the buffer queue length $p(n_i)$ for the most sensible node of the constant-P reference network with $k^{\text{constP}} = 24$, 100 nodes. The results for PCS $= 4$ and PCS $= 16$ are depicted for a subcritical packet creation rate $\mu = 0.009$ in the case of the constant-P reference network. As a reference the analytically obtained curve for PCS $= 1$ from equation (3.20) is included in the plot. Clear deviations from the exponential behaviour are visible. The close correlation of the subsequent arrival of packets from one cluster causes a significant contribution to the probability distribution at higher lengths of the buffer queue.

Similar deviations are observed in the corresponding probability distributions of the interarrival time of that same node, given in Figure 3.17. On the one hand the arrival of a packet to a buffer queue is a clear sign of (PCS $- 1$) other packets to arrive soon. This contributes to rather small interarrival times $t_i^{\text{send}} \lesssim \langle k \rangle = 25$ in the order of the number of neighbours of i. On the other hand the reduced effective packet creation rate $\mu^{\text{eff}} = \frac{\mu}{\text{PCS}}$ leads to longer periods where no clusters of packets are forwarded via i. This accounts for the long tail of the probability distribution towards high interarrival times.

Without showing it shall be remarked that the probability distribution of the sending time $p(t_i^{\text{send}})$ does also differ from the ideal PCS $= 1$ case, altough the deviations are not that significant. The major reason for these rather small changes

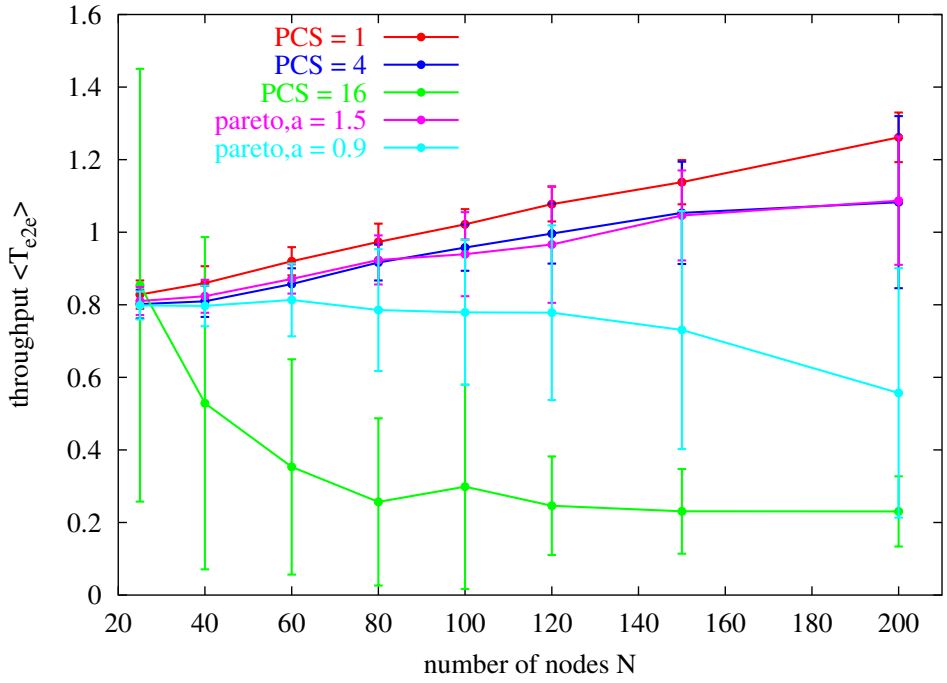

Figure 3.15: Throughput for different user behaviour
The throughput T_{e2e} for different user behaviours is shown depending on the size N of the network. A sampling over 100 network realisations has been performed for each network size. All networks come with the constant-P rule with $k^{\mathrm{constP}} = 24$. The results for PCS $= 4$ (blue), PCS $= 16$ (green) and PCS $= 1$ (red) are depicted. For generic data traffic with a Pareto-like packet generation the curves for $a = 1.5$ (pink) and $a = 0.9$ (light blue) are provided. The error bars correspond to the variance of the sampling process.

has to be seen in the particular way of measuring the sending time which is only counted if the buffer queue is non-empty. So there might be longer periods where the node is just empty, doing nothing. The interarrival time counts these periods, the sending time does not.

Compared to the single-node temporal correlations $r_i(\Delta t)$ studied in Section 3.3.4 for the PCS $= 1$ case Figure 3.18 illustrates the changes introduced by the simultaneous generation of PCS > 1 packets. Basically the correlation decline is shifted towards larger times Δt for higher packet cluster sizes PCS. This observation nicely fits in the above given context that one arriving packet is always correlated with the arrival of the other simultaneously generated packets to a given node. In a simplified picture this causes a buffer queue to grow over a certain period in time. Until all the packets are send the buffer queue yields $n_i > 0$

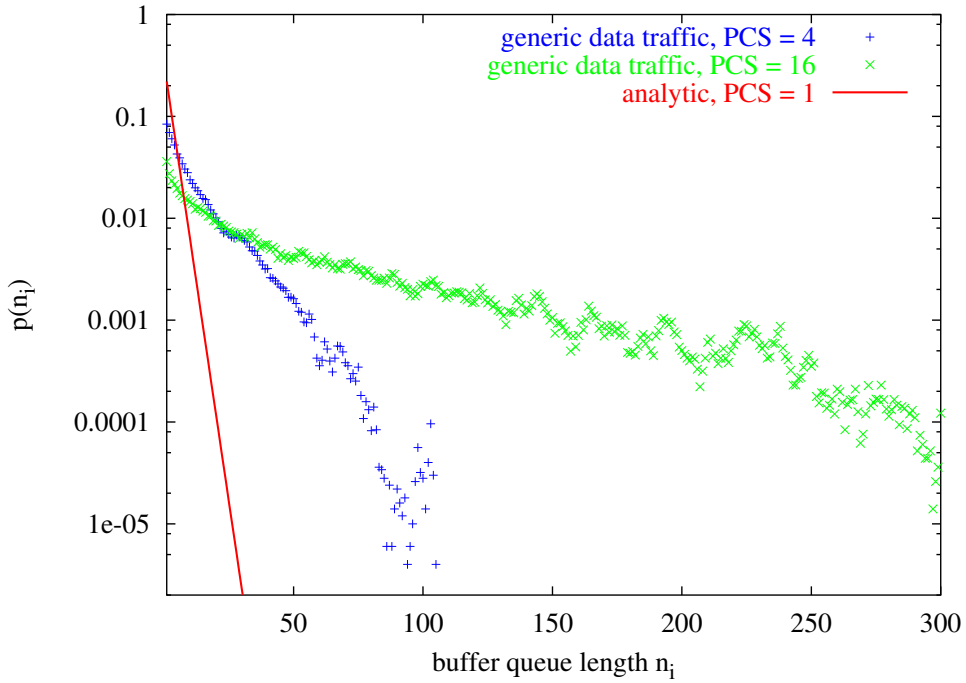

Figure 3.16: Probability distribution of the buffer queue: user-input
The probability distribution of the buffer queue length $p(n_i)$ for the most sensible node of the constant-P reference network is shown. The results for PCS $= 4$ and PCS $= 16$ are depicted for a subcritical packet creation rate $\mu = 0.009$. The simulations covered $5 \cdot 10^5$ time steps. As a reference the analytically obtained curve for PCS $= 1$ from equation (3.20) is included in the plot; $\mu_i^{\text{in}} = 0.117$ and $\mu_i^{\text{out}} = 0.173$ are obtained from the generic data traffic simulation at $\mu = 0.009$.

which holds a significant contribution to $\langle n_i(t + \Delta t)\, n_i(t)\rangle$. This effect causes the graphical representations of the single-node temporal correlations to be almost constant for smaller to intermediate Δt and to shift the onset of the decline.

In the second modification of the generic data traffic the number of packets m generated at a chosen node at one time step follows an truncated Pareto distribution. In most cases the amount of data transmitted is rather small (and so fits in one data packet) but there is a non-zero probability of having so much data to require up to PCS_{max} packets. A similar scale-free behaviour of the probability density functions of the size of data packets to be transmitted is observed in the World Wide Web [12, 68] and similar data networks [13].

In each time step a node is chosen to generate new packets with a probability $\mu^{\text{eff}} = \frac{\mu}{\langle m\rangle}$. The exact number of packets to be generated is then determined by

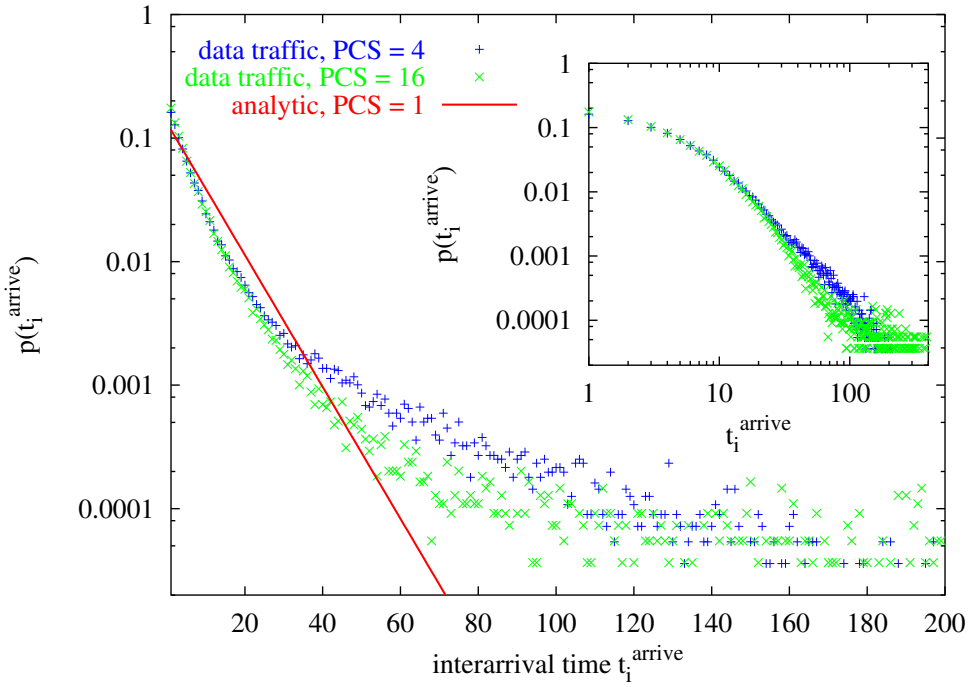

Figure 3.17: Probability distribution of the interarrival time: user-input
The probability distribution of the interarrival time $p(t_i^{\text{arrive}})$ for the most sensible node of the constant-P reference network is shown. Parameters are the same as in Figure 3.16. The inset illustrates the same plot in log-log scale. As a reference the analytical obtained curve for PCS $= 1$ from equation (3.15) is included in the plot; $\mu_i^{\text{in}} = 0.117$ is obtained from the generic data traffic simulation.

the probability density function of the discrete truncated Pareto distribution

$$p(m) = \frac{1}{C} a k^a m^{-a-1} \quad (m \in \mathbb{N}, m = 1 \ldots \text{PCS}_{\max}, m \geq k) \tag{3.62}$$

with $C = \sum_{m=1}^{\text{PCS}_{\max}} a k^a m^{-a-1}$ being the normalisation constant. The mean number of packets created by a Pareto event is given by

$$\langle m \rangle = \sum_{m=1}^{\text{PCS}_{\max}} m\, p(m) \quad . \tag{3.63}$$

The average total number of packets created in one time step is given by $\mu^{\text{eff}} \langle m \rangle N = \frac{\mu}{\langle m \rangle} \langle m \rangle N$ which is again identical to the former version μN. Without loss of generality the parameter k is set 1.

The performance of the network was tested for two different scaling exponents $a = 0.9$ and $a = 1.5$ of the Pareto density function in equation (3.62) with

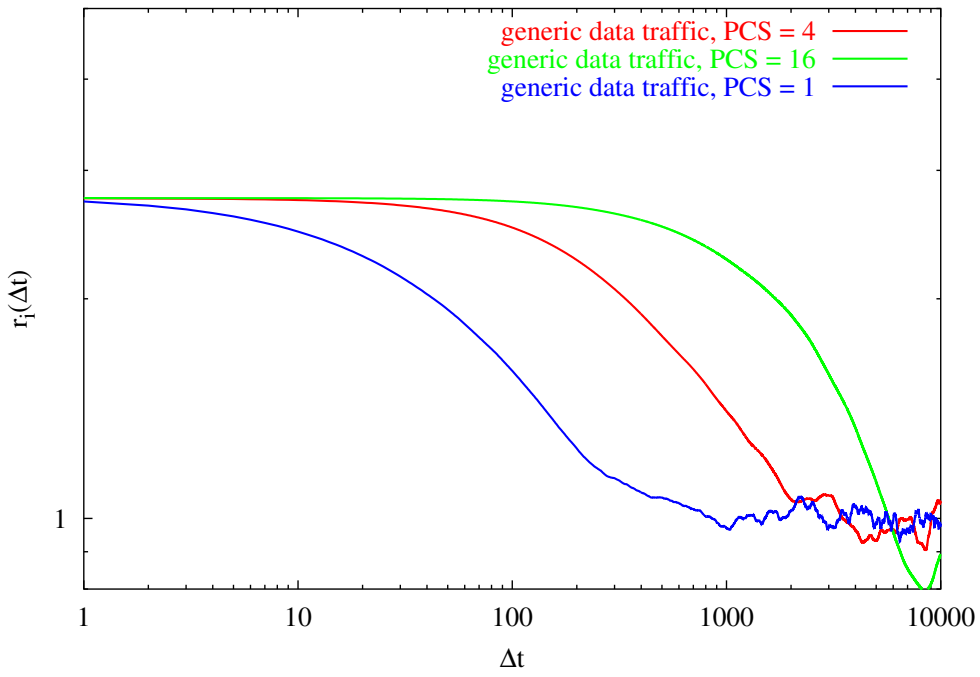

Figure 3.18: Single-node temporal correlations: user-input
The single-node temporal correlations $r_i(\Delta t)$ for the most sensible node of the constant-P reference network are shown. Parameters are the same as in Figure 3.16. For comparison the graph for PCS $= 1$ is included in the plot.

$\text{PCS}_{\text{max}} = 16$. The mean number of generated packets according to equation (3.63) is evaluated to $\langle m \rangle = 2.42$ in the $a = 0.9$ case and $\langle m \rangle = 1.662$ in the $a = 1.5$ case. The corresponding curves for the throughput depending on the network size are included in Figure 3.15. Again, if the average packet cluster size of generated packets increases, as in the $a = 0.9$ case compared to the $a = 1.5$ case, the throughput decreases for higher packet creation rates. The same argumentation, that packets of one cluster interact with each other on their common path to the final receiver, should also hold for the Pareto-like packet generation.

A representative probability distribution of the interarrival time measured at the most sensible node of the constant-P reference network is shown in Figure 3.19 for the two different scaling exponents a as well as a reference plot for a fixed packet cluster size PCS $= 1$. The packet creation rate was fixed to $\mu = 0.0095$ which is in the subcritical regime for the constant-P reference network. The divergence from the exponential behaviour is clearly visible in the log-lin plot. The probability of finding rather large interarrival times does not converge to zero as quickly as in the PCS $= 1$ case. The inset shows the same probability distributions in a log-log

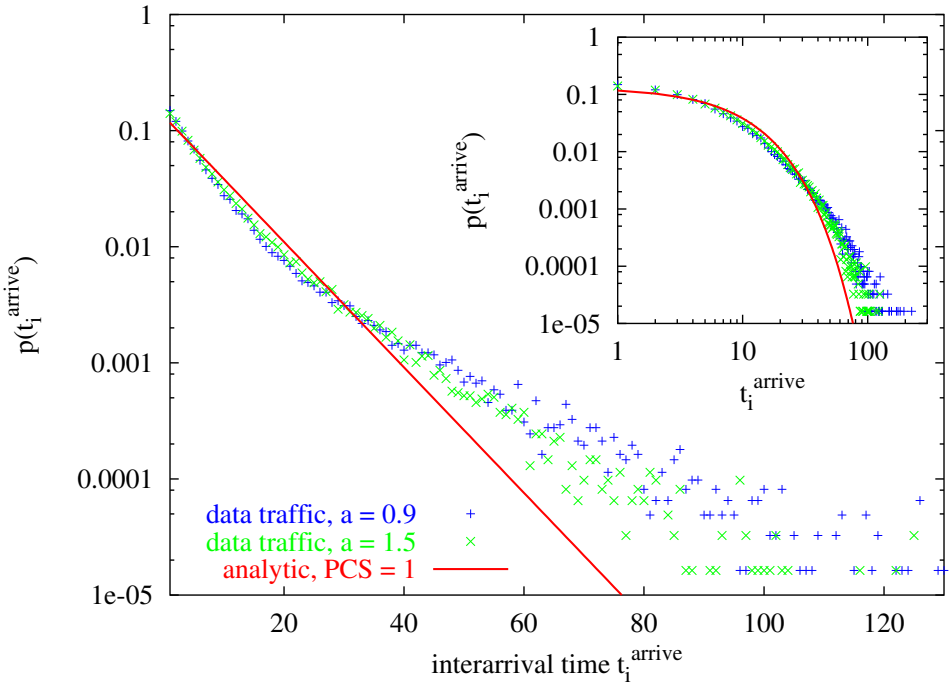

Figure 3.19: Probability distribution of the interarrival time: Pareto-input

The probability distribution of the interarrival time $p(t_i^{\text{arrive}})$ is illustrated for the Pareto-motivated data traffic with scaling exponents $a = 0.9$ (blue) and $a = 1.5$ (green). The analytical curve obtained from (3.15) with $\mu_i^{\text{in}} = 0.117$ for PCS $= 1$ (red) is also included. The most sensible node of a constant-P reference network was observed at $\mu = 0.0095$. The simulation covered $5 \cdot 10^5$ time steps.

plot.

