Longitudinal hopping in intervehicle communication: Theory and simulations on modeled and empirical trajectory data

Christian Thiemann,[1,2,*] Martin Treiber,[1,†] and Arne Kesting[1]
[1]*Technische Universität Dresden, Andreas-Schubert-Strasse 23, 01062 Dresden, Germany*
[2]*Max Planck Institute for Dynamics and Self-Organization, Bunsenstrasse 10, 37073 Göttingen, Germany*
(Received 30 April 2008; published 3 September 2008)

Intervehicle communication enables vehicles to exchange messages within a limited broadcast range and thus self-organize into dynamical and geographically embedded wireless *ad hoc* networks. We study the longitudinal hopping mode in which messages are transported using equipped vehicles driving in the same direction as a relay. Given a finite communication range, we investigate the conditions where messages can percolate through the network, i.e., a linked chain of relay vehicles exists between the sender and receiver. We simulate message propagation in different traffic scenarios and for different fractions of equipped vehicles. Simulations are done with both, modeled and empirical traffic data. These results are used to test the limits of applicability of an analytical model assuming a Poissonian distance distribution between the relays. We found a good agreement for homogeneous traffic scenarios and sufficiently low percentages of equipped vehicles. For higher percentages, the observed connectivity was higher than that of the model while in stop-and-go traffic situations it was lower. We explain these results in terms of correlations of the distances between the relay vehicles. Finally, we introduce variable transmission ranges and found that this additional stochastic component generally increased connectivity compared to a deterministic transmission with the same mean.

DOI: 10.1103/PhysRevE.78.036102 PACS number(s): 89.40.−a, 05.65.+b

I. INTRODUCTION

Complex networks have drawn the attention of physicists for a couple of years and many real-world systems have been studied under this paradigm [1–5]. In particular, percolation theory has been utilized [6–9] to determine the impact of failing nodes on the overall network—either to prevent failures such as power outages or Internet breakdowns [10,11], or to deliberately induce failures, e.g., stopping a spreading disease by vaccination [12,13].

Another type of network relying on short-range percolation is an "*ad hoc*" vehicular network, where nodes represent vehicles equipped with wireless communication devices, and edges connect vehicles that can communicate with each other within a limited communication range. Due to the vehicle movements, links in the communication network will be broken and others generated, resulting in a dynamical network topology. Looking at a snapshot of the network at a given time, the geographic locations of the nodes are stochastic, and, by virtue of the limited communication range, this stochasticity carries over to the existence of links between the nodes. A second layer of stochasticity can be introduced by allowing for variable communication ranges, i.e., the maximum allowed physical distance for a link between a pair of nodes is a stochastic variable itself.

In addition to the academic allure of such a network, the relevance of the introduced system consists in its application to the emerging technology of *ad hoc* wireless networks in individual transport, known as intervehicle communication (IVC). Intervehicle communication has important and promising applications in traffic safety, information, and control. On a small scale, a vehicle involved in an accident or suffering from a breakdown might constantly send out messages stating that there is an emergency situation, thus warning upcoming vehicles in advance. Furthermore, with the emergence of adaptive cruise-control systems (ACC) and associated sensors for the distance to the preceding vehicle, equipped vehicles can inform their environment about the local traffic state, and past actions (e.g., a hard braking maneuver). The next ACC and IVC equipped vehicles in the upstream direction can process this information and automatically adapt their driving style to the situation or at least display suggestions and warnings to the driver, which may serve as a mean to homogenize traffic and thus increasing the overall traffic throughput in a self-organized way [14]. On a larger scale, the information passed along the IVC network can be aggregated to produce live floating car data. In addition to the conventional cross-sectional data from loop detectors, this information can be used by traffic-state recognition and traffic forecast models thereby increasing their performance and reliability.

In most applications, it is necessary to carry messages over distances that are significantly longer than the device's broadcast range. This can be achieved by two ways: First, a message can be passed to the following vehicle which then passes it to its following vehicle, and so on. We call this longitudinal hopping since the message always travels parallel to the desired travel direction. Or, a vehicle may transfer a message to a vehicle driving in the opposite direction. This vehicle can store the message and continuously broadcast it for a certain period of time while mechanically transporting the message upstream. Although this message is of no use for this relay vehicle, it might eventually jump back to the original driving direction by a second transversal hop. We therefore call this method transversal hopping or "store and forward."

*mail@christian-thiemann.de
†martin@mtreiber.de. URL: http://www.traffic-simulation.de

The longitudinal hopping mode has the advantage of virtually instantaneous message transmission, so the transmitted information is always up-to-date. However, the reliability may be low since a fully connected chain of IVC equipped vehicles is needed. This restriction is overcome by the transversal hopping mode where the successful transmission is only a matter of time, but the information may be obsolete when it finally arrives. There is, in fact, a third possibility of intervehicle communication involving roadside units which may act as a repeater between two vehicles which are too far apart for direct communication. However, this concept requires an extensive amount of stationary hardware.

Studies of intervehicle communication often focus on technical details of the IVC devices and the communication protocol resulting in complex message transmission models. In contrast, a simplistic traffic model, if any at all, is then used to simulate IVC systems. To our knowledge, only a few studies exist that use more sophisticated and empirically tested traffic models [15–18]. For a physicist, however, the influence of the traffic dynamics (dynamics of nodes) on the connectivity of the network is more interesting than the communication details. Moreover, to our knowledge, investigations of IVC that are based on real vehicle trajectory data do not exist.

In this paper, we will study the longitudinal hopping of intervehicle messages on a freeway with the focus on how far upstream a message is expected to travel in different traffic situations and for different percentages of equipped vehicles. To this purpose, we compare vehicle networks based on real and simulated vehicle trajectory data with analytical results assuming uncorrelated node positions [19]. We found that long-ranged correlations of the distances arising from traffic instabilities generally lead to a decrease of the connectivity while short-ranged correlations resulting from the need for a minimum gap lead to an increase. Remarkably, stochasticity in the communication ranges increases the connectivity as well.

We start by introducing the analytical model for randomly distributed nodes and deterministic link transmission, which we then adapt to our network. In Sec. III, we will compare the model predictions with a network simulation based on simulated (Sec. III A) and empirical trajectory data (Sec. III B). In Sec. III C, we introduce stochastic transmissions and investigate their properties, both analytically, and based on the simulated trajectories. In Sec. IV, we discuss the implications of our findings.

