Information Transmission over a Finite-Lifetime Channel under Energy-Constraints

Norbert Goertz and Johannes Gonter

Institute of Telecommunications
Vienna University of Technology
Gußhausstr. 25-29 / 389, 1040 Vienna, Austria
Email: {norbert.goertz, johannes.gonter}@nt.tuwien.ac.at

***Abstract*— In this work we investigate the limits on information-transmission in the finite-lifetime channel. We derive an upper bound for the information that can be transmitted, and we introduce an optimum acausal scheme for power-allocation. Furthermore, we introduce more practical but nevertheless efficient schemes for allocating power, and we compare their performance with the optimum scheme in a vehicle-to-vehicle communication scenario.**

I. Introduction

In inter-vehicle communication on motorways, the link between vehicles on opposite lanes is characterised by a channel that most of the time will not allow for any data transmission at all. Only when two vehicles approach each other, the receive power will be high enough to establish a connection; the same arguments apply for communication between a vehicle and a fixed access point at the roadside.

The channel will be highly time-variant and during its lifetime it will suffer from fast fading due to multipath propagation and Doppler shifts. Moreover, it seems reasonable to use a model with a long-term pathloss that, due to large distance, is very large most of the time, apart from a short period when the vehicles approach each other. On top of that we will observe fast fading, with values of the channel coefficients located "around" the long-term path loss.

Such a channel will not allow for time-continuous data transmission and the random process describing the time-series (or sequence) of the channel coefficients is not ergodic; a definition of channel "capacity" by the average of transmitted information over an infinite time-span would lead to the result of "zero".

Although more general definitions for the channel capacity have been stated in the literature (e.g., [1], [2], [3], [4], [5]) those results do not apply in a straightforward way to the problem at hand. Therefore, we propose a pragmatic approach to define theoretical limits: we consider the absolute amount of information that can be transmitted during the "lifetime" of the channel, and we introduce a suitable "rate normalisation" in a second step.

II. Channel Model

We will use a block fading model and i.i.d.[1] Gaussian noise as it is common in communcation theory. This means we assume that the fading-coefficient of the channel will remain constant within a block that contains "many" transmit symbols. If the block size, M, is "large enough", we can closely achieve the classical Gaussian channel capacity within the block.

As we want to avoid the definition of specific "on" and "off" times of the connection, we define the channel coefficients $h(i)$ for all time slots $i = \{-\infty, ..., -1, 0, 1, ... +\infty\}$, and the values of $h(i)$ will be used to implement the limited life-span of the channel (by appropriately chosen zero-values). Hence, we will assume hat $|h(i)| > 0$ only for a limited subset $\mathcal{S} \doteq \{i : |h(i)| > 0\}$ with $|\mathcal{S}| < \infty$.

In each block i the receive power $P_r(i)$ will equal the transmit power $P(i)$ scaled by the magnitude square $|h(i)|^2$ of the channel coefficient $h(i)$, i.e.,

$$P_r(i) = |h(i)|^2 \cdot P(i) \tag{1}$$

We assume coherent detection, i.e., we need full channel knowledge at the receiver, which we will always assume.

III. Limits on Information Transmission

The capacity $C(i)$ (in bits per channel-use) within block i is given by

$$C(i) = \log_2 \big(1 + \frac{|h(i)|^2 \cdot P(i)}{2\sigma^2}\big) \tag{2}$$

where σ^2 is the variance of the Gaussian receiver noise in each real component of the transmit channel (inphase and quadrature component); by definition of (2) we implicitly assume a carrier-modulated system, i.e., complex modulation. The total power $P(i)$ is spread equally across the two real channel parts, as this will, for symmmetric channels, achieve capacity (e.g., [6]).