In a similar fashion a non-zero probability for large buffer queue lengths is introduced in the corresponding probability distribution as depicted in Figure 3.20, analogous to the fixed PCS case. The subsequent arrival of packets in larger clusters causes the buffer queue of the most sensible node to temporarily store significant numbers of packets. The probability distributions for the buffer queue length clearly deviates from the ideal exponential behaviour. Riedi et al. [42] state that the distributions can be better fitted by a Weibull probability density function

$$p_{(a_w,b_w)}(n) = \frac{a_w}{b_w}\left(\frac{x}{b_w}\right)^{(a_w-1)} e^{-(x/b_w)^{a_w}} \tag{3.64}$$

in the case of scale-free data traffic. The corresponding fits in Figure 3.20 support this picture. Parameters are $a_w = 0.77$ and $b_w = 73.46$ in the $a = 0.9$ case and

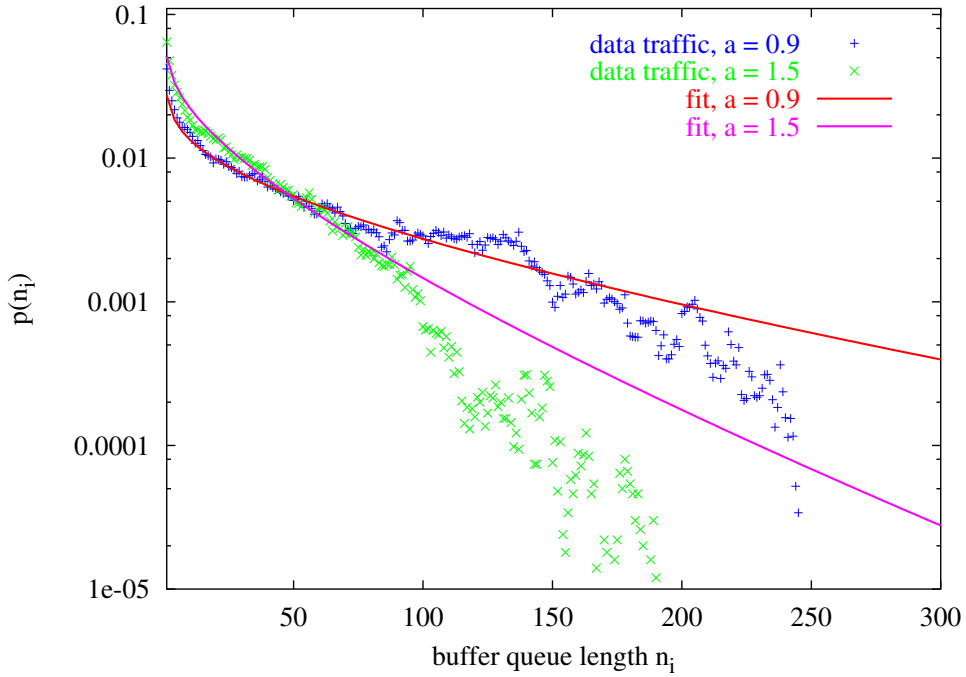

Figure 3.20: Probability distribution of the buffer queue: Pareto-input
The probability distribution of the buffer queue length $p(n_i)$ is illustrated for the Pareto-motivated data traffic with scaling exponents $a = 0.9$ (blue) and $a = 1.5$ (green). The most sensible node of a constant-P reference network was observed at $\mu = 0.0095$. The simulation covered $5 \cdot 10^5$ time steps. The curves are fitted with a weibull probability density function (3.64). Parameters are given in the text.

$a_w = 0.79$ and $b_w = 28.96$ in the $a = 1.5$ case.

An interesting behaviour could be observed looking at the probability distribution of the end-to-end time delay $p(t_{\text{e2e}})$ that has already been discussed in Section 3.2.2 for the PCS $= 1$ case. A best fit with a log-normal probability density function (3.14) was used to describe the probability distribution in the PCS $= 1$ picture. The results were promising but not perfect, consult Figure 3.7. In contrast, with a Pareto-like distribution of the packet cluster size the log-normal description of the probability density function of the end-to-end time delay provides almost perfect results over many orders of magnitude. The plots for the Pareto-motivated data traffic with scaling exponents $a = 0.9$ and $a = 1.5$ as well as the PCS $= 1$ case are shown in Figure 3.21. A best fit with a log-normal probability density function according to equation (3.14) is attached to each of the data sets. The particular fitting parameters are $\mu_L = 3.589$ and $\sigma_L = 1.665$ in the $a = 0.9$ case; $\mu_L = 3.132$

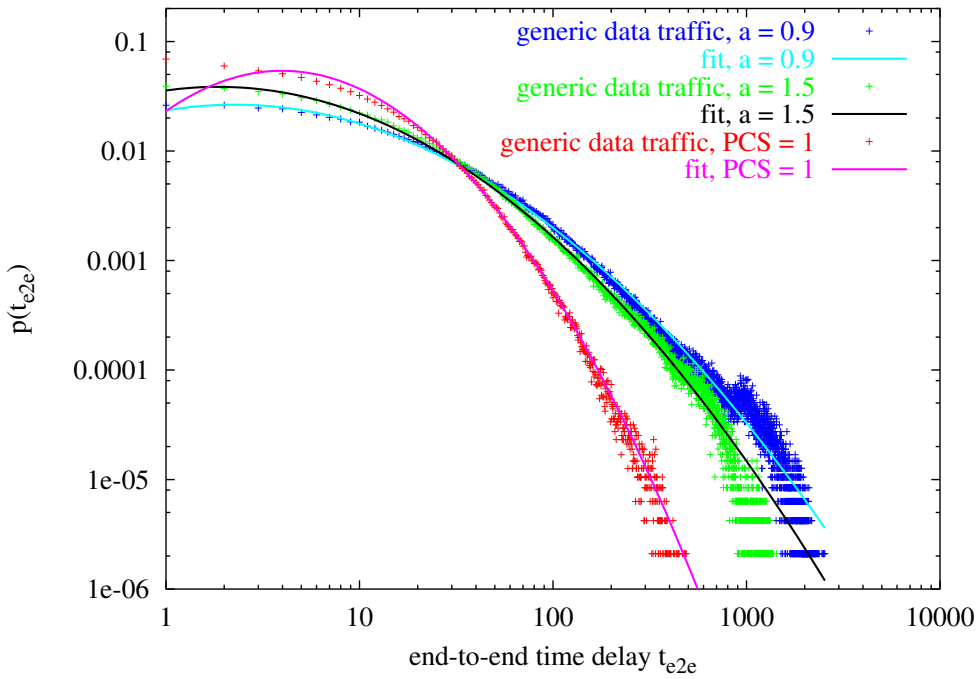

Figure 3.21: Probability distribution of the end-to-end time delay: Pareto-input
The probability distribution of the end-to-end time delay $p(t_{\mathrm{e2e}})$ is illustrated for the Pareto-motivated data traffic with scaling exponents $a = 0.9$ (blue) and $a = 1.5$ (green) as well as the PCS $= 1$ case (red). A constant-P reference network was used with 100 nodes and constant-P rule ($k^{\mathrm{constP}} = 24$). Fitting parameters are given in the text.

and $\sigma_L = 1.584$ in the $a = 1.5$ case; and $\mu_L = 2.504$ and $\sigma_L = 1.060$ in the PCS $= 1$ case.

As already discussed in Section 3.2.2, a detailed understanding of this phenomena could not be provided so far. Especially with non-Poissonian in- and outflux the task becomes even more difficult.

3.7 Summary

Within this chapter a simplified data traffic model was employed to study general properties of network traffic on wireless multihop ad hoc networks. The model is based on the random generation of data packets in the network that are forwarded along shortest paths to their final receiver. The transmission process pays respect to the special limitations of data transfer in a shared medium by employing a

simplified medium access control protocol (MAC).

The identification of the node inbetweenness B_i^{node} as a suitable graph theoretical measure to account for the frequency of use of a given node i and the definition of the influx $\mu_i^{\text{in}} = 1/\langle t_i^{\text{arrive}} \rangle$ and the outflux $\mu_i^{\text{out}} = 1/\langle t_i^{\text{send}} \rangle$ provided valuable tools for the description of global and local network observables. The observation of a Poissonian in- and outflux to the nodes in $\mathcal{N}$ allowed an analytic understanding of a number of dynamical network observables like the buffer queue length and the single-node temporal correlations. A description of the sending time solely based on graph theoretical measures has been given in an approximate form.

A comparison of data traffic on network graphs with constant-P and minimum degree power assignment did not show remarkable differences. It has to be attributed to the packet forwarding along shortest paths that the dynamical properties of network traffic are almost insensitive to a modified topology.

A modified packet generation was included to model specific user behaviour. Especially for the influx a deviation from the Poissonian behaviour can be observed. Based on the previous findings a qualitative understanding of the occuring differences could be provided.

4 Distributed routing control

The fast and efficient flow of data through any communication network represents a highly desired feature. It is a major design principle but also a direct criteria for the performance of a network. In the recent literature many publications focus on the close connection between the topology of a network and its possible advantages and disadvantages for different routing strategies [64, 69, 70]. The main focus was caught by the investigation of internet data traffic [71, 72, 73]

Future wireless mobile ad hoc networks require a fast adaption to changes in the topology caused by the mobility of nodes, their on/off states or their temporal activity. As already mentioned the ad hoc networks can hardly be based on small world scenarios that allow communication over long distances with a fairly small number of hops due to the shared communication media and the emerging interference effects. Unlike structures in the World Wide Web where connections between major hubs are based on powerful high capacity links corresponding communication links in a wireless ad hoc network provide almost identically limited capacity for data transfer. Just taking these special properties of wirless ad hoc communication networks into account it becomes clear that these particular topologies require specific routing strategies for optimal performance. The highly complex interplay between topology of a network and dynamic processes on that network have developed into a focal point of latest network research [64, 74].

Within this chapter the investigations are limited to the constant-P topologies studied so far. Given this restriction the fundamental aim of this chapter is to provide simple algorithms that demonstrate the fundamental possibility of establishing routing schemes that serve the basic needs of wireless ad hoc communication including elements of distributive flow control. Because of the absence of a master authority in the observed networks a decentralised approach is desired that reveals a high degree of self-organisation and stability under different traffic situations. It should be mentioned that in this chapter no explicit protocols will be constructed. It is not the intention to design a complete routing algorithm ready to be implemented in an existing realisations of wireless communication networks. It seems far more valuable to review the specific needs of wireless communication and to present and discuss possible algorithms that might be further developed into complete protocols.

The underlying topology of the communication network is kept fixed. Thus mobility and emergence or loss of communication links as well as the management of vanishing or newly introduced nodes is not taken into account. The question to be asked is how routing can be established in an environment where nodes only

have information about the local neighbourhood and how this information can be employed to provide efficient routing.

Section 4.1 reviews the routing scheme based on shortest paths. By including the choice between degenerate shortest paths depending on the buffer queue length of the next hop neighbour the actual state of the network is used as a decision criteria. Section 4.2 introduces a self-organised routing similar to an asynchronous distance vector routing proposed by Boyan and Littman [75]. The routing scheme is modified in order to cope with wireless ad hoc networks. In the simplified version of Section 4.2 the routing converges to the shortest path algorithm from Section 4.1. This approach is modified in Section 4.3 by introducing network state dependend measures in the update procedures of the routing tables. The routing presented in this section adapts to spatially and temporarily localised congestion and allows significantly increased throughput compared to the so far used fixed shortest path method. Section 4.4 additionally includes a learning rate that keeps a memory of former network states.

4.1 Shortest path algorithms

The algorithm used in the previous chapter to manage the routing of packets from the chosen source node $i \in \mathcal{N}$ to the final receiver $f \in \{\mathcal{N} \setminus i\}$ employed shortest paths $\sigma_{i,f}$ in the sense of Section 1.1. Since this routing algorithm is only based on topological measures of the network graph this approach is classified as static routing [76]. A classical Dijkstra algorithm [57], as explained in Section 3.1.1, is employed for the determination of the shortest paths. Thus for every node i a routing table

$$F_f^i = \mathcal{S} \tag{4.1}$$

can be supplied that provides a list $\mathcal{S}$ of suitable bidirectional neighbours $j \in \mathcal{N}_i$ for every final receiver $f \in \{\mathcal{N} \setminus i\}$ that minimises the hop-distance towards f.

If more than one shortest path to the final receiver f exist that split at i then more than one possible next hop neighbours are listed in $\mathcal{S}$. The number of choices in $\mathcal{S}$ is denoted as $|\mathcal{S}|$. For the data traffic simulation employed in the previous chapter all lists $\mathcal{S}$ are from the very beginning reduced to just one element taken randomly out of possibly different choices in $\mathcal{S}$. In this way path degeneracy is completely suppressed. Every node i only holds its limited routing table $F_f^i = \mathcal{S}^1$. This routing-scheme is referred to as fixed-fifpo, where the "fifpo" means first-in-first-possible-out and takes into account that a node chosen as a sender can sequentially try to send the packets in its buffer queue. This algorithms allows the node to send its second packet in its buffer to its next hop node in case the sending of the first packet is not possible due to a MAC-blocked next hop node. The node only remains quiet if none of the packets in its buffer queue can be

forwarded to any of the designated next hop nodes. Consult again Section 3.1.1 for reference.

The existence of the full information about all degenerate shortest paths in $\mathcal{S}$ allows the introduction of two extended versions of the shortest path algorithms: the random-fifpo and the shortest-fifpo routing scheme. Their presentation provides a simple background for the introduction of the tools necessary to compare routing and network performance in the remainder of this chapter.

The random-fifpo routing scheme only differs that far from the fixed-fifpo that each node contains not only the limited routing lists $\mathcal{S}^1$ with only one predefined element for each F_f^i but the complete list $\mathcal{S}$. For every packet in a nodes buffer queue which is checked for sending, one out of the possible neighbours j_{next} contained in $\mathcal{S}$ is chosen with equal probability $p(j_{\text{next}} \in \mathcal{S}) = 1/|\mathcal{S}|$. If the chosen next hop node is already MAC-blocked the fifpo mode proceeds to the next packet in the buffer queue or terminates the sending assignment for the chosen sender node if the packet is the last in the buffer queue. It does not try to send the same packet to another node in $\mathcal{S}$. This routing-scheme does simply not stick to the one predefined shortest path for every end-to-end data transfer but moves around in the space of possible shortest paths between two nodes. This routing scheme is more suitable to reproduce the node inbetweenness B_i^{node} defined in (3.2). The degeneracy of shortest paths is now present in the routing algorithm, although the different shortest paths are not weighted equally.

The so called shortest-fifpo routing is a first approach towards a routing-scheme that takes the current local network state into account. As in the random-fifpo scheme every node i has the complete information $F_f^i = \mathcal{S}$ about shortest paths to all other nodes f in the network. But instead of randomly choosing one node out of this list that particular node is picked that besides its shortest path feature also maintains the shortest buffer queue ($\min_{j\in\mathcal{S}} n_j$). In that instance the sending node i uses information about its local surrounding for a routing decision. It should be remarked that if more than one node qualifies with a shortest buffer queue one of them is picked randomly.

A comparison of the performance of the three so far introduced shortest path routing schemes is given in Figure 4.1 where the throughput T_{e2e} is plotted as a function of the system size N. The simulations of $10,000$ time steps have been performed on constant-P networks with $k^{\text{constP}} = 24$. Each data point was sampled over 100 different network realisations, the variance of the sampling is given by the error bars. Determination of the critical packet creation rate $\mu^{\text{crit}} = T_{e2e}/N$ was done by a binary search algorithm. Already the simple step from the fixed-fifpo to the random-fifpo routing scheme causes a measurable increase in the throughput T_{e2e}. Even more superior is the performance of the shortest-fifpo routing scheme which is a first indication about the important influence of information about the current network state.