II. ANALYTICAL MODEL

In this section, we will adapt an analytical model for the availability of messages to the IVC network. The model assumptions are (i) longitudinal hopping along a linear chain, (ii) a deterministic and instantaneous transmission mechanism in which a message is available for receiving within a certain radius r from the sender with certainty, but unavailable further away, (iii) randomly distributed nodes (IVC vehicles) with a linear node density λ, i.e., the distances Y between two nodes are i.i.d. exponentially distributed stochastic variables whose density is given by

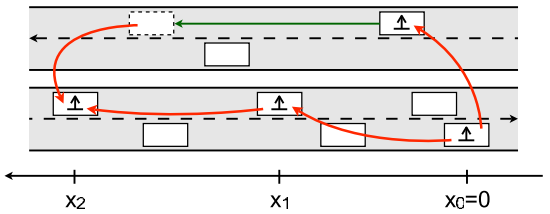

FIG. 1. (Color online) The figure shows a four-lane highway with 50% IVC equipped vehicles. The longitudinal hopping mode uses vehicles driving in the same direction to instantaneously transmit a message, while with transversal hopping a relay vehicle driving into the opposite direction is used and the message is delivered with time delay. In the analytical model for the longitudinal hopping mode, since we are interested in vehicles upstream of the sender only, the vehicle positions are measured antiparallel to the driving direction such that the stochastic variables X_i represent the distance to the sender vehicle.

$$f_Y(y) = \lambda e^{-\lambda y}. \tag{1}$$

Given traffic on m lanes with an average vehicle density ρ per lane and a percentage of equipped vehicles α (which we will refer to as penetration level), the node density is equal to the longitudinal partial density of IVC-equipped vehicles,

$$\lambda = m\alpha\rho. \tag{2}$$

Now, we consider an IVC vehicle at position $x_0=0$, and a chain of n further IVC vehicles in the upstream direction, whose positions are given by the random variables X_i, $i=1,\ldots,n$ (increasing in the upstream direction) (Fig. 1). A vehicle beyond the end of the chain—with distance $x>X_n$ from the sender—that is not within the communication range of the sender will be able to receive the message if it is within the communication range r of the last vehicle at X_n, and if this last vehicle has received the message (event A_n), i.e., the probability of availability is given by

$$P_c(x) = P[(x - X_n < r) \text{ and } A_n]. \tag{3}$$

We can rewrite the first condition in terms of the distance to the last vehicle of the chain $Y=x-X_n$. For any realization y of Y with $0 \leqslant y < r$, the first condition is fulfilled and the second is fulfilled with probability $P_c(x-y)$. Therefore, $P_c(x)$ transforms to an integral over all possible distances, where a communication is not *a priori* excluded,

$$P_c(x) = \int_0^r \lambda e^{-\lambda y} P_c(x-y)dy. \tag{4}$$

From here it is straightforward to construct a delay-differential equation for $P_c(x)$. Taking the derivative

$$\frac{dP_c}{dx} = \int_0^r \lambda e^{-\lambda y} \frac{dP_c(x-y)}{dx} dy,$$

and simplifying this expression using partial integration and Eq. (4), one obtains

$$\frac{dP_c}{dx} = -\lambda e^{-\lambda r} P_c(x-r). \tag{5}$$

In order to solve this delay-differential equation, a complete function $[0,r] \to \mathbb{R}$ must be provided as initial condition. According to model assumption (ii), the message is always available for $x<r$. At the distance $x=r$, however, the message is available only if there is at least one additional relay vehicle between the sender and the destination point x, which is the case with probability $\int_0^r f_Y(y)dy$, resulting in

$$P_c(x) = \begin{cases} 1, & x < r, \\ 1 - e^{-\lambda r}, & x = r. \end{cases} \tag{6}$$

In order to obtain an analytical result, we insert the initial condition into Eq. (5) and obtain

$$\frac{dP_c}{dx} = -\lambda e^{-\lambda r}, \quad r \leqslant x < 2r. \tag{7}$$

With the initial condition $P_c(r)=1-e^{-\lambda r}$ we obtain

$$P_c(x)|_{r \leqslant x < 2r} = 1 - e^{-\lambda r} - \lambda e^{-\lambda r}(x-r). \tag{8}$$

Similarly, for $2r \leqslant x < 3r$,

$$\frac{dP_c}{dx} = -\lambda e^{-\lambda r}\{1 - e^{-\lambda r}[1 + \lambda(x-r)]\}$$

with $P_c(2r)=1-e^{-\lambda r}(1+\lambda r)$ from (8) yields

$$P_c(x)|_{2r \leqslant x < 3r} = 1 - e^{-\lambda r}[1 + \lambda(x-r)] - e^{-2\lambda r}\left[[-\lambda(x-2r)]\left(1 + \frac{\lambda(x-2r)}{2}\right)\right].$$

Generalizing from the previous steps to arbitrary x we obtain

$$P_c(x) = 1 - \sum_{k=1}^{m} e^{-k\lambda r}\left[\frac{[-\lambda(x-kr)]^{k-1}}{(k-1)!}\left(1 + \frac{\lambda(x-kr)}{k}\right)\right]. \tag{9}$$

Here, $m=\lfloor x/r \rfloor$ denotes the integer part of x/r. When scaling distances to the communication range, $x=r\tilde{x}$, Eq. (9) depends, as a function of $\tilde{x}$, only on the scaled partial density $\tilde{\lambda}=\lambda r$. For reference, expression (9) can be transformed into the equation given by Dousse *et al.* [19],

$$P_c(x) = \sum_{k=0}^{m} \frac{[-\lambda e^{-\lambda r}(x-kr)]^k}{k!} - e^{-\lambda r}\sum_{k=0}^{m-1} \frac{\{-\lambda e^{-\lambda r}[x-(k+1)r]\}^k}{k!}. \tag{10}$$

We emphasize that Eqs. (9) and (10) express the probability of a message being available at a certain distance x from the sender. For the transfer of information within the IVC network, however, the relevant quantity is the distribution of the physical length X of the unbroken chain of equipped vehicles connected to a given IVC vehicle in the downstream direction. This can be seen by considering a local information that can be captured by the IVC network only when an equipped vehicle passes the location of the information source (the information, "rear end of a traffic jam," would be an example). For the driver of an IVC vehicle at an upstream distance x from the source, this information is available if the vehicle is part of a chain of length $X>x$.