Let us now assume that each block i contains M uses of the channel (with M very large). Then we can transmit at most

$$N(i) = C(i) \cdot M \tag{3}$$

[1] independent and identically distributed

bits of information in block i. Hence, the total number N of bits of information we can transmit through the channel equals

$$N = M \sum_{i=-\infty}^{\infty} C(i) = M \sum_{i=-\infty}^{\infty} \log_2 \big(1+\frac{|h(i)|^2 \cdot P(i)}{2\sigma^2}\big). \quad (4)$$

Note that N will not be unbounded as we have assumed that $|h(i)| > 0$ only for a limited number $|\mathcal{S}| < \infty$ of blocks i. Hence, we can equivalently write

$$N = M \sum_{\forall i \in \mathcal{S}} \log_2 \big(1 + \frac{|h(i)|^2 \cdot P(i)}{2\sigma^2}\big) . \quad (5)$$

IV. Allocation of Transmit Power

Although the assumption of full acausal (*non*-causal) deterministic channel knowledge for transmit power allocation is not realistic in practice, it provides an interesting upper performance bound for practical transmission systems including coding, modulation and power allocation. For a channel with limited lifetime (as in vehicle-to-vehicle communication) we can record the full history of a real-world channel by measurement und compute the largest possible amount of information that could have been transmitted over this channel, thereby, evaluating and comparing the performance of practical systems that we run on the same set of channel coefficients.

A. Optimisation Problem

We aim to maximise the number N of bits in (5), given $|h(i)|$ for all $i \in \mathcal{S}$. We define the total energy used for the transmission of the information bits by

$$E = \sum_{\forall i \in \mathcal{S}} MP(i) = M \sum_{\forall i \in \mathcal{S}} P(i) . \quad (6)$$

If we limit the value of the total energy according to

$$E \leq E_0 , \quad (7)$$

we can define a constrained optimisation problem. We set up the functional L with the Lagrange multiplier $\lambda > 0$:

$$L \doteq \frac{M}{\log(2)} \sum_{\forall i \in \mathcal{S}} \log \big(1+\frac{|h(i)|^2 \cdot P(i)}{2\sigma^2}\big)+\lambda\big(E_0-M \sum_{\forall i' \in \mathcal{S}} P(i')\big) \quad (8)$$

(with "log" the natural logarithm). Now we take the derivative[2] for the unknown transmit power P_j in block $j \in \mathcal{S}$:

$$\frac{\partial L}{\partial P(j)} = \frac{M}{\log(2)} \frac{1}{1 + \frac{|h(j)|^2 \cdot P(j)}{2\sigma^2}} \cdot \frac{|h(j)|^2}{2\sigma^2} - \lambda \cdot M . \quad (9)$$

We set the derivative to zero, cancel $M \neq 0$, set $\tilde{\lambda} \doteq \lambda \log(2)$ and obtain

$$P^*(j) = \Big(\frac{1}{\tilde{\lambda}} - \frac{2\sigma^2}{|h(j)|^2}\Big)^+ \quad (10)$$

which is a "waterfilling" solution (e.g. [7]). In (10) we have introduced a "max"-operation according to $(x)^+ \doteq \max(0, x)$ to ensure that the solutions for the powers do not take negative values; the Karush-Kuhn-Tucker [8] conditions guarantee that (10) is still an optimal solution to our problem.

The value of $\tilde{\lambda}$ must be chosen such that (7) is fulfilled. The maximum number of information bits N^* is given by substituting (10) into (4) with $P(i) = P^*(i)$.

For a measured sequence $\{h(i)\}$ of channel coefficients that we know in advance we can find $\tilde{\lambda}$ numerically by trying a value for it, solving (6) and checking if E is close to the pre-specified value of E_0. If $E > E_0$ we have to increase $\tilde{\lambda}$ to get closer to E_0: this process can be iteratively repeated to get the optimal solution for $\tilde{\lambda}$ for this particular sequence of channel coefficients.