For a fixed system size N analog results can be obtained from Figure 4.2. For

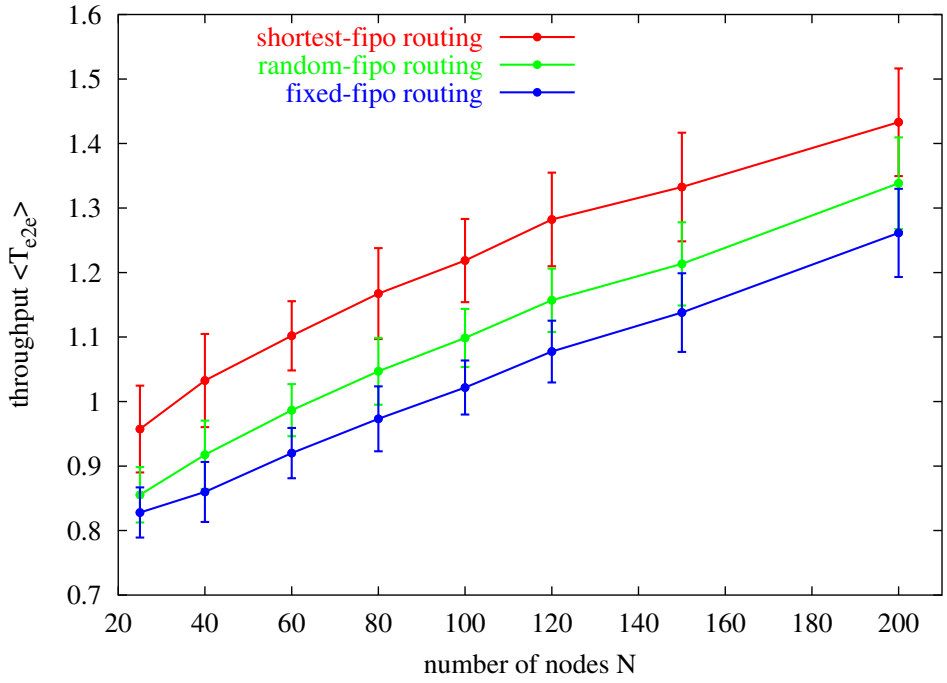

Figure 4.1: Throughput for the shortest path algorithms
The end-to-end throughput $\langle T_{e2e} \rangle$ averaged over 100 different network realisations of constant-P type with $k^{\text{constP}} = 24$ is shown for three shortest path routing algorithms depending on the network size N. A detailed description of the routing schemes is given in the text. The simulations spanned 10,000 time steps, the critical packet creation rate μ^{crit} was determined by a binary search algorithm. Error bars correspond to the variance over the ensemble of network realisations.

different packet creation rates μ the mean of the end-to-end delay $\langle t_{e2e} \rangle$ is shown. Data traffic simulation of $5 \cdot 10^5$ time steps were performed for each shortest path routing scheme depending on the packet creation rate μ. A constant-P reference network with 100 nodes was employed. The sudden increase in the end-to-end delay time is a clear sign of $\mu \to \mu^{\text{crit}}$, that the system is about to reach the supercritical state.

The single-node observables introduced in the previous chapter are not significantly changed by the use of the modified routing schemes. Especially the probability distributions of the interarrival times $p(t_i^{\text{arrive}})$ and the sending times $p(t_i^{\text{send}})$ can still be described by equation (3.15) employing the corresponding mean values $\langle t_i^{\text{arrive}} \rangle$ and $\langle t_i^{\text{send}} \rangle$.

Significant deviations are observed only in the probability distribution function of the buffer queue length $p(n_i)$ for the shortest-fifpo scheme compared to the

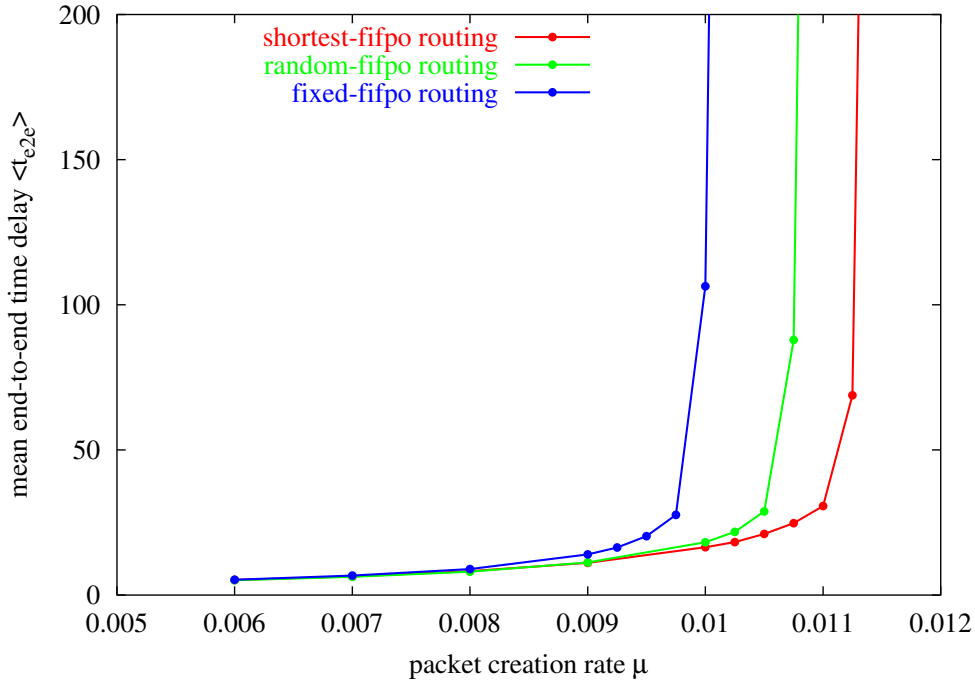

Figure 4.2: End-to-end time delay for the shortest path algorithms
The mean of the end-to-end time delay $\langle t_{e2e} \rangle$ over all packets generated in a $5 \cdot 10^5$ time steps simulation is illustrated for three shortest path routing algorithms depending on the packet creation rate μ. A constant-P reference network ($k^{\text{constP}} = 24$) of 100 nodes served as underlying network graph.

theoretical estimate given by the probability density function (3.20). This is shown in Figure 4.3. The probability distribution of the buffer queue length $p(n_i)$ of the most sensible node in a constant-P reference network is illustrated. The simulation run spanned $5 \cdot 10^5$ time steps at a packet creation rate $\mu = 0.01075$ which is close to the critical value $\mu_{\text{sf}}^{\text{crit}} = 0.0114$ for the shortest-fifpo routing within this network. The red points correspond to the data from the generic data traffic simulation whereas the green curve represents the estimate given by equation (3.20) with the corresponding values $\mu_i^{\text{in}} = 1/\langle t_i^{\text{arrive}} \rangle = 0.099$ and $\mu_i^{\text{out}} = 1/\langle t_i^{\text{send}} \rangle = 0.106$ from the same simulation run. It is the special property of the shortest-fifpo routing to prefer next hop neighbours $j \in \mathcal{N}_i$ with rather short buffer queues n_j if there exist a degeneracy of shortest paths. These correlations between the in- and outflux to a node cause the deviations between the generic data traffic results and the estimate by equation (3.20). This estimate is based on an independent Poissonian in- and outflux to a node which does not hold for the shortest-fifpo routing. The probability distribution $p(n_i)$ can still be well described

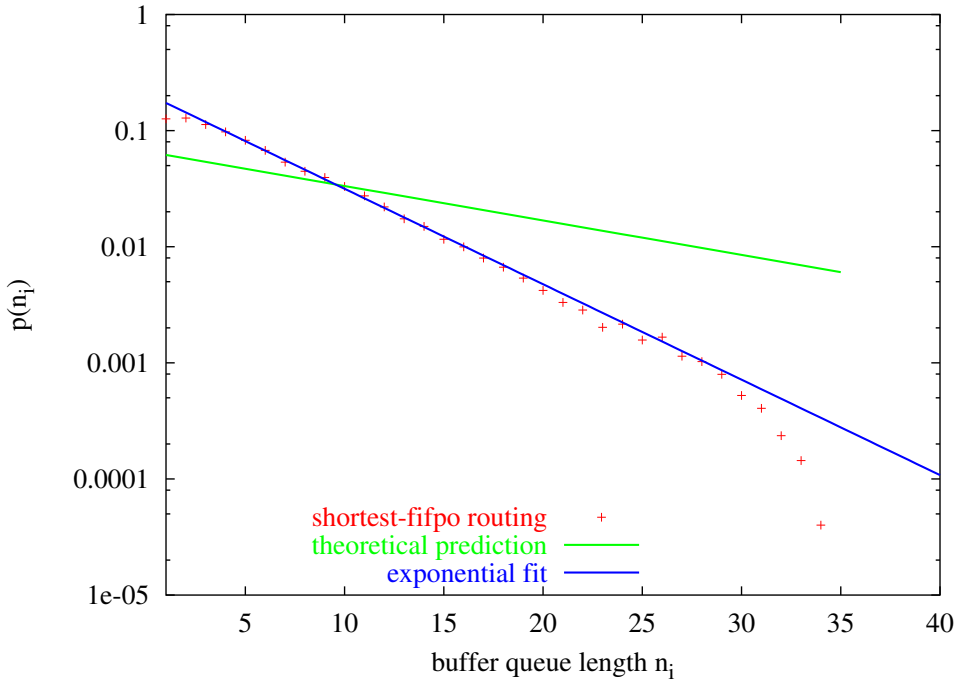

Figure 4.3: Probability distribution of the buffer queue: shortest-fifpo
The probability distribution of the buffer queue length $p(n_i)$ of the most sensible node in a constant-P reference network is shown. The simulation run covered $5 \cdot 10^5$ time steps with a packet creation rate $\mu = 0.01075 < \mu_{\text{sf}}^{\text{crit}} = 0.0114$. The red points correspond to the data from the generic data traffic simulation whereas the green curve represents the estimate given by equation (3.20) with $\mu_i^{\text{in}} = 0.099$ and $\mu_i^{\text{out}} = 0.106$ obtained in the same simulation run. The blue line is an exponential fit with (3.13). Parameters are $a_E = 0.531$ and $b_E = 5.284$.

by a exponential probability density function similar to (3.13). A fit (blue line) is included in Figure 4.3 with $a_E = 0.531$ and $b_E = 5.284$.

In Section 3.1.2 a graph theoretical definition of the node inbetweenness B_i^{node} was supplied. It was defined as the fraction of shortest paths within a network that pass node i out of all possible shortest paths. Since the fixed-fifpo routing employed in the previous chapter is exclusively based on shortest paths the node inbetweenness B_i^{node} proved to be a sufficient measure to predict local data traffic for the nodes of the network. Additionally it is more or less implicitly required that degenerate shortest paths are taken with fixed probability independent of any former decisions. This is simply fulfiled by the reduction of the routing tables $F_f^i = \mathcal{S} = \mathcal{S}^1$ that eliminates all choices in the fixed-fifpo routing. Also in the random-fifpo routing the requirement is satisfied by the randomness in the choice

of a degenerate shortest path. The graph theoretical node inbetweenness B_i^{node} should hold as a measure of local traffic density.

As soon as a routing scheme is coupled to an information about the current network state, the routing decisions do not longer maintain their statistical independence. A routing decision at t might depend on earlier routings at $(t - \Delta t)$. The correlations between several path choices make an analytic approach a highly complex task.

In the shortest-fifpo routing a node tries to avoid next hop nodes with long buffer queues given that alternative shortest routes exist. This behaviour manipulates the input to a heavily used node, in fact it reduces the inflow of packets. To quantify that modified behaviour an effective node inbetweenness B_i^{eff} is introduced:

$$B_i^{\text{eff}} = \frac{\mu_i^{\text{in}}}{\mu N} \tag{4.2}$$

The rate of incoming packets μ_i^{in} characterises the average number of packets arriving at the buffer queue of i. Thus B_i^{eff} describes the fraction of totally generated packets that are forwarded via i. In the fixed-fifpo and the random-fifpo routing it holds $B_i^{\text{eff}} \cong B_i^{\text{node}}$. In Figure 4.4 the maximum inbetweenness $(\max_{i\in\mathcal{N}} B_i^{\text{eff}})$ is shown for the three so far introduced routing schemes. The simulations of $5 \cdot 10^5$ time steps have been performed on the constant-P reference network. The curve for the fixed-fifpo routing is constant for different packet creation rates μ in accordance with the previous results. The curve for the shortest-fifpo routing shows a clear dependence on the packet creation rate μ. For higher packet creation rates the system tends towards a lower effective inbetweenness. Here the influence of the incorporation of the network state forces the system to avoid heavily used nodes, which causes the significant reduction of inflow especially to the most sensible nodes. For smaller packet creation rates μ the most sensible nodes obviously perform superiorly compared to the alternatives and serve a superproportional amount of data traffic. The inset in Figure 4.4 shows the variance of all sampled effective node inbetweenness of the network. The shortest-fifpo routing clearly tends to a narrower distribution of the effective node inbetweenness for $\mu \to \mu_{\text{sf}}^{\text{crit}}$, whereas the other algorithms are not effected by a higher packet creation rate. The higher overall variance for the shortest-fifpo routing is attributed to a certain number of nodes collaboratively managing most of the data traffic. In the fixed-fifpo and the random-fifpo only one node is heavily used whereas all other nodes are more or less unburdened.

One may observe the slight dependence of the maximum effective node inbetweenness $\max_{i\in\mathcal{N}}(B_i^{\text{eff}})$ on the packet creation rate μ for the random-fifpo routing in Figure 4.4. This is caused by an artefact of the fifpo algorithm. For the most sensible node a higher packet creation results in higher in- and outflux to this node. So the node might already be involved in a transmission if one of its neigh-

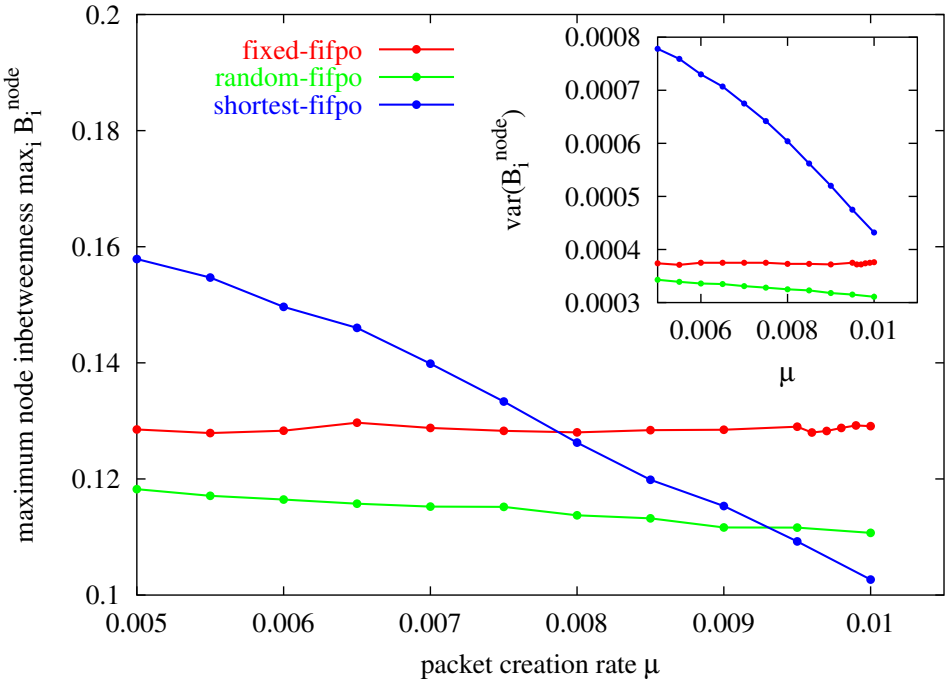

Figure 4.4: Effective inbetweenness for the shortest path algorithms
The maximum effective node inbetweenness ($\max_{i\in\mathcal{N}} B_i^{\text{eff}}$) defined in (4.2) for the shortest path algorithms depending on the packet creation rate μ is depicted. The simulations of $5 \cdot 10^5$ time steps have been performed on the constant-P reference network. The insets shows the corresponding variance for the nodes in $\mathcal{N}$.

bours tries to forward a packet to it. The fifpo algorithm then proceeds to the next packet in the buffer queue. Alternative shortest routes that omit the most sensible node are in that way favored because the probability of finding their nodes blocked is below the corresponding probability of the most sensible node. This effect results in a slight reduction of the maximum effective inbetweenness ($\max_{i\in\mathcal{N}} B_i^{\text{eff}}$) for higher packet creation rates μ.

As already visible by the simple modification from the fixed- to the shortest-fifpo routing there is a huge potential if information of the network state is used to construct more sophisticated routing algorithms. A possible approach is given in the following sections based on self-organisation and an adaptive learning process employing network state information.