In our model, the availability of the message always extends beyond the last connected vehicle by the communication range r. Thus, the probability $P_{\text{chain}}(x)$ that the chain length X is larger than x is given by

$$P_{\text{chain}}(x) = P(X > x) = P_c(x+r). \tag{11}$$

The assumed exponential distribution of node distances, $f_Y(y)$, may be wrong under certain conditions, e.g., a single-lane road where vehicles keep a minimum distance to avoid accidents. This can be overcome by postulating a compressed node distance distribution

$$f'_Y(y) = \Theta(y-s)\lambda' e^{-\lambda' y}. \tag{12}$$

Here, $\Theta(\cdots)$ denotes the Heaviside function, and the "hard-core radius" s is an additional model parameter stating the minimum distance between two nodes. To reflect the same average node density as with the exponential distribution $f_Y(y)$, the parameter λ' is chosen to be $(\lambda^{-1}-s)^{-1}>\lambda$, resulting in lower probabilities for very large distances (thus the term "compressed"). From this distribution we can derive a delay-differential equation in the same manner as above,

$$\frac{dP'_c}{dx} = -\lambda' e^{-\lambda'(r-s)} P'_c(x-r) - \lambda'[P'_c(x) - P'_c(x-s)]. \tag{13}$$

Since this equation is more complex than Eq. (5) and contains an additional model parameter, we will focus on the assumption of an uncompressed exponential distance distribution in the rest of this paper.

In Fig. 2 we discuss some general characteristics of the model. The probability of availability $P_c(x)$ is equal to 1 for $x<r$ which corresponds to the initial condition (6). In the interval $r<x<2r$, the probability decays linearly with the distance, as stated in (8). As expected, the probability decays slower when increasing the range r or the partial density λ. In fact, the decay (as a function of x/r) is controlled by the product λr. Figure 2(c) illustrates Eq. (11) expressing the difference between the probability of a message being available at a certain distance and the probability that the last vehicle receiving the message is at a certain distance from the sender. Furthermore, Fig. 2(c) shows the increased probability of availability $P'_c(x)$ when using the compressed node distance distribution (12). Finally, Fig. 2(d) gives the expectation value $\langle X \rangle$ of the physical length of the communication chain for different values of r and λ.

III. COMPARISON TO SIMULATION RESULTS

In order to test the assumptions of the analytical model we have performed several simulations of a longitudinal message hopping with the following two-stage approach. First, vehicle trajectories are obtained, either from the empirical data provided by the next generation simulation community (NGSIM) or by simulating the vehicle dynamics on a road

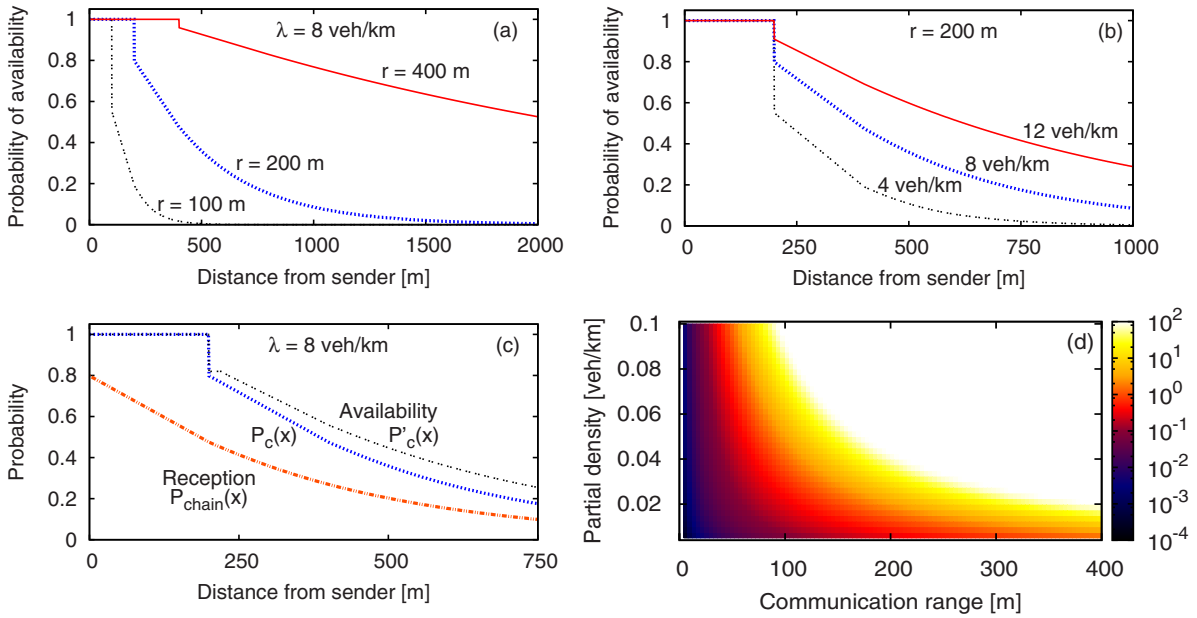

FIG. 2. (Color online) Characteristics of the analytical model. (a) Availability of an intervehicle message for a fixed density of equipped vehicles but different communication ranges r. (b) The same for fixed communication range but different equipped vehicle densities. (c) Difference between the probability of availability, $P_c(x)$, and the probability $P_{\text{chain}}(x)=P_c(x+r)$ that the last vehicle receiving the message is, at least, at a distance x from the sender. Notice that $P_{\text{chain}}(0)$ denotes the probability for any IVC connection at all. Also, the probability of availability $P_c'(x)$ with the compressed node distance distribution $f_Y'(y)$ ($s=20$ m) is compared against $P_c(x)$. (d) Expectation value $\langle X\rangle$ of the chain length for different values of the communication range and the density of equipped vehicles [color (gray)-value coded; in the white area the value exceeds 100 km which is not plotted].

section using a car-following model. Then, the communication is simulated by analyzing snapshots of the vehicle trajectories for several times t. From the NGSIM data (see Sec. III B for details), the complete time interval of ≈30 min has been sampled with a rate of 15 Hz, resulting in 25 500 snapshots. From the simulated data, snapshots have been taken every 10 s resulting in 3750 snapshots of homogeneously congested traffic, 4550 of stop-and-go traffic, and 5650 of free traffic.