B. Normalisation of the Solution

The problem and its solution can be rewritten equivalently by dividing (6) by M and $2\sigma^2$. We obtain

$$\frac{E/M}{2\sigma^2} \;=\; \sum_{\forall i \in \mathcal{S}} \frac{P(i)}{2\sigma^2} \;\leq\; \frac{E_0/M}{2\sigma^2} . \quad (11)$$

We can interpret $\frac{P(i)}{2\sigma^2} \doteq SNR(i)$ as a "transmit" signal-to-noise ratio (as it does not contain the channel fading factor $h(i)$) for block i. The ratio E/M on the left-hand side can be interpreted as the average transmit power for each use of an equivalent substitute channel that would transmit the same amount of information but use only one block (also of size M) instead of $|\mathcal{S}|$ blocks of size M that are used in the channel. Hence, we obtain an alternative form of (11) according to

$$\frac{E/M}{2\sigma^2} \doteq SNR \;=\; \sum_{\forall i \in \mathcal{S}} SNR(i) \;\leq\; \frac{E_0/M}{2\sigma^2} . \quad (12)$$

The solution for the optimal transmit $SNR^*(i)$ follows directly from (10) by dividing through the non-negative (total) noise variance $2\sigma^2$. The solution reads[3]

$$SNR^*(i) = \Big(\frac{1}{\lambda} - \frac{1}{|h(i)|^2}\Big)^+ , \quad \lambda > 0 \quad (13)$$

with the upper limit for the largest possible number N^* of information bits given by

$$N^* = M \sum_{\forall i \in \mathcal{S}} \log_2 \big(1 + |h(i)|^2 SNR^*(i)\big) , \quad (14)$$

and λ in (13) has to be chosen (typically by a numerical approach) such that with $SNR(i) = SNR^*(i)$ (12) is fulfilled.

As it is inconvenient to deal with an absolute number N^* of information bits we introduce the relative quantity

$$C^* \doteq \frac{N^*}{M} = \sum_{\forall i \in \mathcal{S}} \log_2 \big(1 + |h(i)|^2 SNR^*(i)\big) \quad (15)$$

that we interpret as a "capacity" measured in *information bits per use of the equivalent substitute channel* that transmits a single block of size M only. As long as the lifetime of the channel is limited (i.e., $|\mathcal{S}| < \infty$), the sum in (15) will always converge.

[2] To avoid confusion with the sum index i, we change the index to j for the power for which we take the derivative.

[3] Although the Lagrange multiplier λ in (13) is different from that in (8) we use the same variable-name as there is no risk of confusion.

C. Causal Knowledge of the Channel Coefficients

For causal channel knowledge at the transmitter, i.e. the case we will face in a practical system, it seems impossible to solve *in advance* for the value of λ that, via "power" allocation in (13), will satisfy (12) (with $SNR(i) = SNR^*(i)$), because the future channel coefficients $h_{i+1}, h_{i+2}, ...$ are unknown at time i.

a) Non-adaptive scheme: In the ergodic case and with full statistical channel knowledge we could simulate by a random generator a "fake-sequence" of channel coefficients which would exhibit the same statistical properties as the true sequence of channel coefficients. Therefore, we could use the fake-sequence of channel coefficients to determine λ, and this value of λ would then be used, together with the true value of the current channel coefficient, to allocate transmit power for the current transmit block.

This approach will not work perfectly in our setting, though, as the random process is not ergodic and of limited size. If the lifetime of the channel is "large enough", however, the idea might still be helpful to find a good approximation. A possible approach would be to take a simulated fake-sequence, compute λ and use it for power allocation. As λ won't fit the real set of channel coefficients, there is a risk that more power would be consumed. This, however, could be prevented by just stopping power allocation, once the given energy-budget has been used up. More difficult would be a situation, in which the power-budget available is *not* fully used. In this case λ is too large for the real set of channel coefficients, but once the "peak"-values of the channel coefficients have passed, there is hardly any chance, due to waterfilling with the incorrect value of λ, to allocate any power in the future, and, once we realise the problem as we run along the time axis, it is too late to make efficient use of power as then all channel coefficients tend to have small values. We will still use this "*non-adaptive*" scheme for comparison in the simulations.