4.2 Self-organised routing

For the previously introduced routing algorithms a central authority is in charge of the route discovery that provides all nodes in $\mathcal{N}$ with the necessary information. A decentralised routing protocol performs a route discovery algorithm at each node independently. In a so called link state algorithm [76] each node maintains a dynamical map of the complete network. Optimal shortest routes can now be calculated by the nodes employing a suitable algorithm, usually Dijkstra [57]. A periodical broadcast of the nodes' routing information to all other nodes in the network, called flooding, becomes necessary [77, 78]. As long as the network topology does not change frequently and the estimate of shortest paths does not rely on highly volatile network state observables the number of such flooding events is reasonable. The "open shortest path first" (OSPF) protocol [78], which is used in the internet, is based on such a link state routing.

A suitable approach to self-organised routing in wireless ad hoc networks should be able to adapt to local changes in the environment, but also to extensively employ information about the current network state. This would demand frequent communication about any changes in the network state to all nodes causing extensive flooding of the whole network and subsequent recalculations of the routing tables. Considering the fact that all information exchange has to happen in the multihop mode and taking into account the possibly high number of changes in a dynamical network topology an approach based on link state routing does not seem promising in the very first place.

The main focus is now drawn on a self-organised routing algorithm inspired by adaptive distance vector routing [75, 79, 24], that is modified for the specific use in wireless ad hoc communication networks.

For a distance vector routing, also known as Bellman-Ford distance vector routing [79], each node $i \in \mathcal{N}$ maintains a routing table $D^i_{f,j}$ where for each final receiver f an estimate of the cost of sending a packet to f denoted as $w_{i,f|j}$ is stored given that the packet is send to next hop node $j \in \mathcal{N}_i$. These routing tables are periodically updated by exchanging information between neighbouring nodes. The "routing information protocol" (RIP), formerly employed in the internet, uses such a Bellman-Ford distance vector routing algorithm [78, 79].

Littman and Boyen [75] proposed an asynchronous version of the Bellman-Ford distance vector routing algorithm based on reinforcement learning [80]. The routing policy tries to send a data packet from node i to a neighbouring node $j \in \mathcal{N}_i$ in order to reach the final destination f with minimum associated cost. The routing table $D^i_{f,j}$ of i can then be improved by employing a direct feedback if the packet has reached f. Following reinforcement learning [80] the estimate in $D^i_{f,j}$ can already be updated locally before the packet reaches its final receiver f. The concept employs the fact that the node $j \in \mathcal{N}_i$ is nearer to f and thus has a better estimate of the remaining cost $w_{j,f}$ to transmit the packet. The

asynchronous distance vector routing proposed by Littman and Boyen [75] is not based on the periodical exchange of cost estimates but couples them to actual data transmissions.

The particular organisation of data traffic in a wireless ad hoc communication network allows an even more elegant solutions. If a data transfer between two connected nodes i and j takes place both nodes have to block all their outgoing neighbours $\{(\mathcal{N}_i^{\text{out}} \setminus j) \cup (\mathcal{N}_j^{\text{out}} \setminus i)\}$ due to the medium access control protocol. This MAC-blocking requires the broadcasting of a "Keep-quiet" message that can be used for the spreading of the routing information.

The routing table $D_{f,j}^i$ of i can be updated if i receives a so called reduced routing table $\{\min_{j_2 \in \mathcal{N}_j} D_{f,j_2}^j\}$ from a bidirectional neighbour $j \in \mathcal{N}_i$. The broadcasted reduced routing table of j does only contain the lowest cost estimates for transmissions from j to any final receiver $f \in \mathcal{N}$. The update operation of node i reads as follows [76]:

$$D_{f,j}^i \leftarrow w_{i,j} + \min_{j_2 \in \mathcal{N}_j} D_{f,j_2}^j \tag{4.3}$$

$w_{i,j}$ is the measure of the cost to send a packet from i to its neighbour $j \in \mathcal{N}_i$. For the self-organised routing protocol studied in this section it is set to $w_{i,j} = 1$ which represents the hop distance between the two nodes. With this definition the estimated cost $w_{i,f|j} = D_{f,j}^i$ of a packet delivery from i to f is just the number of hops from i to f on a given route specified by the next hop node $j \in \mathcal{N}_i$. It has been shown that this process converges in finite time to shortest paths [78]. The full ability of self-organised routing approach will be visible by including information about the network state in the cost estimates $w_{i,j}$ in the next section. For demonstrating the basic functionality within this section the distance between two nodes is the only parameter on which to decide which particular route is taken.

The converged routing tables $D_{f,j}^i$ do not only contain the information about shortest paths from i to f but they also have information about the length of the shortest paths from any of the bidirectional neighbours $j \in \mathcal{N}_i$ to f. Thus a node now has the ability to not only choose a shortest but also a second shortest path. This choice is paid respect to by assigning a Boltzmann-like probability [80] to each of the possible routes represented by all neighbours of i. Trying to send a packet in the fifpo mode from a node i, node $j \in \mathcal{N}_i$ is chosen as a next hop with probability

$$p(j|i \to f) = \frac{(D_{f,j}^i)^{-\beta}}{\sum_{j' \in \mathcal{N}_i} (D_{f,j'}^i)^{-\beta}} \quad . \tag{4.4}$$

The parameter β acts like an inverse temperature. For high values of β the probability of choosing other paths than the shortest is significantly reduced. In the limit of $\beta \to \infty$ the routing scheme can be mapped on the random-fifpo modus described in the previous section.

As mentioned above within the self-organised routing the routing information

is coupled to the MAC-blocking signal in order to not cause additional communication. In particular the reduced routing table $\{\min_{j\in\mathcal{N}_i} D^i_{f,j}\}$ of the sending/receiving node i is broadcasted with the MAC signal. The bidirectional neighbours of the sending and the receiving node do subsequently update their routing tables according to (4.3).

For the simulation employed to verify the performance of the self-organised routing all routing tables are initially set to

$$D^i_{f,j} = \left\{ \begin{array}{ll} 1 & (f \in \mathcal{N}_i) \\ 0 & (i = f) \\ \infty & (\text{else}) \end{array} \right\} \quad . \tag{4.5}$$

After a sufficient number of updates the routing tables $D^i_{f,j}$ of all nodes $i \in \mathcal{N}$ converge towards a fixed configurations strictly representing shortest paths

$$D^i_{f,j} = w_{i,j} + w_{j,f} = 1 + d_{j,f} \tag{4.6}$$

where $d_{j,f}$ is the shortest path distance from j to f. At this point further updates according to (4.3) do not cause changes in the routing tables $D^i_{f,j}$. This behaviour was verified in the data traffic simulations. Convenient measures are the number of end-to-end routes, that have so far not converged to the shortest path $\sigma_{i,f}$, referred to as imperfectness, and the number of changes in the routing tables $D^i_{f,j}$ per ten time steps. Both observables are shown in Figure 4.5 for networks of size $N = 100$ and $N = 200$ and different packet creation rates μ. The simulations were performed on networks with constant-P power assignment with $k^{\text{constP}} = 24$. As indicated in both graphs, the convergence time for the routing tables to reach the shortest path representation is shifted to higher time scales for higher values of the system size N. This is caused by the higher diameter $D \sim \sqrt{N}$, requiring more propagation steps for the complete distribution of any information through the network. It is also intuitive that for a given network size N a higher packet creation rate μ causes a faster convergence of the routing tables due to the fact that more packets in the network result in more sending events that are themselves coupled to the broadcasting of the routing information.

Another critical parameter is the inverse temperature β for the Boltzmann-like probability distribution in equation (4.4). The lower the value of β the higher is the probability to choose a path that is not the shortest. In the generic data traffic simulations higher values of β proved to be superior to lower ones (see Figure 4.6). This might be understood by looking at the final configuration of the routing table $D^i_{f,j}$. The values for $(\min_{j\in\mathcal{N}_i} D^i_{f,j})$ representing the shortest paths from i to f and $(\max_{j\in\mathcal{N}_i} D^i_{f,j})$ can only differ by a maximum of 2. This can be easily proved: The distance from any neighbour $j^* \in \mathcal{N}_i$ to any node f can not be larger than $(\min_{j\in\mathcal{N}_i} D^i_{f,j} + 1)$ because node i itself represents the next hop of node j^* to reach f in $(\min_{j\in\mathcal{N}_i} D^i_{f,j} + 1)$ hops. Thus, it follows for the reduced routing table

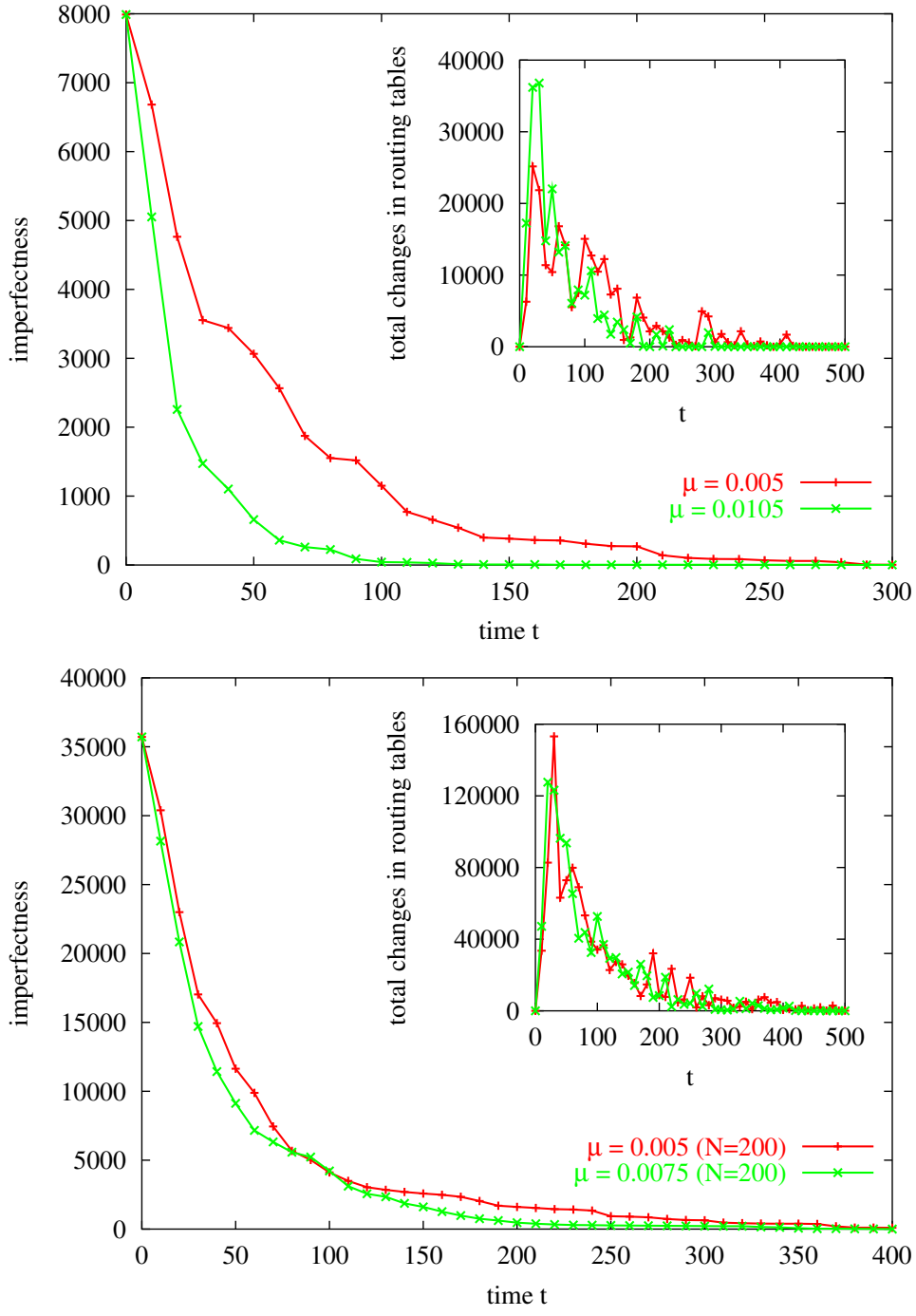

Figure 4.5: Convergence of the routing tables

The upper graph corresponds to a network with 100 nodes, the lower one has 200 nodes, both are constructed with constant-P power assignment with $k^{\mathrm{constP}} = 24$. The inverse temperature is set to $\beta = 100$. The imperfectness indicates, how many end-to-end routes i to f in $D^i_{f,j}$ are not represented by at least one shortest route $\sigma_{i,f}$. The insets show at how many positions the routing tables $D^i_{f,j}$ have been improved in the last ten time steps.

from node j^* that $(\min_{j_2 \in \mathcal{N}_{j^*}} D^{j^*}_{f,j_2}) \leq (\min_{j \in \mathcal{N}_i} D^i_{f,j} + 1)$ which following equation (4.3) causes $D^i_{f,j^*} = (\min_{j \in \mathcal{N}_i} D^i_{f,j} + 2)$. This implies that signing a significant probability to paths with $D^i_{f,j} = (\min_{j \in \mathcal{N}_i} D^i_{f,j} + 2)$ encourages the sending of data packets to next hop nodes j that might just send the packet back after a couple of time steps because node i is already the best suitable base for the journey to f. This effect can only be reduced by higher values of β.

From the perspective to avoid congestion it might still seem reasonable to opt for intermediate values of the inverse temperature β that avoids overloading of shortest paths and uses alternative second shortest routes instead. Generic data traffic simulations disprove that assumption. Self-organised routing schemes with higher values of β are always superior to lower ones as indicated in Figure 4.6. The mean of the end-to-end time delay $\langle t_{e2e} \rangle$ over all packets of a $5 \cdot 10^5$ time steps simulation is shown depending on the packet creation rate μ. The simulations were performed on a constant-P reference network. One has to draw the conclusion that the choice of non-shortest paths does not happen in an intelligent way but randomly. In most cases the choice of a longer path is not suitable because the routes show little congestion.

As already mentioned the self-organised routing with an exponent $\beta \to \infty$ in (4.4) only discriminates between different routes if they represent shortest paths. According to (4.4) they are chosen with the same probability. In that respect the routing decision is identical to the random-fifpo case from the previous section. Both routing schemes show the same behaviour under generic data traffic conditions.

4.3 Self-organised adaptive routing

Employing the self-organised routing introduced in the previous section a node i does not gather essentially more or better information about the network structure compared to the centralised shortest path routing approaches presented in section 4.1. Including information about the recent or current dynamical state of the network seems to be the key element of any advanced routing approach [76, 77, 78]. Only by employing knowledge about highly frequented nodes or areas, travelling times on certain paths or quality of links the management of an "intelligent" packet forwarding can be achieved.

Imagine a typical routing decision: A node $i \in \mathcal{N}$ intends to send a packet to the final receiver $f \in \{\mathcal{N} \setminus i\}$. If f is contained in the set of neighbouring nodes $\mathcal{N}_i$ of i the packet will be directly forwarded to f. If $f \notin \mathcal{N}_i$, i can send the packet to any of its neighbouring nodes $j \in \mathcal{N}_i$. How should node i decide where to send the packet? Before answering that question one has to define, which criteria shall be applied for comparing different possible paths represented by the next hop nodes $j \in \mathcal{N}_i$. Within wireless multihop ad hoc communication

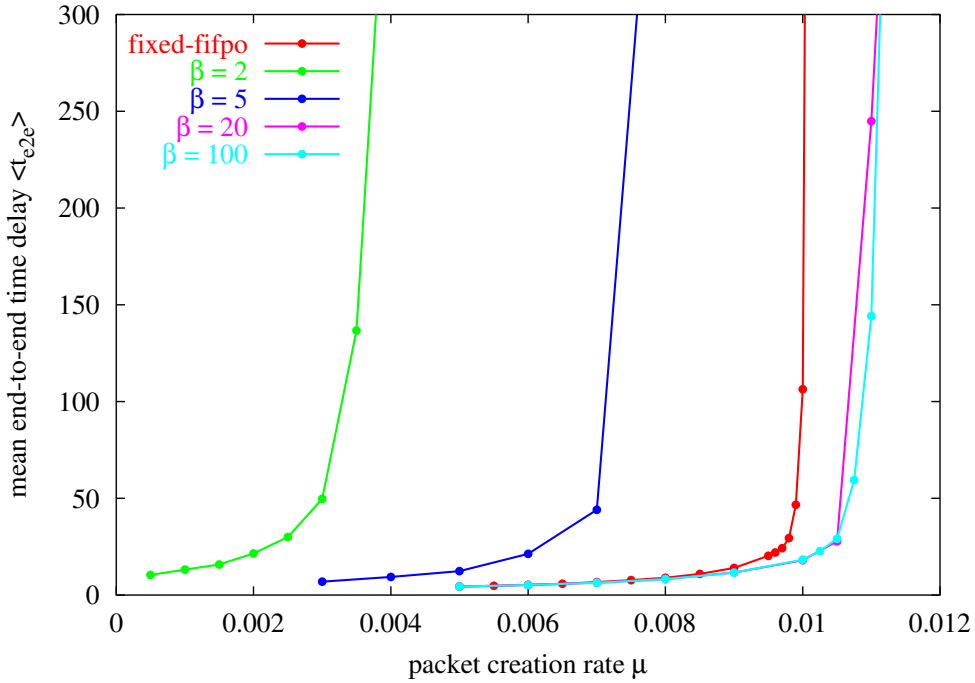

Figure 4.6: End-to-end time delay for the self-organised routing
The mean of the end-to-end time delay $\langle t_{e2e} \rangle$ over all packets of a $5 \cdot 10^5$ time steps simulation is illustrated for different inverse temperatures β depending on the packet creation rate μ. A constant-P reference network of 100 nodes with $k^{\text{constP}} = 24$ served as underlying network graph.

the most important feature seems to be the minimisation of the end-to-end time delay. Short delay times are essential, especially for voice transmission as in mobile phones. According to Little's Law (3.12) the minimisation of the end-to-end time delay also reduces the number of active packets in a network which in general leads to systems with a lower risk of congestion.