For each snapshot, the following stochastic experiment has been repeated 100 times: First, the population of equipped vehicles is determined by independently assigning to each vehicle the property "is an IVC vehicle" with probability α (the penetration level). All other vehicles are discarded. Then, a vehicle issuing a message is randomly chosen from the last 10% of the remaining vehicles at the downstream end. Iterating upstream over all vehicles, the message hops to the next vehicle if and only if its distance to the previous vehicle is less than r. Otherwise, the message chain is broken and the chain length is output.

The final result of the simulation is the cumulative distribution function $F_{\text{cl}}(x)$ of the chain length distribution, which states the probability of a message not being received beyond a given distance x from the sender vehicle. Thus, $P_{\text{chain}}(x)=1-F_{\text{cl}}(x)$ is the probability for a successful message propagation over a distance x. In order to compare the result to the theory, the mean vehicle density ρ is determined from the trajectory data, and the expectation value of the partial density of IVC vehicles [which is, besides r, the only input variable of Eq. (11)] is estimated by the relation $\lambda=\alpha\rho$.

A. IVC in simulated traffic data

We will now present the application of this method to generated trajectory data, obtained by simulating a single-lane road section of 52 km length with open boundary conditions using the Gipps model [20]. Similar spatiotemporal patterns have been obtained for other models [21,22]. At position $x_b=50$ km we added an on-ramp with a 100 m merging zone inducing a bottleneck. By varying the ramp flow and the inflow boundary conditions, three data sets with different traffic situations upstream of the bottleneck were obtained: Free traffic with an average density of 19.5 veh/km, congested traffic with density 49.5 veh/km and no internal structure, and oscillatory congested traffic of average density 33.7 veh/km exhibiting regular stop-and-go waves with a wavelength of approximately 2 km.

Figures 3(a) and 3(b) show the resulting simulated connectivities $P_{\text{chain}}(r)$ for free traffic assuming a range $r=200$ m and various penetration levels α. For $\alpha<40\%$, we found a good agreement with the analytical model (11), while, for larger penetration levels, the observed connectivity was higher than the analytical result. Notice that the apparent "quantization" of the connectivity for small densities leading to steps of width $1/\rho$ is a consequence of simulating identical vehicles with a deterministic acceleration rule. Since, for such densities, there is very little interaction, the initial distance between the vehicles remains essentially unchanged.

In congested traffic [Figs. 3(c) and 3(d)], the situation is similar with a good agreement for small α, while the theory underestimates the observed connectivity for larger penetration levels. In quantitative terms, the deviation reaches the

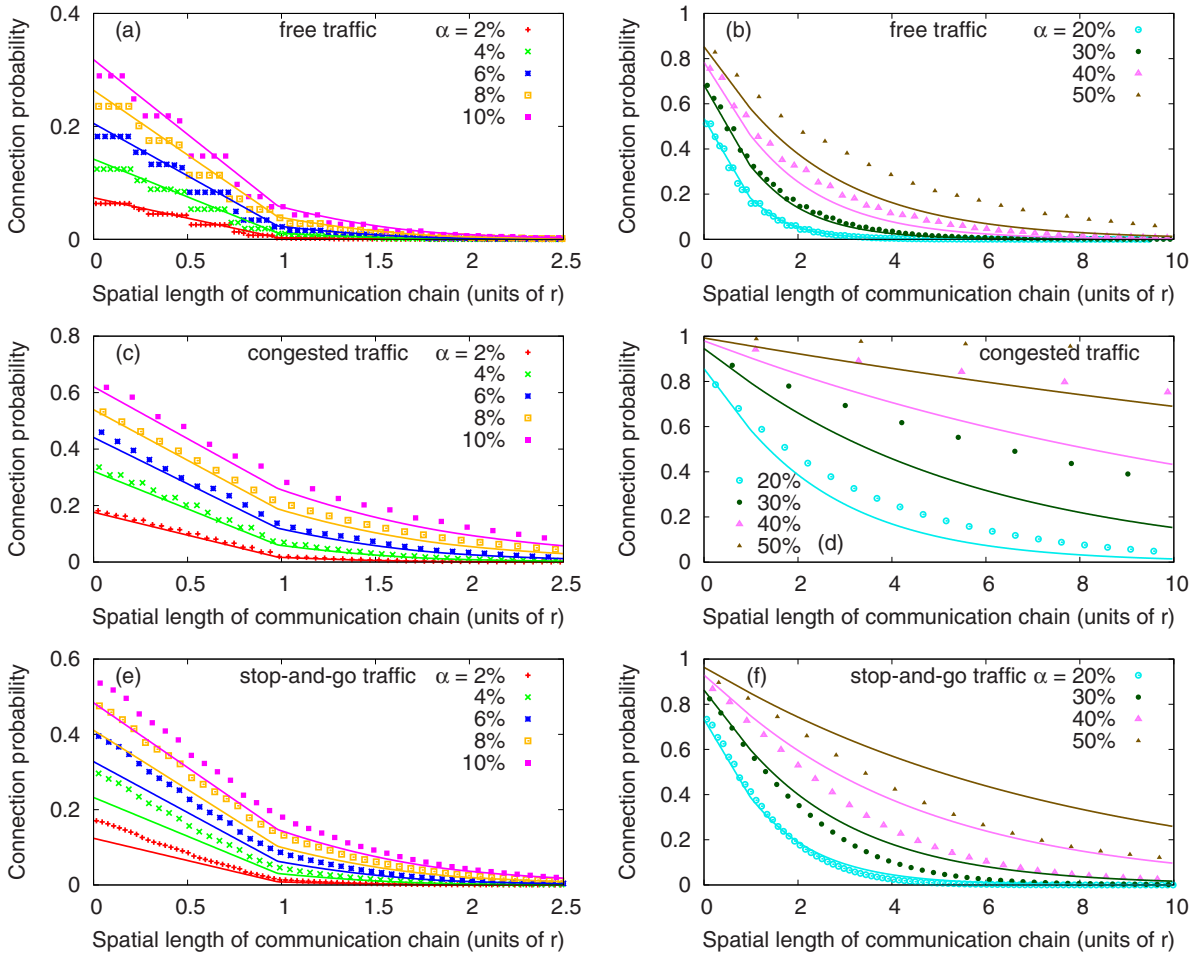

FIG. 3. (Color online) Probability of message reception at a certain distance from the initial sender in free traffic for $r=200$ m and several values of the penetration level α. Shown are the analytical values $P_{\text{chain}}(x)$ (solid lines), and simulation results (symbols) based on trajectories generated with the Gipps model for free traffic (upper row), homogeneous congested traffic (middle row), and stop-and-go traffic (bottom row).