b) Adaptive scheme: A more elaborate, *adaptive* strategy to solve the problem is to re-calculate λ for every block, using the newly available current channel coefficent $h(i)$ and append a fake-sequence $\{h'(j), j = i+1, ..., |\mathcal{S}|\}$ of *future* channel coefficients. With every new known channel coefficient $h(i)$, the corresponding value in the fake sequence is replaced and the new sequence of channel coefficients is used for power allocation in the current block. After taking the decison what part of the available energy is to be used in the current block, we have to update the remaining energy budget that is left for the future use. If $|\mathcal{S}|$ – the lifetime of the channel – is sufficiently large, the error introduced by this procedure will (hopefully) be small and we will use exactly the energy-budget available. The price to pay is the complexity for an update of λ for every transmit block.

In the simulations below, we will use this *adaptive* approach and we compare the results with the acausal "genie"-solution according to Section IV-A.

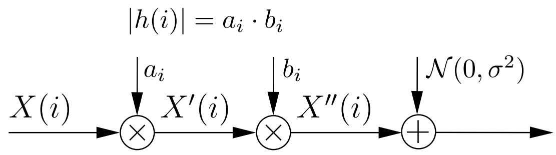

Fig. 1. Channel model used for simulations; depicted is one of two independent, statistically identical parallel transmit channels ("I" or "Q"-path).

V. Simulation Results

A. Channel Model

For a numerical comparison of the power allocation schemes discussed in Section IV, we use the channel model depicted in Fig. 1. We assume zero-mean Gaussian receiver noise with a variance of σ^2 in each real channel component. Due to coherent detection, there are no phase rotations at the receiving end and we only obtain an amplitude scaling that is equal in both the inphase and the quadrature component of the channel.

The channel coefficients in our model are composed of a long-term path loss a_i and a factor b_i that is to model "fast" fading, due to intersymbol interference and multipath propagation. Both factors a_i, b_i are assumed to be fixed during each block i, with the path-loss a_i changing between blocks but much slower than the fading factor b_i. For simplicity we assume Rayleigh-fading for the random variable B_i with the pdf

$$f_{B_i}(b_i) = \frac{b_i}{\sigma_b^2} e^{-b_i^2/2\sigma_b^2}\Big|_{\sigma_b^2=1/2} = 2b_i e^{-b_i^2} \ . \qquad (16)$$

We use $\sigma_b^2 = 1/2$ in our model, so $\mathrm{E}\{b_i^2\} = 1$, i.e., the Rayleigh-fading component does not change average power of the signal, i.e.., $\mathrm{E}\{X'^2(i)\} = \mathrm{E}\{X''^2(i)\}$ (see Figure 1).

For the long-term path loss we assume a deterministic function which represents the physical model of two cars on opposite lanes on the highway which pass each other at $v/2 = 30$m/s each, at a minimum distance to each other of $\delta = 30$m. The channel lifetime is assumed to be 6s in total, i.e. $T_s = 3$s to either side of the symmetric scenario, and the pathloss exponent is assumed to be $\gamma = 2$. The pathloss can therefore be calculated according to

$$a_i = \begin{cases} 0 & , \quad i < 1 \\ \delta^\gamma \cdot \left(\delta^2 + v^2 \cdot \left((i-\Delta) \cdot \frac{T_s}{\Delta}\right)^2\right)^{-\gamma/2} & , \quad 1 \le i \le 2\Delta \\ 0 & , \quad i > 2\Delta \end{cases} , \qquad (17)$$

so the channel lifetime equals 2Δ and the communication partners closely approach each other around time $i = \Delta$ (the factor δ^γ amounts to the normalization of the minimum attenuation to 0dB).

In practice, a_i will also be a random variable, although a system model (such as equations of motion) may underlie it. The intention with our approach to use a deterministic function is to rate the performance of the more practical schemes proposed in Sections IV-C against the "genie" performance as obtained from acausal channel knowledge (Section IV-A).