A promising approach has to obtain a routing policy which balances the minimisation of the number of "hops" a packet will take with the possibility of congestion along popular routes [76]. Since the buffer queue length n_i of a node is a good measure of "how busy" it is [75, 76], n_i is directly incorporated for the cost estimation of a communication route. In a small modification of the self-organised routing scheme the cost estimate $w_{i,j}$ for sending a data packet from node i to its neighbouring node $j \in \mathcal{N}_i$ in (4.3) now yields

$$w_{i,j} = n_j \pm 1 + 1 \quad . \tag{4.7}$$

The + sign is used if j acts as receiver in the upcoming communication and the

$-$ sign if n is the sender. The update operation of the routing tables in (4.3) is now composed of the buffer queue length of the neighbour $j \in \mathcal{N}_i$ represented by $(n_j \pm 1)$ in equation (4.7), the hop distance to j represented by the $(+1)$ term in the same equation and the equivalent best cost estimate of node j for a transmission to f represented by $(\min_{j_2 \in \mathcal{N}_j} D^j_{f,j_2})$. The summarised buffer queue length of the intermediate nodes are used as a measure of how long a communication via a certain path might take. Despite the use of the first-in-first-possible-out strategy this assumption seems reasonable since in general a longer buffer queue results in a longer waiting at the particular node. Coupling the broadcasting of the reduced routing table $\{\min_{j_2 \in \mathcal{N}_j} D^j_{f,j_2}\}$ of j and the subsequent update of its neighbours $j_2 \in \mathcal{N}_j$ with the MAC-blocking results in an immediate information of the local neighbourhood about any changes in the buffers of node j caused by a transmission event.

The update rule defined in (4.3) modifies the routing tables of the local neighbourhood $\{\mathcal{N}_i \cup \mathcal{N}_j\}$ for a transmission between i and j but not the routing tables of further distant nodes. These nodes can only obtain information about more than one hop away communications if their neighbours broadcast their already updated reduced routing tables when they are themselves involved in a communication event. This results in a delay as the update information propagates through the whole network MAC-step for MAC-step. An estimate of the time scale of such a propagation is given in Figure 4.5. For a network without any prior knowledge about possible routes except for the initial configuration (4.5), the time for the convergence to the shortest path in the self-organised routing scheme can be approximated depending on packet creation rate μ and system size N. A local change in the network state is somewhat similar to a new network configuration all nodes have to learn about. As long as this time range is not significantly larger than the time scales on which the network experiences tremendous changes in its load states a routing algorithm based on dynamical observables should provide sufficient results.

A clear advantage of the self-organised adaptive routing is the coupling of the broadcasting of the reduced routing table $\{\min_{j \in \mathcal{N}_i} D^i_{f,j}\}$ of node i with the MAC-blocking signal that allows all neighbours $j \in \mathcal{N}_i$ to quickly learn about changes of i's buffer queue n_i. It is precisely this set of neighbours $j \in \mathcal{N}_i$ that can themselves forward own data packets to i. Ensuring that every node $j \in \mathcal{N}$ has access to information about its bidirectional neighbours which is provided by the coupling to the MAC signal is a valuable tool especially for the local management of traffic.

Unlike the self-organised routing introduced in the previous section, the self-organised adaptive routing described above does not result in fixed routing tables for any node $i \in \mathcal{N}$. The self-organised routing converges towards shortest paths in terms of the hop-metric as indicated in equation (4.6). Because of the inclusion of the buffer queue content n_j via equation (4.7) in the routing tables $D^i_{f,j}$ of the self-organised adaptive routing scheme such a convergence is not reached.

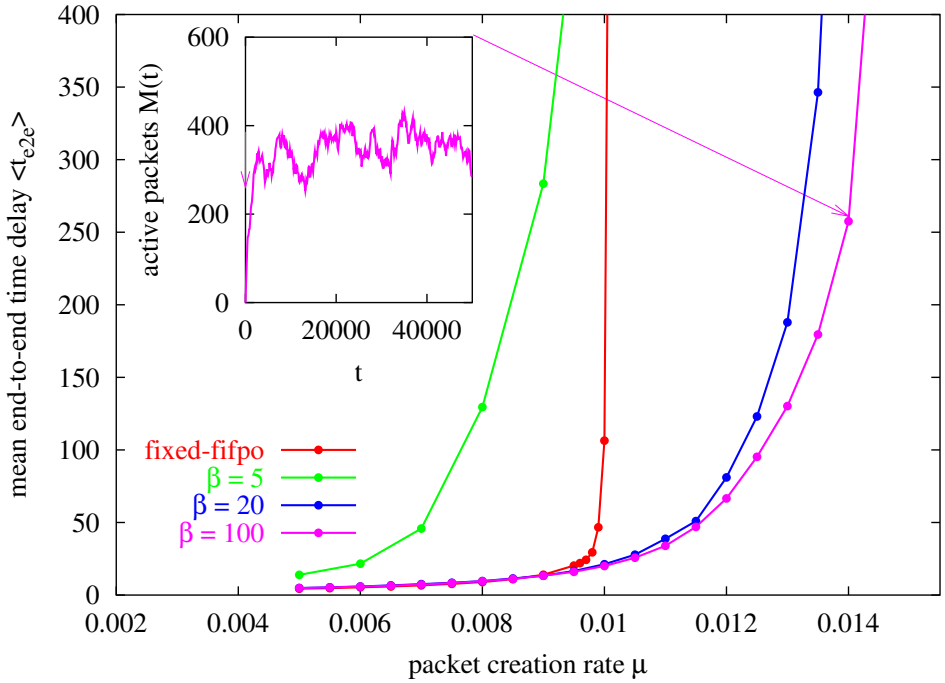

Figure 4.7: End-to-end time delay for the self-organised adaptive routing The mean of the end-to-end time delay $\langle t_{e2e} \rangle$ is illustrated for different inverse temperatures β depending on the packet creation rate μ. The average is taken over all packets of a $5 \cdot 10^5$ time steps simulation on a constant-P reference network. The inset shows the number of active packets $M(t)$ for the case $\beta = 100$ at a packet creation rate $\mu = 0.014$ depending on time t. The observation that there is no linear dependence of $M(t)$ on the time step t is a clear sign that the network is still in the subcritical regime.

Every sending event in the network causes the propagation of the modified buffer queues which enforces changes in the routing tables of the neighbouring nodes and subsequently their further surroundings. The fluctuations in the length of the buffer queues imply a non-steady state for the routing tables of all nodes of the network.

Mean end-to-end time delay The actual choice of a path for a packet to be sent from i to f is determined according to the probability for each next hop node $j \in \mathcal{N}_i$ given by (4.4). The mean end-to-end time delay $\langle t_{e2e} \rangle$ of the self-organised adaptive routing shows a strong dependence on the inverse temperature β as indicated in Figure 4.7. The simulations of the generic data traffic were performed on a constant-P reference network. The average of the end-to-end

time delay was taken over all packets transmitted within the $5 \cdot 10^5$ time steps of each simulation run. For reference the respective plot of the fixed-fifpo regime is included in the graph. Two things should be observed: First, higher inverse temperatures β in equation (4.4) are superior to lower ones. This behaviour is in accordance with the results from self-organised routing in Section 4.2. High inverse temperatures β ensure that there exists almost no probability of choosing another path than the one with the lowest associated cost estimate. Second, the self-organised adaptive routing with high β causes a significant shift in the critical packet creation rate μ_{crit}. Although more packets are present in the network, such a routing scheme can handle packet creation rates up to a factor of around 1.4 compared to the fixed-fifpo case. The inset of Figure 4.7 shows the number of active packets $M(t)$ for the self-organised adaptive routing with $\beta = 100$ at a packet creation rate $\mu = 0.014$ as the arrow indicates. The observation that there is no linear dependence of $M(t)$ on the time step t is a clear sign that the network is still in the subcritical regime.

Interarrival and sending time In order to get an impression about the networks behaviour under data traffic governed by self-organised adaptive routing selected single node observables, already introduced in Section 3.3, have been studied. Figure 4.8 shows a plot of the probability distribution for the interarrival and the sending time for the most sensible node of the constant-P reference network of 100 nodes. The simulations spanned $5 \cdot 10^5$ time steps and were performed for different subcritical packet creation rates μ. The inverse temperature was set to $\beta = 100$. As already mentioned the graph theoretical inbetweenness B_i^{node} introduced in Section 3.1.2 is not longer suitable to characterise the influx of packets to the nodes of the network because the particular routing decisions do not necessarily employ shortest paths. The probability distributions for both the interarrival times $p(t_i^{\text{arrive}})$ and the sending times $p(t_i^{\text{send}})$ show an exponential behaviour which can be described as a geometric probability density function

$$p(t = t^*) = (1 - \eta)^{t^* - 1} \eta \tag{4.8}$$

in accordance with the findings for the fixed-fifpo case in section 3.3.1. In complete analogy the values for η can be obtained from the mean of the interarrival times $\eta = \mu_i^{\text{in}} = 1/\langle t_i^{\text{arrive}} \rangle$ and the mean of the sending times $\eta = \mu_i^{\text{out}} = 1/\langle t_i^{\text{send}} \rangle$. The corresponding approximations are included in Figure 4.8. Parameters are given in Table 4.1. Since a graph theoretical description of the mean interarrival time $\langle t_i^{\text{arrive}} \rangle$ based on the node inbetweenness B_i^{node} is no longer possible neither μ_i^{in} nor μ_i^{out} can be expressed analytically.

Compared to Figure 3.8 the corresponding curves in Figure 4.8 show a small peak for low values of the interarrival times that is especially pronounced for lower packet creation rates μ. It can be understood in terms of the routing algorithm for

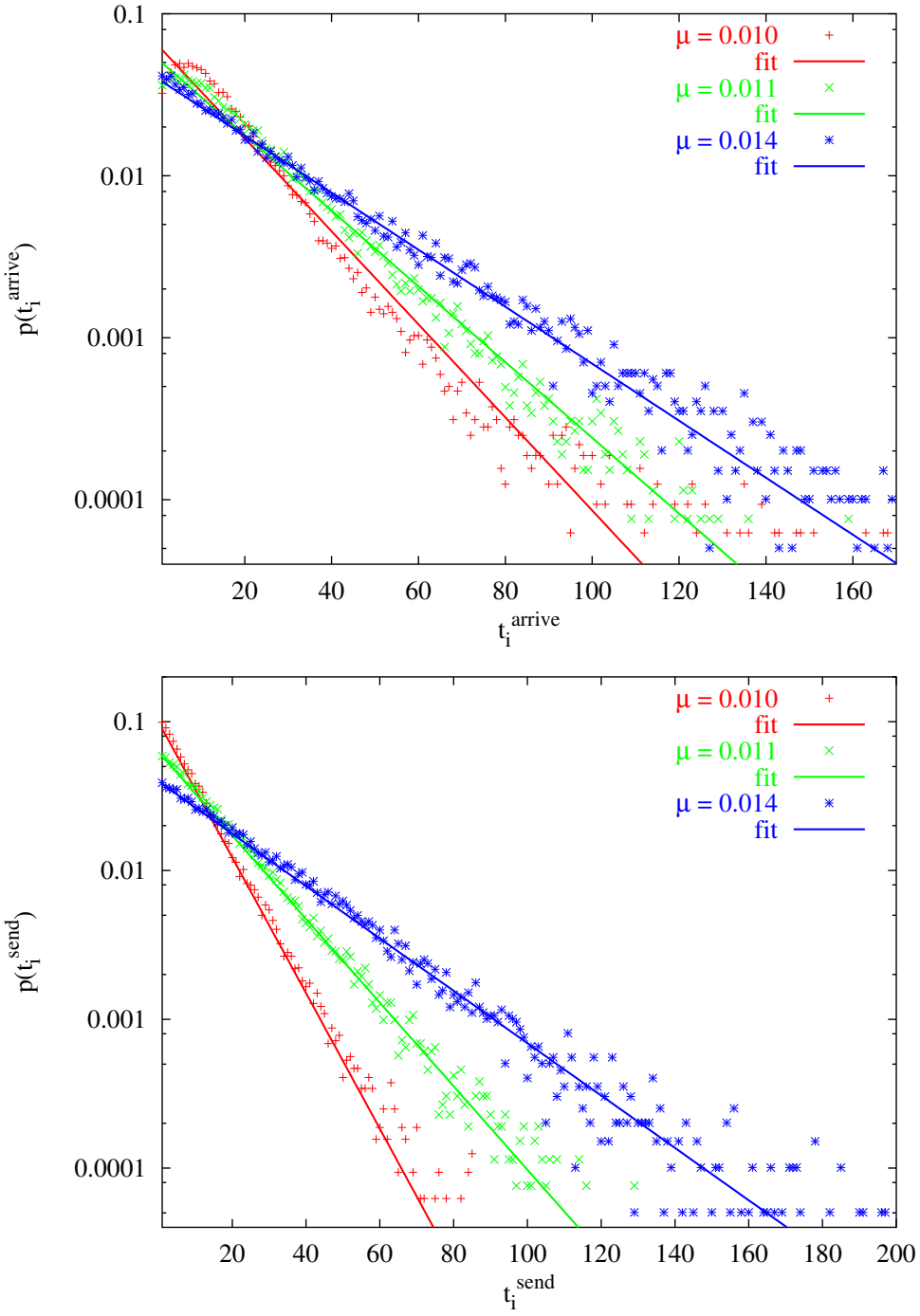

Figure 4.8: Probability distribution of interarrival- and sending time
The probability distribution for the interarrival (upper graph) and the sending time (lower graph) for the most sensible node of the constant-P reference network is depicted. The simulations covered $5 \cdot 10^5$ time steps and were performed for different subcritical packet creation rates μ, indicated by different colours. The inverse temperature of the self-organised adaptive routing was set to $\beta = 100$. A fit according to (4.8) is included for each data set.

networks with few packets. As soon as a node i receives a packet its neighbourhood gets the corresponding message due to the MAC-blocking signal. This effect causes an update of the routing tables with an increased cost estimate for transfers via i and so reduces the probability of sending another packet there. After node i itself could forward the packet it is again a good alternative for packet transfer and receives a higher probability of gaining new packets. For higher packet creation rates μ the average number of packets in the network $M(t)$ is higher which makes routing decisions not too sensitive for single packets events.

μ	$\langle t_i^{\text{arrive}} \rangle$	$\langle t_i^{\text{send}} \rangle$
0.0010	15.616	10.037
0.0011	19.066	15.971
0.0014	25.204	25.192

Table 4.1: Mean interarrival and sending time
Mean values of the interarrival and sending time, $\langle t_i^{\text{arrive}} \rangle$ and $\langle t_i^{\text{send}} \rangle$, as used for the fit in Figure 4.8 for different packet creation rates μ. The node degree of node i is measured to be $k_i = 25$. Since for $\mu \to \mu^{\text{crit}}$ it competes with all its neighbours for sending permission it is intuitive that the mean sending time is in the same range $\langle t_i^{\text{send}} \rangle \approx 25$.

Buffer queue length Significant changes can be observed in the probability distribution for the buffer queue length $p(n_i)$ compared to the fixed-fifpo case studied in Section 3.3.2. For packet creation rates μ closer to μ^{crit} the graphs of the probability distribution get bell shaped as presented in Figure 4.9. The most sensible node of the constant-P reference network was observed. The simulations spanned $5 \cdot 10^5$ time steps. The inverse temperature of the self-organised adaptive routing was fixed to $\beta = 100$. In networks with $\mu \to \mu^{\text{crit}}$ the node's buffer queue n_i always contains a certain number of packets fluctuating around a given mean $\langle n_i \rangle$ that shows a clear dependence on μ. This almost stable distribution of packets is a key stone for the functionality of the routing policy that emerges in self-organised manner.