order of 0.05 for $\alpha \approx 15\%$ while, in free traffic, such discrepancies are only observed for $\alpha \geqslant 40\%$. Remarkably, the situation is drastically different for stop-and-go traffic [cf. Figs. 3(e) and 3(f)] where, for sufficiently high penetration levels, we observed lower reception probabilities compared to theory, while, for penetration levels below 20%, they were higher.

The findings can be understood in terms of correlations that are observed in real (and simulated) traffic, but are absent in the analytical communication model. Obviously, one source of correlations are repulsive interactions that are necessary to keep a certain minimum "safe" distance to the leading vehicle. This leads to short-ranged correlations of the distances between the vehicles decaying rapidly over a few vehicle distances [23]. With respect to the theoretical exponential distribution (1), this effect compresses the actual distance distribution towards the mean value $1/\rho$. If the IVC penetration level α is sufficiently high such that the mean distance $1/\lambda = 1/(\alpha\rho)$ of the equipped vehicles does not significantly exceed the correlation length, the correlation carries over to that of the equipped vehicles. If, in addition, the mean distance satisfies $1/\lambda < r$, the probability of distances exceeding r thereby breaking the chain is reduced. Since the correlation length and the assumed communication range $r = 200$ m are of the same order, both conditions are satisfied simultaneously for sufficiently high penetration levels explaining the higher connectivities observed in this regime [cf. also Fig. 2(c)]. If, however, the average node distance $1/\lambda$ exceeds the correlation length, the positions of the equipped vehicles are nearly independent (even if that of all vehicles are not), and the analytical model is approximately valid. Since the repulsive force acts only between vehicles in the same lane, the effect described above is expected to be weaker for multilane traffic.

In order to understand the opposite effects of a lower observed connectivity in the case of collective traffic instabilities such as stop-and-go traffic, one needs to consider the long-range correlations associated with them. In the low-density areas, the local connectivity is lower than that at average density which is assumed in the theory, and the message chain breaks with a higher probability. Moreover, since the linking of the chain is a nonlinear and multiplicative process (the weakest link matters), the higher local connectivities in the high-density areas cannot compensate for the loss in the low-density areas. This mechanism is only effective if (i) the probability of a connected chain with several

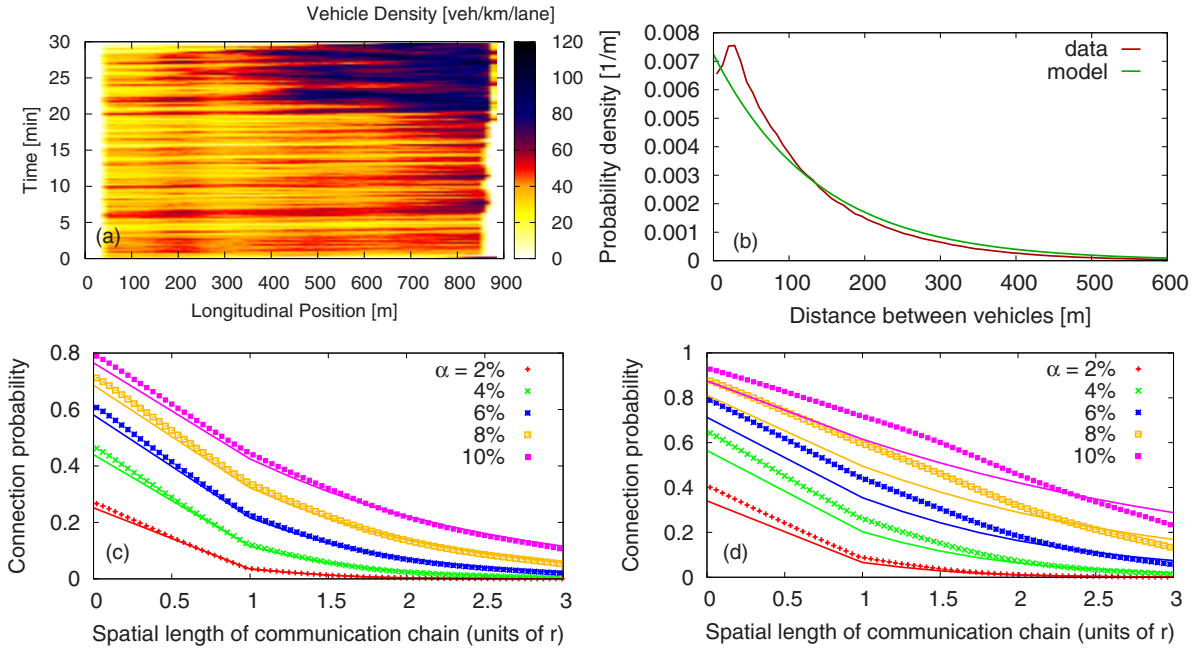

FIG. 4. (Color online) (a) Spatiotemporal vehicle density in the NGSIM data. (b) Distance distribution found in the first 20 minutes of the data for $\alpha=0.1$ and the corresponding exponential model. (c) Analytical and simulation results ($r=200$ m) for the first 20 minutes, and (d) for the last 10 minutes of the observation interval.

links is sufficiently high, and (ii) the average distance $1/\lambda$ between two connected vehicles is significantly smaller than the wavelength of the traffic waves (≈ 2 km). Both criteria are satisfied for higher penetration levels which is in agreement with Fig. 3(f).