B. Numerical Results

In Figures 2, 3, and 4 we show some simulation results that were obtained for $\Delta = 100$, i.e., for a channel lifetime of 200 blocks, and a sum SNR of 20dB according to (12). We compare the acausal power allocation scheme (Section IV-A) in Fig. 2 with the two causal schemes described in Section IV-C: in Fig. 3 we show results for power allocation using λ (for use in 13) that we calculate in advance from a "fake" coefficient sequence of length 2Δ. In Fig. 4 we show the results for block-apdative λ as also described in Section IV-C.

A channel coefficient is used for transmission if the magnitude square $|h(i)|^2$ of the coefficient is larger than λ and this is indicated by "circles" in the upper plots of all figures.

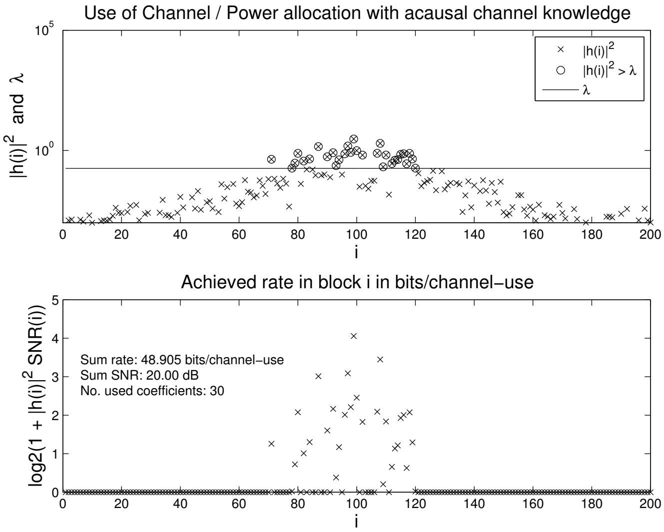

Fig. 2. Simulation of a vehicle-to-vehicle with optimal acausal power allocation; upper plot: channel coefficients with indication which coefficients are used for transmission, lower plot: achieved rates; channel lifetime is 200 blocks, the sum transmit SNR is limited to 20dB.

The optimal acausal scheme in Fig. 2 is the benchmark as it indicates the best possible theoretical performance. Note that an SNR of 20dB is not very much as it is a sum SNR over 200 blocks (according to (12)). Similarly, the achieved rate of almost 49 bits per channel-use for the equivalent single-block transmission is also a sum rate over all blocks. The achieved rates for the individual transmit blocks are shown in the lower part of Fig. 2: those numbers are not higher than 4 bits per channel use and, hence, reasonable for a practical implementation, e.g., by QAM modulation. As we know the full sequence of channel coeffcients in advance we can determine λ such that exactly 20dB sum SNR are comsumed by the power allocation.

The situation is generally different when we calculate λ using a "fake" sequence from a random generator that has exactly the same statistics as the original channel (non-adaptive approach from Section IV-C). The results in Fig. 3 show that (due to limited blocksize) λ is slightly too large (in this example), so the allocated power is slightly too low

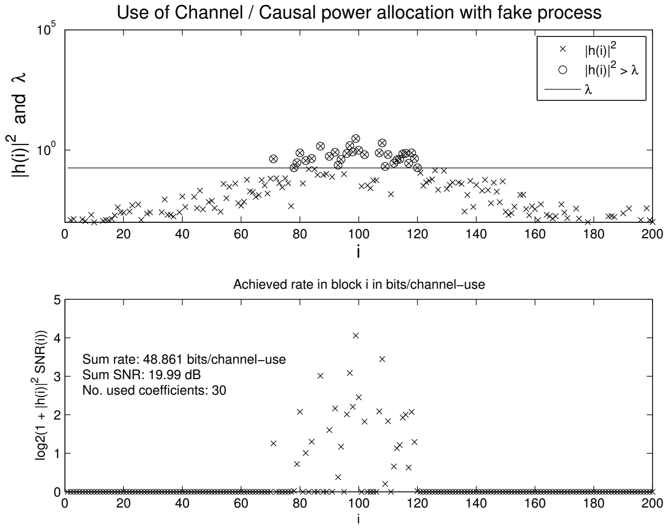

Fig. 3. Simulation of a vehicle-to-vehicle channel with causal power allocation using a fake random process (with correct statistics) with non-adaptive λ; upper plot: channel coefficients with indication which coefficients are used for transmission, lower plot: achieved rates; channel lifetime is 200 blocks, the sum transmit SNR is limited to 20dB.