Good fits of the experimental data can be obtained by a continuous Gamma-distribution with scaling parameter b

$$p_{(b,k)}(x) = \frac{1}{\Gamma(k) b^k} x^{k-1} e^{-\frac{x}{b}} \quad . \tag{4.9}$$

The inset in Figure 4.9 shows an fit for $\mu = 0.013$. Parameters are $k = 10.019$ and $b = 0.487$. For $k \to 1$ the Gamma-function converges towards an exponential

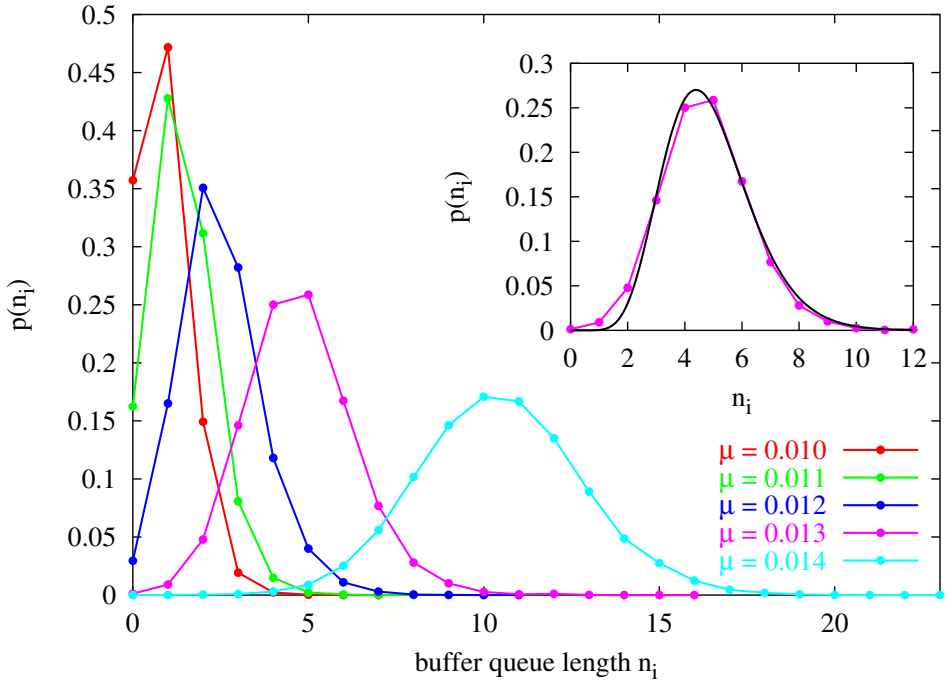

Figure 4.9: Probability distribution of the buffer queue length
The probability distribution of the buffer queue length $p(n_i)$ for the most sensible node of the constant-P reference network is shown. The data traffic simulations spanned $5 \cdot 10^5$ time steps and were performed for different subcritical packet creation rates μ, indicated by different colours. The inverse temperature of the self-organised adaptive routing was set to $\beta = 100$. The inset illustrates a fit with a Gamma-distribution according to (4.9) for $\mu = 0.013$. Parameters are $k = 10.019$ and $b = 0.487$.

probability distribution function which is the continuous analog to the the geometric probability distribution used in equation (3.20) for the description of the buffer queue lengths in Section 3.3.2. In that respect the exponential behaviour in the fixed-fifpo case might be a special case of a more general description for the probability density function of the buffer queue length $p(n_i)$.

Single-node temporal correlations The rather narrow interval where the probability density function of the buffer queue length $p(n_i) \neq 0$ is the key feature for the understanding of the single-node temporal correlations $r_i(\Delta t)$ as defined in (3.27) for packet creation rates near the critical value μ^{crit}. Unlike the fixed-fifpo routing where the probability density function of the buffer queue length $p(n_i)$ is significantly skewed, in the self-organised adaptive routing the values of n_i are

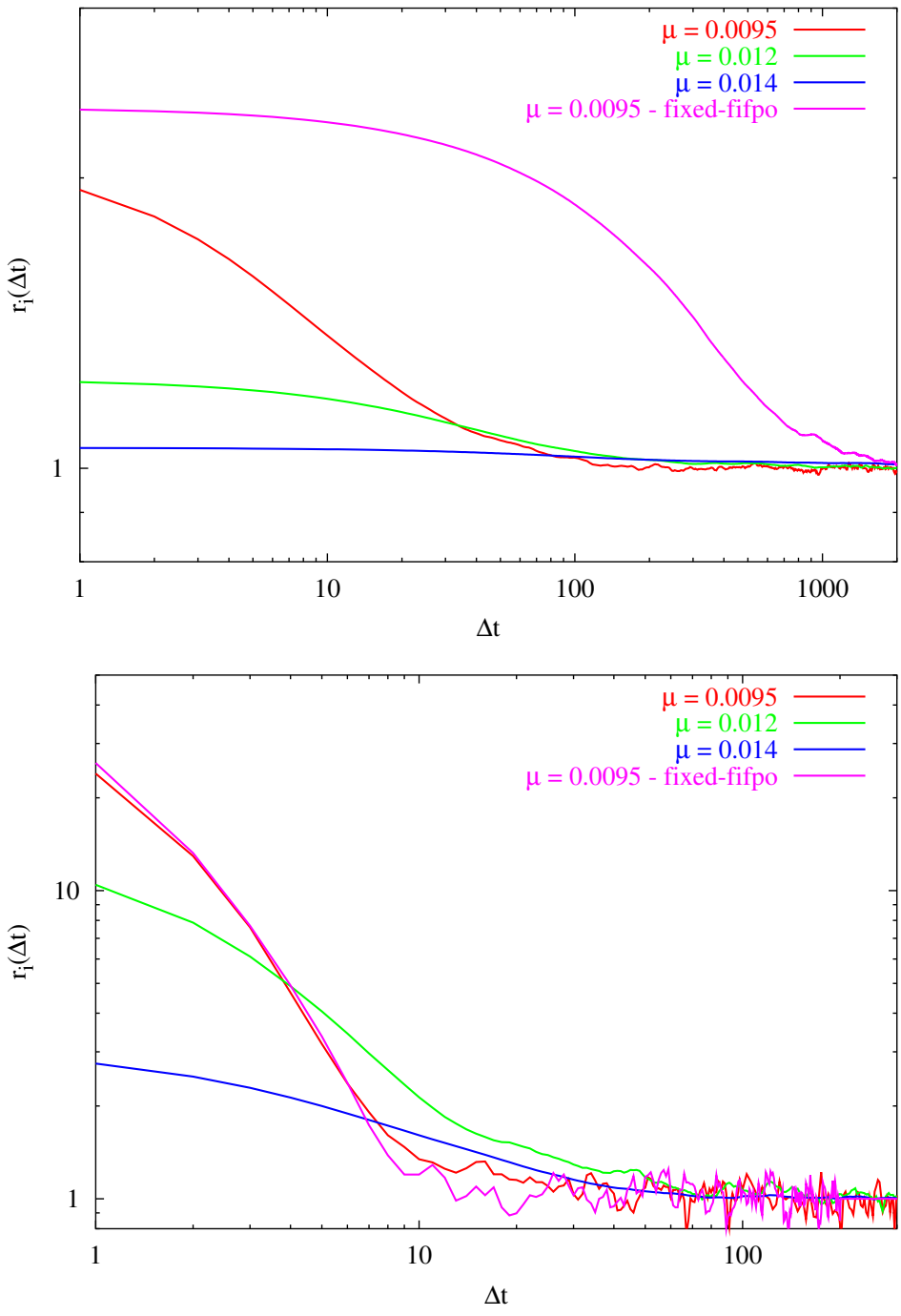

Figure 4.10: Single-node temporal correlations

Single-node temporal correlations $r_i(\Delta t)$ for self-organised adaptive routing are shown for different packet creation rates. Simulation parameters are identical to Figures 4.8 and 4.9. The upper graph shows $r_i(\Delta t)$ for i being the most sensible node of the network with 25 communication neighbours. In contrast $r_i(\Delta t)$ is also illustrated for a less used node with only 8 neighbours in the lower graph. The corresponding single-node temporal correlations for these nodes in the fixed-fifpo case are presented as reference.

centered around a mean value $\langle n_i \rangle$, at least for the more sensible nodes as shown in Figure 4.9. This causes $\langle n_i(t+\Delta t) n_i(t) \rangle \approx \langle n_i(t) \rangle^2$ for smaller Δt. Thus the single-node temporal correlations are expected to be drastically reduced. The upper graph of Figure 4.10 verifies that result where the curve for the fixed-fifpo case near its critical packet creation rate $\mu_{\text{ff}}^{\text{crit}}$ is shown for reference. The simulations of $5 \cdot 10^5$ time steps were performed on the constant-P reference network. In the upper graph the single-node temporal correlations are studied for the same most sensible node that was observed for the interarrival/sending times and the buffer queue length in Figures 4.8 and 4.9, respectively. For this node the transition to the bell-shaped gamma-like density function of the buffer queue length $p(n_i)$ is most pronounced and causes the significantly reduced correlations. In the lower graph of Figure 4.10 a less frequented node was observed. This node only has eight communication neighbours compared to 25 for the most sensible node. In classical shortest path routing this node does handle almost no traffic that does not originate or terminate at the node itself. Also in the case of higher packet creation rates μ that might occur in a network governed by the self-organised adaptive routing this node is not extensively used as an alternative router nor does its performance be a limiting factor of the network's overall behaviour. In that respect the statistical properties of the node are not significantly changed by the different routing scheme. This can be verified in the corresponding plot of the single-node temporal correlation where the curves for the packet creation rate $\mu = 0.0095$ in the fixed-fifpo and the self-organised adaptive routing almost perfectly match.

Node inbetweenness In extension to Figure 4.4 a more detailed picture is provided by Figure 4.11 including the corresponding measures of the effective node inbetweenness for the self-organised adaptive routing (see definition in equation (4.2)). The generic data traffic simulation covered $5 \cdot 10^5$ time steps and was based on the constant-P reference network. In the upper graph of Figure 4.11 the maximum value of the effective node inbetweenness ($\max_{i \in \mathcal{N}} B_i^{\text{eff}}$) is plotted for different packet creation rates μ. Up to around the critical packet creation rate for the fixed-fifpo routing $\mu_{\text{ff}}^{\text{crit}}$ the maximum effective node inbetweenness ($\max_{i \in \mathcal{N}} B_i^{\text{eff}}$) shows a similar behaviour like the shortest-fifpo routing. The more packets the most sensible node would have to serve the more packets are routed on alternative routes. For $\mu > \mu_{\text{ff}}^{\text{crit}}$ the maximum effective node inbetweenness ($\max_{i \in \mathcal{N}} B_i^{\text{eff}}$) increases. It has to be suspected that at that point other nodes than the most sensible one get near to a congested regime, too, which makes them less attractive for rerouting. The most sensible node is then again a good alternative in spite of a possible longer waiting time there. For other network configurations not just a single peak in the interval $\mu_{\text{ff}}^{\text{crit}} < \mu < \mu_{\text{soa}}^{\text{crit}}$ between the critical rate of the fixed-fifpo and the self-organised adaptive routing was observed but two of

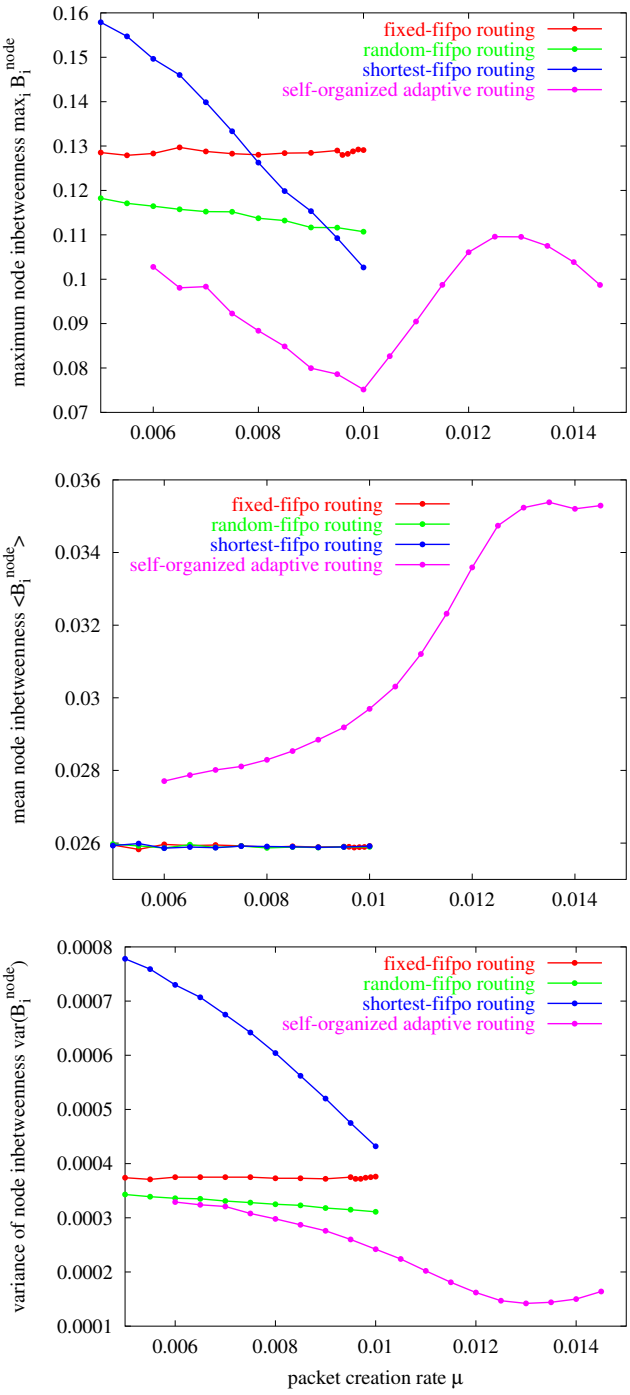

Figure 4.11: Effective inbetweenness for different routing approaches
The maximum ($\max_{i\in\mathcal{N}} B_i^{\text{eff}}$), the mean $\langle B^{\text{eff}}\rangle$ and the variance $\text{var}_{i\in\mathcal{N}}(B_i^{\text{node}})$ of the effective node inbetweenness are presented in the three graphs depending on the packet creation rate μ. The data is obtained from generic data traffic simulations on a constant-P reference network each covering $5 \cdot 10^5$ time steps. The red curves corresponds to the fixed-fifpo routing, the green to the random-fifpo, the blue to the shortest-fifpo and the pink to the self-organised adaptive routing.

them. This could be an indication that two nodes play a significant role in the limitation of the network's performance.

The middle graph of Figure 4.11 shows the mean of the effective node inbetweenness $\langle B_i^{\text{eff}} \rangle$ averaged over all nodes in $\mathcal{N}$. For the shortest path algorithms the number of hops for each packet is fixed thus determining the total number of inflows per time step given by $\sum_{i \in \mathcal{N}} \mu_i^{\text{in}} = \mu N D$. It follows

$$\langle B_i^{\text{eff}} \rangle = \frac{\langle \mu_i^{\text{in}} \rangle}{\mu N} = \frac{\sum_{i \in \mathcal{N}} \mu_i^{\text{in}}}{\mu N^2} = \frac{\mu N D}{\mu N^2} = \frac{D}{N} = \text{const} \tag{4.10}$$

for a given network realisation. This behaviour is verified for the three shortest path routing schemes as indicated in Figure 4.11. The higher mean for the effective node inbetweenness $\langle B_i^{\text{eff}} \rangle$ of the self-organised adaptive routing is due to the fact that not only shortest paths in terms of the hop metric are chosen. This causes a higher number of packet inflows per time step but does not cause a higher mean end-to-end time delay $\langle t_{e2e} \rangle$ as seen in Figure 4.7. For packet creation rates near the critical packet creation rate of the self-organised adaptive routing $\mu \to \mu_{\text{soa}}^{\text{crit}}$ the effective node inbetweenness $\langle B^{\text{eff}} \rangle$ saturates. This is a possible indicator for reaching the critical creation rate $\mu_{\text{soa}}^{\text{crit}}$ since a higher influx of packets to the nodes $\mathcal{N}$ can not be handled anymore.

An analog interpretation to the one given in Section 4.1 holds for the decrease of $\text{var}_{i \in \mathcal{N}}(B_i^{\text{eff}})$ in the lower graph of Figure 4.11. The routing algorithm tends to a more homogeneous distribution of the packets on all nodes of the network. If more central nodes get longer buffer queues alternative routes get more and more likely to be taken.

A pleasant visualisation is provided by Figure 4.12 which is set up in complete analogy to Figure 3.1. The usual constant-P reference network of 100 nodes with the constant-P rule ($k^{\text{constP}} = 24$) is shown. The two figures only differ thus far that in Figure 3.1 the graph theoretical node inbetweenness B_i^{node} from (3.2) is used for the colouring of the nodes whereas in Figure 4.12 the effective node inbetweenness B_i^{eff} from (4.2) is used for a self-organised adaptive routing. Since the effective node inbetweenness B_i^{eff} represents the "real" flow of data in the network, which for the self-organised adaptive routing is significantly different compared to the fixed-fifpo routing, one gets an impression of "how" data traffic is rerouted. In Figure 4.12 more nodes have red and blue colour, indicating that they are frequently used. Note that the scale of the node inbetweenness changed compared to Figure 3.1. Even the most sensible nodes have a smaller inbetweenness than in the fixed-fifpo case.