For low penetration levels, however, the typical chain length is very short (or there is no valid connection at all), and the above "weakest-link effect" is dominated by a statistical "selection effect" acting in the opposite direction: When selecting the sending vehicle randomly (as in the simulation), the probability of this vehicle being in a high-density region is comparatively high (simply because there are more vehicles in high-density than in low-density regions). If, in addition, the typical chains are significantly shorter than the size of one density cluster (a few 100 m), there is a high probability that the whole chain is located inside the cluster, and the weakest-link effect is not applicable. The resulting net effect is a higher observed connectivity compared to the theory explaining the findings of Fig. 3(e). Notice that even the transition from a negative to a positive deviation of the theoretical connectivity with respect to the data, observed for $\alpha=0.2$, can be understood in terms of these two counteracting mechanisms.

B. IVC in empirical traffic data

In order to check the applicability of the model to real traffic situations, we simulated a dynamical network based on empirical trajectory data collected on the I-80 (Emeryville, California) in 2003 by the Berkeley Highway Laboratory [24]. The trajectory data was obtained from video footage recorded by six cameras mounted on a 97 m tall building. The cameras captured 30 minutes of traffic on an 800 m segment of a six-lane highway with an on-ramp at the upstream end and an off-ramp at the downstream end. We extracted the four left-most lanes ignoring on-or off-ramp traffic and used it as input for our IVC simulations. As in Sec. III A, we have carried out the simulations with $r=200$ m, and various values for α (cf. Fig. 4). The sender vehicle was randomly chosen from the most downstream region (width ≈ 100 m) ensuring that at least 700 meters of measurement area were available for the hopping chain.

Figure 4(a) gives an impression of the vehicle densities in the data and clearly identifies two different regimes: In the first 20 minutes we see fairly homogeneous (dense) traffic while, in the last 10 minutes, we observe a traffic jam covering the downstream one-half of the observation area. In order to sample over comparable traffic situations, we have split the data into these two regimes and carried out the simulations on both datasets, separately.

Figure 4(c) shows the simulated connectivities for the first 20 minutes. As in Fig. 3(a) or 3(c), the simulated connectivities are consistent with, or only slightly above, the analytical values (calculated for $\rho=72.4$ veh/km). For higher penetration levels (not shown), the difference becomes more significant, similar, though not as pronounced, as in Figs. 3(b) or 3(d).

Analysis of the distance distribution for the first 20 minutes allows for a direct check of the assumed Poisson distribution used in the derivation of the analytical model in Sec. II. For each equipped vehicle, we have calculated the longitudinal distance to the next equipped vehicle, regardless of the lane. Figure 4(b) shows the resulting probability density for $\alpha=0.1$. Although the correlations are expected to become weaker for increasing number of lanes, the compression effect discussed in the preceding section clearly can be seen. As a consequence, the probability of distances beyond

$r=200$ m is lower than in the exponential distribution, which is consistent with the increase of the simulated connectivity shown in Fig. 4(c). Nevertheless, the discrepancy is remarkably low, in agreement with [16].

The simulation for the last 10 minutes of the data [Fig. 4(d)] reports mostly larger reception probabilities than theoretically predicted. However, for chain lengths exceeding $\approx 2.5r$ (≈ 500 m), there is a crossover to lower values compared to theory. This is consistent with the network simulations based on artificial trajectories of stop-and-go traffic, Figs. 3(e) and 3(f). Moreover, it seems that the observed crossover for long chains and higher penetration levels can be understood in terms of the interplay between the weakest link and statistical selection effects already discussed in Sec. III A.

C. Stochastic message transmission model

A variety of causes can reduce or increase the actual transmission range. In the following, we therefore relax the assumption that the communication range r for each hop is exactly given by $r=200$ m. Instead, we now model the communication ranges as i.i.d. Gaussian distributed stochastic variables, $R \sim N(r,\sigma^2)$ with unchanged mean $r=200$ m and standard deviation $\sigma=75$ m.

Equations (4) and (6) can be generalized according to the following reasoning. Given a distance x and the range $R_0 \sim N(r,\sigma^2)$ of the sender vehicle, a direct communication is possible with probability $1-F(x)$ where $F(x)$ denotes the cumulative distribution function of the range. With the complementary probability $F(x)$, however, one or more relays are needed. The communication probability for this case is given by summing, over all possible intervehicle distances y, the probability $P_c(x-y)$ that the message is available at the location of the relay vehicle, multiplied with the conditional probability that the range R of the last relay is larger than y,

$$P_c(x|R_0 < x) = \int_0^x \lambda e^{-\lambda y} P_c(x-y) P(R > y|R_0 < x) dy. \tag{14}$$

Assuming independent ranges, $P(R>y|R_0<x)=P(R>y)=1-F(y)$, the unconditional probability can be written as

$$P_c(x) = 1 - F(x) + F(x) \int_0^x \lambda e^{-\lambda y} [1 - F(y)] P_c(x-y) dy. \tag{15}$$

For nontrivial distribution functions $F(y)$, there is generally no possibility to obtain an explicit nonintegral expression.

We point to the fact that (15) does not include all ways of transmission since non-Markovian effects generalizing the network topology (it is no longer a linear chain) are neglected. For example, even if the link between the nodes $i-1$ and i is broken, a message may nevertheless be transmitted from node $i-2$ to i. As a necessary condition, the range R_{i-2} of node $i-2$ must be larger than the range R_{i-1} of the neighboring node by at least the distance Y_{21} between the nodes $i-2$ and $i-1$. This "bridging effect" becomes significant if $\sigma\lambda$ is of the order of unity, or larger.

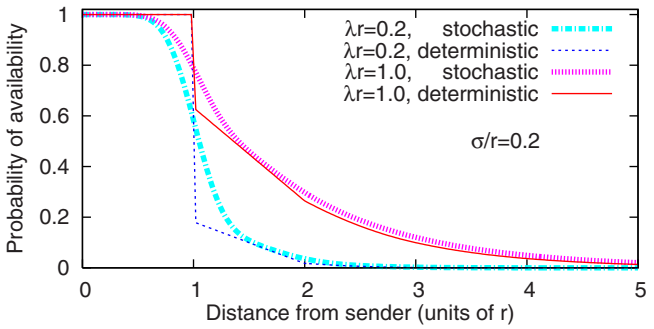

FIG. 5. (Color online) Availability of the intervehicle message in the analytic stochastic model assuming i.i.d. Gaussian distributed communication ranges with a coefficient of variability $\sigma/r=0.2$.

Figure 5 shows typical results for the stochastic model (including the non-Markovian effects), in comparison with the reference case of deterministic communication. The kinks and discontinuities of the deterministic case are smoothed out. Moreover, the availability in the stochastic case compared to the deterministic situation tends to be lower if there is a high probability for a direct communication ($x<r$), while it tends to be higher, otherwise.

Figure 6 shows the results for the three simulated trajectory datasets in comparison with the deterministic analytical results. In agreement with the above considerations, the kinks are smoothed out. Remarkably, the simulated connectivity generally turned out to be higher than in theory, with the exception of stop-and-go traffic at high penetration levels. For $\lambda r>1$, it is also significantly larger than the connectivity simulated for fixed ranges (cf. Fig. 7). Nevertheless, the correlations of the node positions typically influence the results more strongly than the stochasticity of the communication range, at least for the displayed parameter settings.

IV. DISCUSSION

Intervehicle communication (IVC) is an interesting concept, both theoretically and with respect to applications. It can be described as a "double-dynamic" network where both the nodes (the drivers or vehicles), and the links (messages hopping from one vehicle to the next) obey their own dynamics. Since the transmitted information can be used to change the driving behavior (thereby potentially increasing the efficiency of traffic flow) [14,22], a feedback between the two dynamic processes is possible.

Since the IVC technology is not yet implemented and no large-scale empirical data can be produced, an understanding of this type of network is also relevant for assessing the performance of IVC designs under various conditions. Like all technologies relying on local communication, IVC is only effective if the node density exceeds a certain minimum. It is, therefore, particularly relevant to investigate the order of magnitude and the influencing factors of the corresponding "penetration threshold."

In this contribution, we have investigated the connectivity of the longitudinal hopping mode where the network is formed by vehicles driving in only one driving direction, and the messages hop from vehicle to vehicle against the traffic

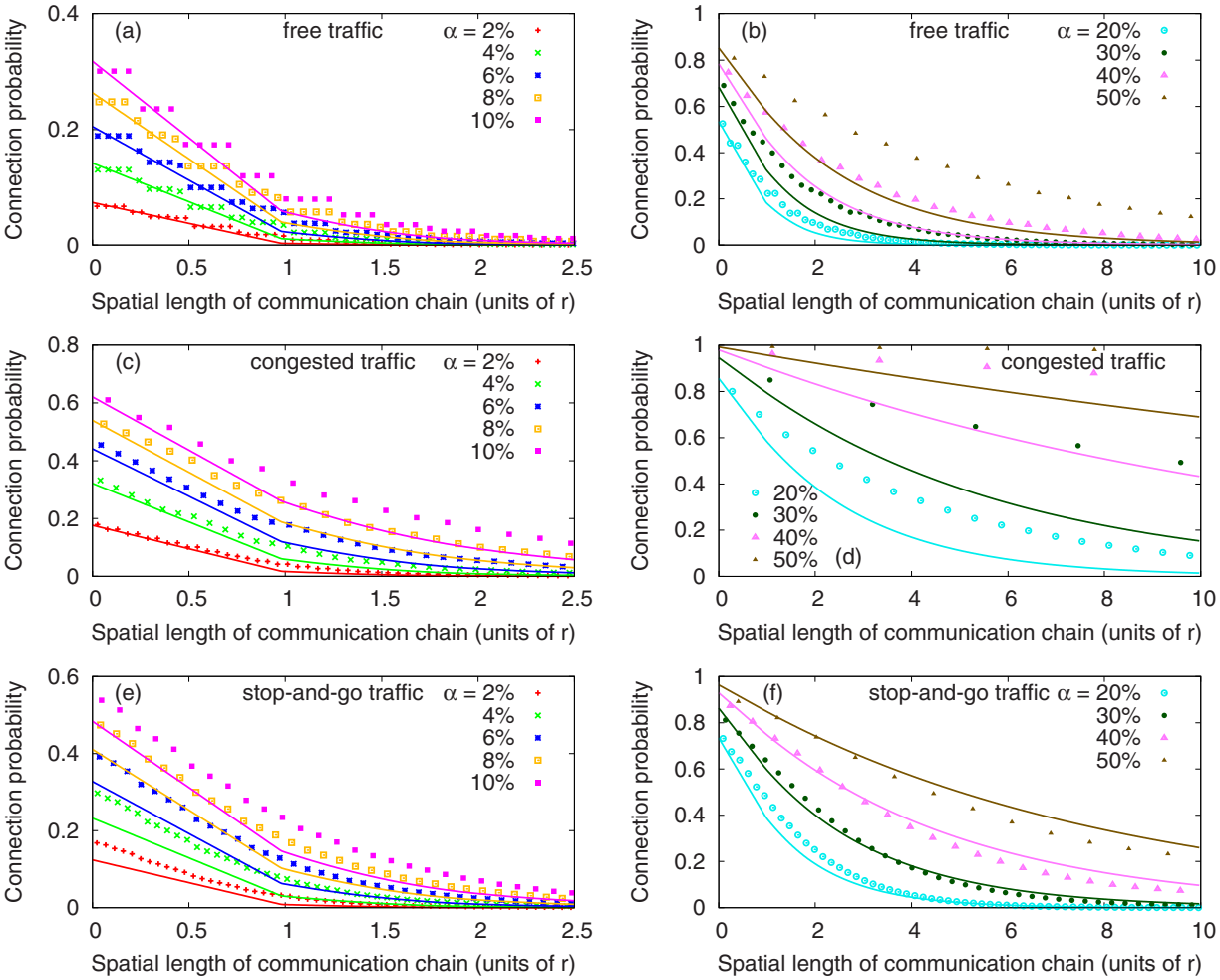

FIG. 6. (Color online) Probability of message reception at a certain distance from the initial sender for the stochastic transmission model based on simulated trajectories. The traffic situations are identical to the corresponding plots of Fig. 3. The lines represent the deterministic analytical result $P_{\text{chain}}(x)$.

stream. We have analyzed networks resulting from three different models for the node dynamics (independently positioned nodes, trajectories resulting from simulations of a car-following model, and real-world trajectory data), and two different assumptions for the links (fixed and stochastic communication ranges). For each type of network, we compared the resulting connectivities as a function of the penetration level and the type of traffic flow. Since analytical expressions are available for the simplest network (independently positioned nodes, fixed broadcast ranges) it is particularly interesting to compare the more realistic and complicated network variants to this reference.