Probability distribution of the end-to-end time delay Comparing the self-organised adaptive routing to the formerly used shortest path routing algorithms it is worth to have a look on the probability distribution of the end-to-end time

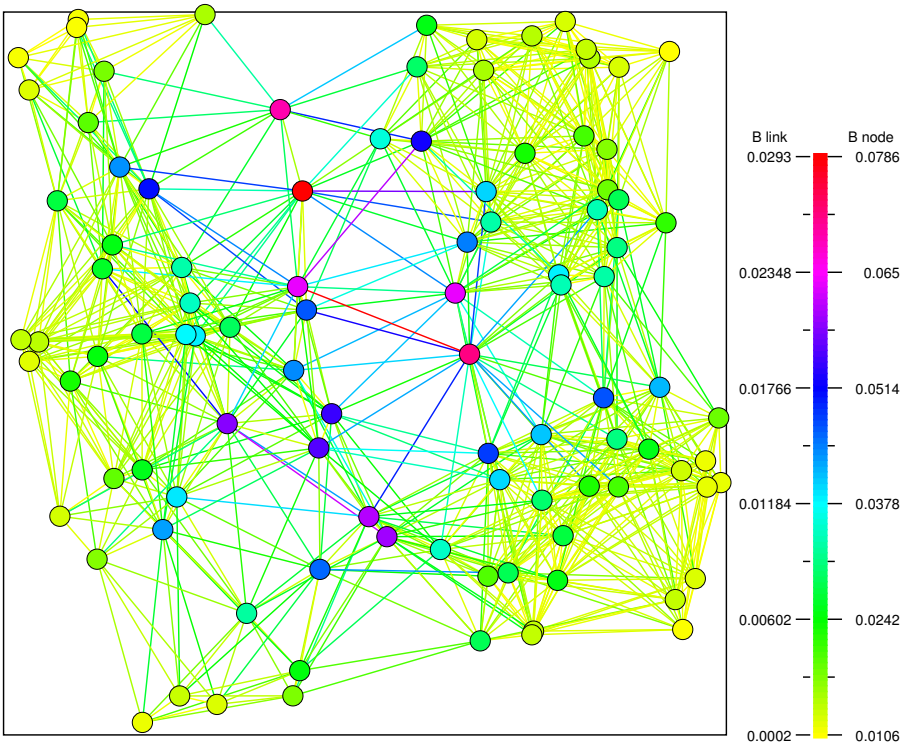

Figure 4.12: Visualisation of the effective node inbetweenness
The colour of the nodes correspond to the effective node inbetweenness B_i^{eff}. The network graph of 100 nodes with constant-P power assignment ($k^{\text{constP}} = 24$) is the usual constant-P reference network. The effective node inbetweenness was measured at $\mu = 0.0095$.

delay $p(t_{e2e})$ in Figure 4.13. The simulations on the constant-P reference network covered $5 \cdot 10^5$ time steps. It is important to make sure that an alternative routing algorithm also works for lower packet creation rates meaning that one does not have to pay for a good performance in the high traffic regime with worse conditions in the low traffic regime. As observed for the mean delay times $\langle t_{e2e} \rangle$ in Figure 4.7 the self-organised adaptive routing works as well as the other approaches in the range $\mu < \mu_{\text{ff}}^{\text{crit}}$. Although small differences in the probability distribution of the end-to-end time delay $p(t_{e2e})$ at $\mu = 0.0095$ can be seen between the fixed-fifpo and the self-organised adaptive routing in Figure 4.13 the overall shape of the curves match quite well. While the fixed-fifpo routing might show a slightly better performance towards short end-to-end time delays the self-organised adaptive routing reduces the probability for finding packets with a very large end-to-end time delay. For higher packet creation rates $\mu > \mu_{\text{ff}}^{\text{crit}}$ the mean delay times $\langle t_{e2e} \rangle$ of the fixed-fifpo routing becomes a linear function of time according to Little's Law

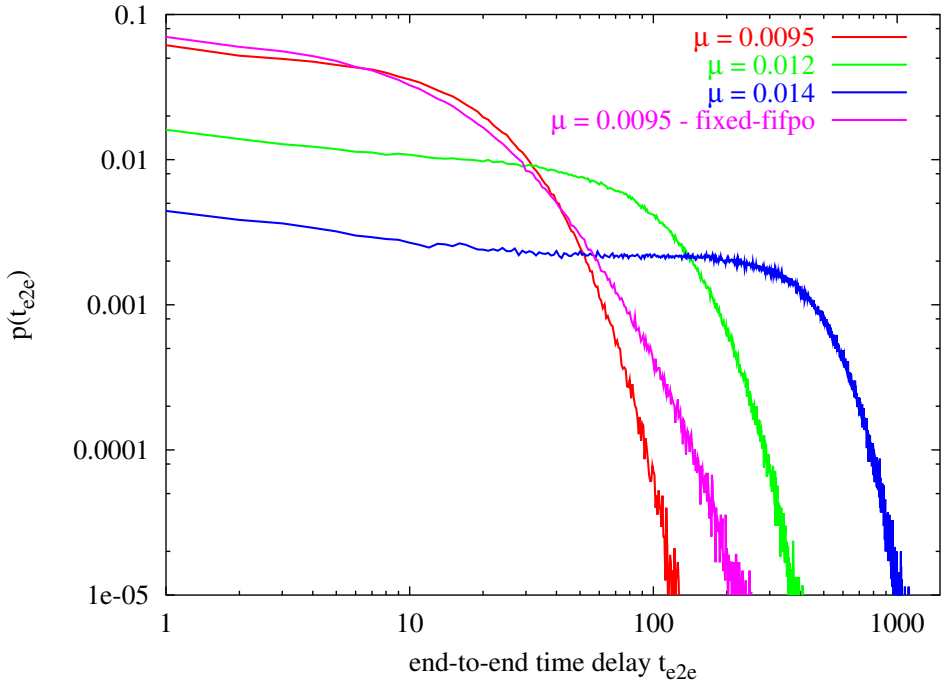

Figure 4.13: Probability distribution of the end-to-end time delay
The probability distribution of the end-to-end time delay $p(t_{e2e})$ is presented for the self-organised adaptive routing for different packet creation rates μ. As reference the corresponding data set for the fixed-fifpo routing at $\mu = 0.0095$ is included. The data is obtained from generic data traffic simulations on the constant-P reference network each covering $5 \cdot 10^5$ time steps.

(3.12). This causes the probability density function of the end-to-end time delay $p(t_{e2e})$ to not converge to 0 for $t_{e2e} \to \infty$. In this critical regime $\mu_{\text{ff}}^{\text{crit}} < \mu < \mu_{\text{soa}}^{\text{crit}}$ the self-organised adaptive routing can still handle the upcoming data traffic although longer end-to-end delay times have to be accepted (see Figures 4.7 and 4.13).

4.4 Self-organised adaptive routing using reinforcement learning

So far the routing table $D^i_{f,j}$ of node i was updated according to update rule (4.3) as soon as either $w_{i,j}$ or $(\min_{j_2 \in \mathcal{N}_j} D^j_{f,j_2})$ was changed and propagated by the neighbouring node $j \in \mathcal{N}_i$. Changes propagated further through a network could so possibly result in heavily oscillating cost measures in the routing tables $D^i_{f,j}$. As proposed by Littman and Boyan [75] a reinforcement learning approach can be

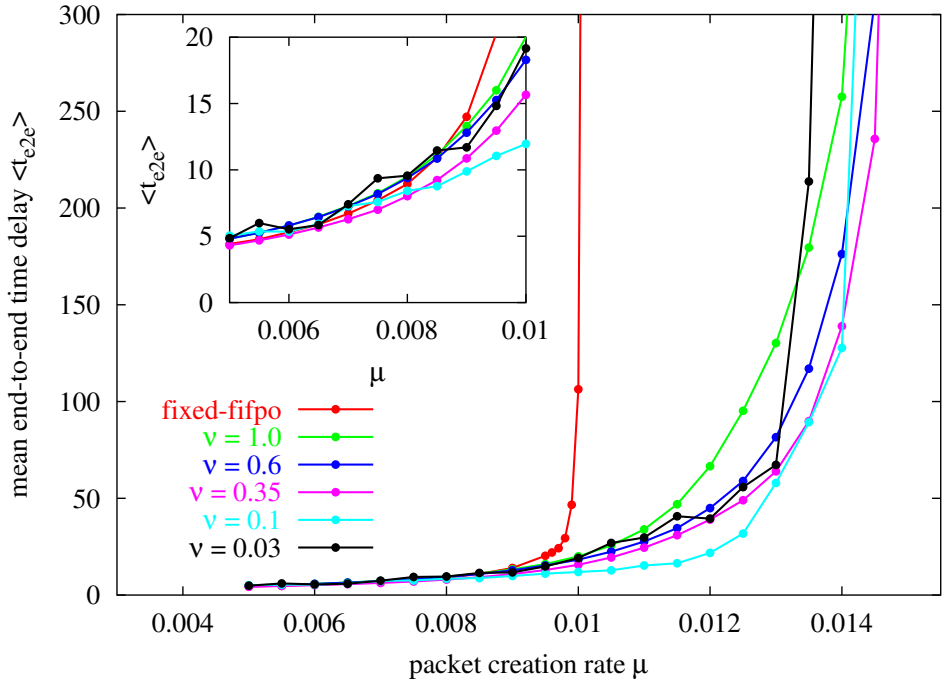

Figure 4.14: End-to-end time delay for different learning rates
The mean of the end-to-end time delay $\langle t_{e2e} \rangle$ over all packets of a $5 \cdot 10^5$ time steps simulation is illustrated for different learning rates ν depending on the packet creation rate μ. The constant-P reference network served as underlying network graph. The inset shows the enlarged part of the plot for smaller packet creation rates μ.

combined with the distance vector routing introduced so far. Instead of employing the update rule (4.3) an additional learning rate $\nu \in [0, 1]$ is incorporated. The modified update rule now reads [24]

$$D^i_{f,j} \leftarrow (1-\nu) D^i_{f,j} + \nu \left[w_{i,j} + \left(\min_{j_2 \in \mathcal{N}_j} D^j_{f,j_2} \right) \right] \tag{4.11}$$

For $\nu < 1$ a fraction of the old cost estimate is kept as a part of the new estimate. This modification includes a memory effect of formerly used routes that are only updated if significant changes in the network state have occurred. This approach is known as Q-routing [81].

The simple extension proved to cause a small but clearly measurable effect towards a further improvement in network performance. As indicated in Figure 4.14 the self-organised adaptive routing was performed for different learning rates ν employing the update rule (4.11) instead of (4.3). For values of ν decreasing from 1.0 to about $0.1 - 0.4$ the critical packet creation rate is pushed to slightly

larger values thus extending the subcritical regime. For $v < 0.1$ the effect reverses. It is intuitive that at a certain small learning rate the update rule collapses only keeping the once obtained values for the cost estimates in $D^i_{f,j}$. This behaviour can be verified observing the "update" measure introduced in Section 4.2. After an initial route discovery has been performed the system sticks to the once acquired cost estimates. Further changes in the network state do not result in updates of the routing tables $D^i_{f,j}$.

The implementation of the modified update rule (4.11) with the learning rate ν in the generic data traffic simulation is based on integer value representation in the routing tables $D^i_{f,j}$ which makes rounding operations in (4.11) necessary. The integer based implementation is retained in order to not slow down the already very time intensive simulations.

Further modifications of the update operations given by (4.7) and (4.11) to maintain the routing tables $D^i_{f,j}$ can be thought of. So far the incorporation of the buffer queue length n_j in (4.7) is only a crude measure of the time a data packet has to spend at a possible next hop node j. A multiplicative parameter depending on the average sending time $\langle t^{send}_j \rangle$ might prove useful to better evaluate the consequences of a buffer queue of length n_j. Also a reasonable treatment of packet generation and destruction could be included. So far no particular message is broadcasted if packets are generated or destroyed at a certain node accounting for the different influences of such events. By referring to the primary goal of this chapter – to develop a suitable routing algorithm that encompasses the features of self-organisation, local independence and exploitation of load dependent observables – a further, mostly numerical study of details will now be omitted.

4.5 Summary

As indicated in the very beginning of this chapter the major aim was not to provide a detailed routing protocol for communication in wireless ad hoc networks but to show the general ability of a self-organised routing algorithm to cope with the limitations of this particular network.

In a detailed study the special features of routing algorithms based on shortest paths have been reviewed. By extending the routing policy with a load dependent element in case of degenerate shortest paths a first promising hint for the predominance of adaptive routing algorithms could be identified.

By a sensible modification of an asynchronous distance vector routing first proposed by Littman and Boyen [75] a routing algorithm was constructed especially suited for routing in wireless ad hoc networks. This self-organised adaptive routing outperforms the previous routing schemes based on shortest paths.

It should be explicitly remarked that further detailed studies of the particular routing algorithm have to be carried out. The statistical properties of the routing

tables and the temporal and spatial distribution of the routing information in the network need a detailed review. As a typical measure relaxation times are often used in statistical physics. Additionally the performance of the routing algorithm for different spatial distribution of nodes, other power assignments, specific user behaviours and on/off switches of nodes are worth studying.

5 Conclusion and outlook

The importance of technological, biological and social networks has become more and more obvious in the last years [14, 15, 16]. These networks are the key elements of complex dynamical processes. Local interactions of the constituents contribute to the overall behaviour and often result in self-organising systems [7, 27, 54, 82]. Methods from modern statistical physics are widely applicable in the analysis of the underlying processes. Much of the growing insight in the self-organised control of complex networked systems is derived from the current research to understand genetic and metabolic networks as they occur in cells and higher organisms. But also within continuously growing technical systems similar questions arise, although the approaches might start from an opposite position. Whereas in most biological systems one tries to identify the interactions that lead to robustness and self-organisation, in a technological system one has to set up the rules in order to reach a desired state. Still the central problems remain the same: How can the local interactions between the networked constituents result in an overall optimised system?

Wireless multihop ad hoc networks represent an example of a complex technological system that calls for the integration of self-organising routines. High priority is given to the setup of distributed interactions rules that result in the emergence of efficient dynamical networked systems serving the needs of wireless communication. Especially for these particular communication networks different issues have to be addressed like connectivity, throughput, routing efficency, robustness, power consumption, quality of service, that, although viewed independently, make up the overall performance by their common interplay.

Connectivity is an essential prerequisite for a communication network. This issue was addressed in Chapter 2 where by introducing the minimum node degree rule a first step towards a distributed set of rules has been presented. The regularly used constant-P power assignment has serious disadvantages. Since this rule is by definition global it is not suited for a distributed approach. Furthermore it can not compensate local inhomogeneities in the density of nodes. In contrast the minimum node degree rule is solely based on local observables namely the number of direct communication neighbours of a node. By enforcing a minimum number of such neighbours each node itself adjusts its transmission power according to the local environment an thus counterbalances spatial inhomogeneities. The algorithm proved to provide efficiently connected networks with reduced total transmission power compared to the constant-P power assignment. In that respect the minimum node degree rule is an example for a distributed algorithm

that leads towards a self-organised connected network structure.

In a next step this underlying network structure is used for the transfer of data. These dynamical processes are closely connected to the particular network topology. By use of an appropriate generic data traffic model in Chapter 3 important characteristics of the dynamical processes have been studied. It could be shown that the throughput, which characterises the overall capacity of a communication network for data transfer, is limited by certain heavily used nodes. A detailed study of so called single-node observables proved highly valuable for the understanding of the influence of these nodes on the overall system dynamics. Sufficient analytical expressions could be derived for the description of the buffer queue content and the single-node temporal correlations. An estimate of the average sending time of a node allows evaluations about its criticality and thus about the overall limitations on the networked system.

The transfer of a data packet through a communication network on a specified path, called routing, is a highly complex problem. On the one hand one requires short transfer times, on the other hand one has to take care to keep the system in a non-congested state. On top of it local rules are required to manage all these tasks. Based on the idea of adaptive distance vector routing a distributed routing approach especially suited for ad hoc networks has been introduced in Chapter 4. Based on a local cost estimate each node does by itself decide where to forward a certain data packet. Since the estimates are especially good for the local surrounding the approach managed to distribute data traffic such that heavily used nodes are widely avoided. By employing this strategy the subcritical regime of data traffic could be remarkably extended. The particular set of rules independently discovers and maintains efficient routes through the network by employing load dependent observables. In that respect the self-organised adaptive routing is a first step towards a distributed routing scheme that serves the needs of wireless ad hoc communication.

Continuing the line of this work the next questions have to address the complex interplay between routing and topology. It seems certain that different network topologies require different routing strategies. At that point it is legitimate to ask what combination of both leads to the best overall performance. It is so far not clear to what extend a sophisticated routing algorithm can counterbalance a less than optimal topology.