As a main result, we have found that the analytical model agrees remarkably well with the simulated results of the more realistic model variants. Deviations can be explained by the assumptions made in the model. The assumption of exponentially distributed distances made in the analytical model neglects both the short-range correlations arising from the necessity to keep a certain minimum distance, and the long-ranged correlations caused by traffic instabilities such as stop-and-go traffic. However, two factors greatly decrease the consequences of these correlations. First, the (short-ranged) correlations decrease with the number of lanes since vehicle positions on different lanes are essentially independent. Moreover, even on single-lane roads, the random distribution of the equipped vehicles leads to a significantly reduced correlation of the relevant node positions (equipped vehicles) with respect to that for all vehicles. When comparing the communication models with stochastic and fixed

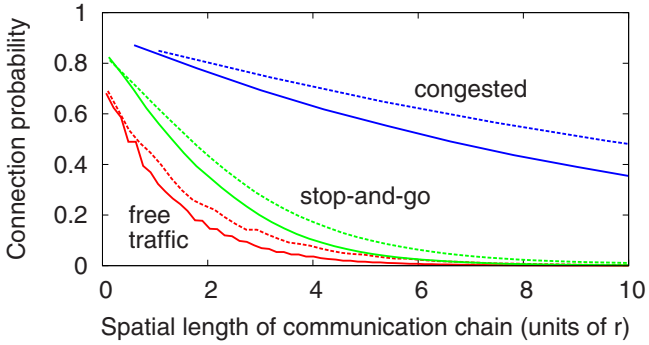

FIG. 7. (Color online) Direct comparison of the connectivities using the deterministic transmission model (solid lines) and the stochastic transmission model (dashed lines) for $r=200$ m and penetration level $\alpha=30\%$.

ranges, we found that the stochastic transmissions generally lead to a higher connectivity. This can be explained by the non-Markovian property of this model allowing additional links, but the effect is even smaller than that caused by the correlations. We conclude that the analytic expressions of the simplest model provide a good estimate, particularly for the low penetration levels that are practically relevant for the first stages of the introduction of IVC systems.

Simulations with all models consistently suggest that longitudinal hopping leads to communications chains of less than 0.5 km for a typical communication range of 200 m, and realistic equipment levels below 10%. In contrast, IVC systems using the alternative "store and forward" mode (using vehicles driving in the opposite direction to transport the messages) can be effective for penetration levels as low as 1%–2% [16], but the communication is not instantaneous.

[1] M. E. J. Newman, SIAM Rev. **45**, 167 (2003).
[2] J. Davidsen, H. Ebel, and S. Bornholdt, Phys. Rev. Lett. **88**, 128701 (2002).
[3] M. E. J. Newman, S. Forrest, and J. Balthrop, Phys. Rev. E **66**, 035101(R) (2002).
[4] V. Kalapala, V. Sanwalani, A. Clauset, and C. Moore, Phys. Rev. E **73**, 026130 (2006).
[5] L. Hufnagel, D. Brockmann, and T. Geisel, Proc. Natl. Acad. Sci. U.S.A. **101**, 15124 (2004).
[6] R. Albert, H. Jeong, and A.-L. Barabasi, Nature (London) **406**, 378 (2000).
[7] D. S. Callaway, M. E. J. Newman, S. H. Strogatz, and D. J. Watts, Phys. Rev. Lett. **85**, 5468 (2000).
[8] P. Holme, B. J. Kim, C. N. Yoon, and S. K. Han, Phys. Rev. E **65**, 056109 (2002).
[9] E. López, R. Parshani, R. Cohen, S. Carmi, and S. Havlin, Phys. Rev. Lett. **99**, 188701 (2007).
[10] R. Cohen, K. Erez, D. ben-Avraham, and S. Havlin, Phys. Rev. Lett. **85**, 4626 (2000).
[11] A. E. Motter and Y.-C. Lai, Phys. Rev. E **66**, 065102(R) (2002).
[12] R. Pastor-Satorras and A. Vespignani, Phys. Rev. E **65**, 036104 (2002).
[13] R. Cohen, S. Havlin, and D. ben Avraham, Phys. Rev. Lett. **91**, 247901 (2003).
[14] A. Kesting, M. Treiber, M. Schönhof, and D. Helbing, Transp. Res., Part C: Emerg. Technol. (to be published).
[15] X. Yang and W. Recker, Transp. Res., Part C: Emerg. Technol. **13**, 370 (2005).
[16] M. Schönhof, A. Kesting, M. Treiber, and D. Helbing, Physica A **363**, 73 (2006).
[17] W. Jin and W. Recker, Transp. Res., Part B: Methodol. **40**, 230 (2006).
[18] M. Schönhof, M. Treiber, A. Kesting, and D. Helbing, Transp. Res. Rec. **1999**, 3 (2007).
[19] O. Dousse, P. Thiran, and M. Hasler, in Proceedings of the IEEE Infocom. Twenty-First Annual Joint Conference of the IEEE Computer and Communications Societies, Vol. 2 (IEEE Operations Center, Piscataway, N. J., 2002) pp. 1079–1088.
[20] P. G. Gipps, Transp. Res., Part B: Methodol. **15**, 105 (1981).
[21] M. Treiber, A. Hennecke, and D. Helbing, Phys. Rev. E **62**, 1805 (2000).
[22] M. Treiber, A. Kesting, and D. Helbing, Physica A **360**, 71 (2006).
[23] W. Knospe, L. Santen, A. Schadschneider, and M. Schreckenberg, Phys. Rev. E **65**, 056133 (2002).
[24] Federal Highway Administration, Next generation simulation community, http://www.ngsim.fhwa.dot.gov.