Examples of highly complex, self-organised systems are manyfold in nature. Most of them are based on rather simple interactions rules. The investigation of these systems and the understanding of the complex processes that lead to self-organisation are a major challenge of current research. The profit is not only the academic gain of knowledge but also the possible adaption in new “man-made” systems of which wireless mobile ad hoc networks are just one example.

Acknowledgment

The presented thesis is based on the research I carried out at the Institute for Theoretical Physics of the Dresden University of Technology and at Siemens AG, Corporate Technology, Information & Communications in Munich.

I'm deeply thankful to Dr. Martin Greiner who was always a brilliant and persistant teatcher and a competent advisor. He guided me during the research for my thesis and taught me much about networks in particular and statistical physics in general. Since the time we spent at Duke University in Durham, NC, Martin encouraged my interest in theoretical physics and also opened my eyes for new and unconventional thoughts. I also want to express my thanks to the other members of the "ad hoc network group"at the Siemens AG in Munich, namely to Dr. Rudolf Sollacher and Wolfram Krause. I especially owe a lot to Wolfram Krause. He did not only introduce me to programming in OCAML but he was also a great help by answering all my questions about computers and what is around it. He always had a watchful eye on our common programming code and the finite running times are attributed to his commitment. Especially Figures 1.1, 3.1, 3.4, 3.15, 4.1 and 4.12 could be prepared thanks to his effort. I also gratefully acknowledge the advices in network routing from Uwe Röttgermann.

I want to thank Prof. Gerhard Soff for giving me the opportunity to write this thesis at his Chair at the Dresden University of Technology and for supporting my work. I enjoyed the friendly and pleasant atmosphere in his group. Especially the members of the Research Group in Theoretical Biophysics contributed much to it by the fruitful and diverse discussions we had.

One of the nicest aspects beyond my research work was the chance to share an office with my friend Michael Graupner. The warm atmosphere, the extensive discussions and the occasionally served black tea made it an unforgetable time with him.

I grateful acknowledge the support from Martin Wolf, Mathias Kuhnt and Lena Bachmann. Since our common time in North Carolina Martin Wolf was a great companion also beyond physics. Nevertheless I could profit a lot from his superior knowledge in mathematical physics. Mathias Kuhnt had an impartial view on parts of my thesis and I am thankful for his comments. I thank Lena Bachmann for her encouragement and all the good times we spent together.

I want to express my deepest gratitude to the help of my parents, Rosemarie and Rüdiger Glauche. Their constant encouragement were a great and valuable support during my whole studies. I also want to thank them for the permanent

financial support I could rely on.

I want to thank Claudia and Martin Greiner as well as their children Annkathrin, Josephine and Johanna for providing a pleasent and always welcome accomodation during my stays in Munich.

I thank Prof. Bernd Schürmann from Siemens AG Munich and Prof. Gerhard Soff for providing financial help for my research work.

Finally, I want to express my most intimate thanks to Veronika Börner. Her understanding and her constant support were always a safe backup during the whole time of my studies.

Bibliography

[1] J. Proakis, *Digital Communications* (McGraw-Hill, Singapore, 1995).

[2] R. Prasad, *Universal Wireless Personal Communications* (Artech House, Boston, 1998).

[3] T. Rappaport, *Wireless Communications - Principles & Practice* (Prentice Hall, Upper Saddle River, 1999).

[4] Z. J. Hass et al., editors, Special Issue on Wireless Ad Hoc Networks, IEEE J-SAC **17** (1999).

[5] J. P. Hubaux, T. Gross, J. Y. L. Boudec, and M. Vetterli, *Towards self-organized mobile ad hoc networks: the Terminodes project*, IEEE Comm. Mag. **31**, 118 (2001).

[6] Mobile Ad Hoc Networks (manet) Working Group (2003), URL `http://www.ietf.org/html.charters/manet-charter.html`.

[7] S. Valverde and R. Sole, *Self-organized Critical Traffic in Parallel Computer Networks*, Physica A **312**, 636 (2002).

[8] H. Fuks and A. Lawniczak, *Performance of data networks with random links*, Math. Comput. Simulation **51**, 103 (1999).

[9] R. Guimera, A. Arenas, A. Diaz-Guilera, and F. Giralt, *Dynamical properties of model communication networks*, Phys. Rev. E **66**, 026704 (2002).

[10] Y. Moreno, R. Pastor-Satorras, A. Vázquez, and A. Vespignani, *Critical load and congestion instabilities in scale-free networks*, Europhys. Lett. **62**, 292 (2003).

[11] K. Park and W. Willinger, *Self-Similar Network Traffic and Performance Evaluation* (John Wiley, Chichester, 2000).

[12] P. Abry, R. Baraniuk, P. Flandrin, R. Riedi, and D. Veitch, *The Multiscale Nature of Network Traffic: Discovery, Analysis, and Modelling*, IEEE SPM **19**, 28 (2002).

[13] W. Leland, M. Taqqu, W. Willinger, and D. Wilson, *On the Self-Similar Nature of Ethernet Traffic*, IEEE/ACM TON **2**, 1 (1994).

[14] R. Albert and A.-L. Barabási, *Statistical mechanics of complex networks*, Rev. Mod. Phys. **74**, 47 (2002).

[15] M. Newman, *The structure and function of complex networks*, SIAM Rev. **45**, 167 (2003).

[16] S. Dorogovtsev and J. Mendes, *Evolution of networks*, Adv. Phys. **51**, 1079 (2002).

[17] S. Bornholdt and H. Ebel, *World wide web scaling exponent from Simon's 1955 model*, Phys. Rev. E **64**, 035104 (2001).

[18] G. Parisi, *Complex Systems: a Physicist's Viewpoint*, Physica A **263**, 557 (1999).

[19] D. Stauffer and A. Aharony, *Introduction to Percolation Theory* (Tayler & Francis Ltd, London, 1994), 2nd ed.

[20] R. Meester and R. Roy, *Continuum Percolation* (Cambridge University Press, 1996).

[21] I. Glauche, W. Krause, R. Sollacher, and M. Greiner, *Continuum percolation of wireless ad hoc communication networks*, Physica A **325**, 577 (2003).

[22] E. Gelenbe and G. Pujolle, *Introduction to Queueing Networks* (John Wiley, Chichester, 1998), 2nd ed.

[23] D. Helbing, *Modelling supply networks and business cycles as unstable transport phenomena*, New J. Phys. **5**, 90.1 (2003).

[24] U. Roettgermann, *Decentralized Throughput Optimization in Industrial Networks*, Ph.D. thesis, Technische Universität München (2003).

[25] W. Krause, I. Glauche, R. Sollacher, and M. Greiner, *Selforganization in Communication Networks and beyond*, in Proc. of the Idea-Finding Symposium for the Frankfurt Institute of Advanced Studies, Frankfurt (2003).

[26] J. Davidsen, H. Ebel, and S. Bornholdt, *Emergence of a small world from local interaction: Modeling acquaintance networks*, Phys. Rev. Lett. **88**, 128701 (2002).

[27] S. Kauffman, *The Origins of Order: Self-Organization and Selection in Evolution* (Oxford University Press, New York, 1993).

[28] R. V. Sole, R. Pastor-Satorras, E. D. Smith, and T. Kepler, *A model of large-scale proteome evolution*, Santa Fe Institut Working Paper 01-08-041 (2001).

[29] H. Jeong, B. Tombor, R. Albert, Z. Oltvai, and A.-L. Barabasi, *The large-scale organization of metabolic networks*, Nature **407**, 651 (2000).

[30] S. Jain and S. Krishna, *Emergence and growth of complex networks in adaptive systems*, Computer Phys. Comm. **121**, 116 (1999).

[31] A.-L. Barabási, R. Albert, and H. Jeong, *Scale-free characteristics of random networks: the topology of the world-wide web*, Physica A **281**, 2115 (2000).

[32] R. Ferrer i Cancho and R. Sole´, *Optimization in complex networks*, arXiv:cond-mat/0111222 (2001).

[33] P. Holme, B. Kim, C. Yoon, and S. Han, *Attack vulnerability of complex networks*, Phys. Rev. E **65**, 056109 (2002).

[34] J. Dall and M. Christensen, *Random Geometric Graphs*, arXiv:cond-mat/0203026 (2002).

[35] W. Krause, I. Glauche, R. Sollacher, and M. Greiner, *Impact of network structure on the performance of wireless multihop ad hoc communication*, submitted (2003).

[36] D. Helbing, *Traffic and related self-driven many-particle systems*, Rev. Mod. Phys. **73**, 1067 (2001).

[37] J. Feder, *Fractals* (Plenum Press, New York, 1988).

[38] U. Frisch, *Turbulence* (Cambridge University Press, 1995).

[39] M. Greiner, H. Eggers, and P. Lipa, *Analytic multivariate generating function for random multiplicative cascade processes*, Phys. Rev. Lett. **80**, 5333 (1998).

[40] T. Lux, *Turbulence in Financial Markets: The Surprising Explanatory Power of Simple Cascade Models*, Quant. Finance **1**, 632 (2001).

[41] J. Muzy, J. Delour, and E. Bacry, *Modeling fluctuations of financial time series: from cascade process to stochastic volatility model*, Euro. Phys. J. B **17**, 537 (2000).

[42] R. Riedi, M. Crouse, V. Ribeiro, and R. Baraniuk, *A Multifractal Wavelet Model with Application to Network Traffic*, IEEE Special Issue on Information Theory **45**, 992 (1999).

[43] E. De Wolf, I. Dremin, and W. Kittel, *Scaling laws for density correlations and fluctuations in multiparticle dynamic*, Phys. Rep. **270**, 1 (1996).

[44] T. Halsey, M. Jensen, L. Kadanoff, I. Procaccia, and B. Shraiman, *Fractal measures and their singularities: The characterization of strange sets*, Phys. Rev. A **33**, 1141 (1986).

[45] M. Isichenko, *Percolation, statistical topography, and transport in random media*, Rev. Mod. Phys. **64**, 961 (1992).

[46] P. Gupta and P. Kumar, *The Capacity of Wireless Networks*, IEEE Trans. Inf. Th. **46**, 388 (2000).

[47] M. Marina and S. Das, *Routing Performance in the Presence of Unidirectional Links in Multihop Wireless Networks*, in Proc. 3rd ACM Int. Symposium on Mobile Ad Hoc Networking and Computing (MOBIHOC), Lausanne, Switzerland, pp. 12 –13 (2002).

[48] D. Watts and S. Strogatz, *Collective dynamics of small-world networks*, Nature **363**, 202 (1998).

[49] C. Bettstetter, *On the minimum node degree and connectivity of a wireless multihop network*, in Proc. of the 3rd ACM Int. Symposium on Mobile Ad Hoc Networking and Computing (MOBIHOC), Lausanne, Switzerland, pp. 80 – 91 (2002).

[50] F. Xue and P. Kumar, *The number of neighbors needed for connectivity of wireless networks*, to appear in Wireless Networks (2003).

[51] O. Dousse, P. Thiran, and M. Hasler, *Connectivity in ad-hoc and hybrid networks*, in Proc. INFOCOM, New York, pp. 1079–1088 (2002).

[52] M. Penrose, *On k-connectivity for a geometric random graph*, Rand. Struct. Alg. **5**, 145 (1999).

[53] D. Baker, G. Paul, S. Sreenivasan, and H. Stanley, *The Critical Percolation Threshold for Randomly Oriented Interpenetrating Squares and Cubes*, Phys. Rev. E **66**, 46136 (2002).

[54] B. Tadic, *Modeling Traffic of Information Packets on Graphs with Complex Topology*, Lecture notes in computer science **2657**, 136 (2003).

[55] A. Wuensche, *Basins of Attraction in Network Dynamics: A Conceptual Framework for Biomolecular Networks*, Santa Fe Institute Working Paper 02-02-004 (2002).

[56] R. Sole and S. Valverde, *Information Transfer and Phase Transition in a Model of Internet Traffic*, Physica A **289**, 595 (2001).

[57] E. Dijkstra, *A Note on Two Problems in Connexion with Graphs*, Numer. Math. **1**, 269 (1959).

[58] R. Sollacher, M. Greiner, and I. Glauche, *On the Dynamics of Wireless Mobile Ad-Hoc Networks*, submitted (2002).

[59] M. Newman, *A measure of betweenness centrality based on random walks*, arXiv:cond-mat/0309045 (2003).

[60] M. Newman, *Scientific collaboration networks. II. Shortest paths, weighted networks, and centrality*, Phys. Rev. E **64**, 016132 (2001).

[61] N. Hock, *Queueing Modelling Fundamentals* (John Wiley, Chichester, 1997).

[62] B. Huberman and R. Lukose, *Social dilemmas and internet congestion*, Science **277**, 535 (1997).

[63] D. Helbing, B. Huberman, and S. Maurer, *Traffic and Granular Flow '99: Social, Traffic, and Granular Dynamics* (Springer, Berlin, 2000), chap. Optimizing Traffic in Virtual and Real Space, pp. 193–204.

[64] R. Guimera, A. Arenas, A. Diaz-Guilera, F. Vega-Redondo, and A. Cabrales, *Optimal network topologies for local search with congestion*, Phys. Rev. Lett. **89**, 248701 (2002).

[65] C. Gardiner, *Handbook of Stochastic Methods for Physics, Chemistry and the Natural Sciences* (Springer, Berlin, 1983).

[66] M. Abramowitz, *Handbook of mathematical functions* (Dover Publications, New York, 1972).

[67] I. Csabai, *1/f Noise in Computer Network Traffic*, J. Phys.: Math. Gen. **27**, L417 (1994).

[68] M. Crovella and A. Bestavros, *Self-Similarity in World Wide Web Traffic: Evidence and Possible Causes*, IEEE/ACM TON **5**, 835 (97).

[69] L. Adamic, R. Lukose, and B. Huberman, *Handbook of Graphs and Networks - From the Genome to the Internet* (Wiley-VCH, Weinheim, 2003), chap. Local Search in Unstructured Networks.

[70] P. Holme, *Congestion and centrality in traffic flow on complex networks*, Advances in Complex Systems **6**, 163 (2003).

[71] R. Cohen, K. Erez, D. ben Avraham, and S. Havlin, *Resilience of the Internet to Random Breakdown*, Phys. Rev. Lett. **85**, 4626 (2000).

[72] A. Broder, R. Kumar, F. Maghoul, P. Raghavan, S. Rajagopalan, R. Stata, A. Tomkins, and J. Wiener, *Graph structure in the web*, Computer Networks **33**, 309 (2000).

[73] T. Huisinga, R. Barlovic, W. Knospe, A. Schadschneider, and M. Schreckenberg, *A microscopic model for packet transport in the Internet*, Physica A **294**, 249 (2001).

[74] R. Albert and H. Othmer, *The topology of the regulatory interactions predicts the expression pattern of the Drosophila segment polarity genes*, J. Theo. Bio. **223**, 1 (2003).

[75] M. Littman and J. Boyan, *A Distributed Reinforcement Learning Scheme for Network Routing*, in Proc. of the International Workshop on Applications of Neural Networks to Telecommunications, Princeton, pp. 45–51 (1993).

[76] M. Heusse, S. Guérin, D. Snyers, and P. Kuntz, *Adaptive Agent-driven Routing and Load Balancing in Communication Networks*, Adv. Compl. Sys. **1**, 237 (1998).

[77] R. Perlman, *Interconnections: Bridges, Routers, Switches, and Internetworking Protocols* (Addison-Wesley, Reading, MA, 2000), 2nd ed.

[78] A. Tanenbaum, *Computer Networks* (Prentice-Hall, Upper Saddle River, NJ, 1996).

[79] R. Bellman, *On a routing problem*, Quart. Appl. Math. **16**, 87 (1958).

[80] L. Kaelbling, M. Littman, and A. Moore, *Reinforcement learning: A survey*, JAIR **4**, 237 (1996).

[81] S. Choi and D.-Y. Yeung, *Predictive Q-routing: A memory-based reinforcement learning approach to adaptive traffic control*, Adv. Neur. Inf. Proc. Sys. **8**, 945 (1996).

[82] S. Bornholdt and T. Rohlf, *Topological Evolution of Dynamical Networks: Global Criticality from Local Dynamics*, Phys. Rev. Lett. **84**, 6114 (2000).

Declaration

Hereby I declare that this diploma thesis is the result of my own work except where explicit reference is made to the work of others. This work has not been submitted for another qualification to this or any other examination board.

Ingmar Glauche Dresden, 23nd October 2003

Erklärung

Hiermit bestätige ich, dass ich diese Diplomarbeit ohne unzulässige Hilfe Dritter angefertigt und alle Quellen als solche gekennzeichnet habe. Die Arbeit wurde bisher keiner anderen Prüfungsbehörde in dieser oder ähnlicher Form vorgelegt.

Ingmar Glauche Dresden, 23. Oktober 2003