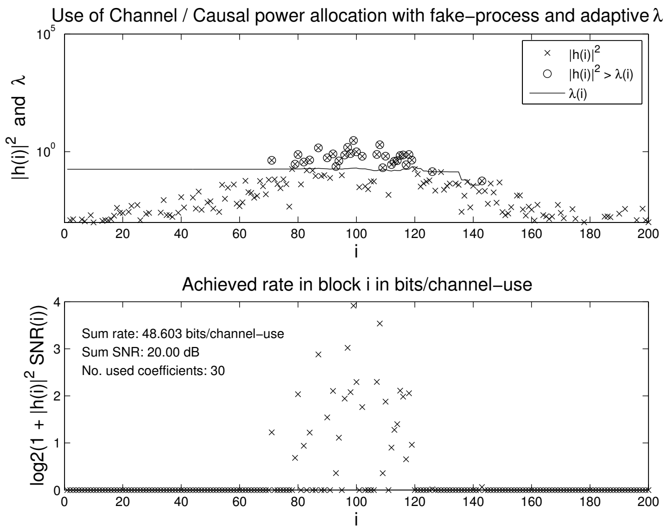

Fig. 4. Simulation of a vehicle-to-vehicle channel with causal power allocation using a fake random process (with correct statistics) with adaptive λ; upper plot: channel coefficients with indication which coefficients are used for transmission, lower plot: achieved rates; channel lifetime is 200 blocks, the sum transmit SNR is limited to 20dB.

(19.99dB instead of 20.00). However, this small mismatch is not really relevant and also the loss in sum rate is marginal (48.86 bits per equivalent channel-use instead of 48.91 in the acausal case).

In Fig. 4 we show the causal scheme with adaptation of λ in every block i as described in Section IV-C. The solid line indicates the values of λ in the blocks; at around $i = 140$ the line disappears from the plot: this is to indicate that at this point the power budget has been used up and transmissions stop. As indicated by the lower plot in Fig. 4 the adaptive scheme can ensure that the given power budget fully is used. Moreover, the loss in rate compared to the acausal scheme is only 0.3 bits/channel use and, hence, rather small. However, the adaptive scheme also makes use of a fake sequence, and if the statistics of the fake sequence are significantly different from the true statistics of the channel, we observe a loss in performance. This is demonstrated by Fig. 5, where a path loss was assumed to generate the fake sequence that was 10dB lower than in the real channel. Although we face a significant

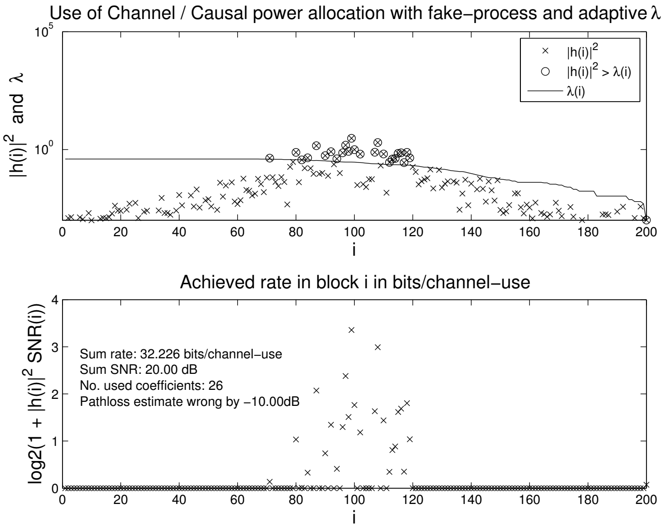

Fig. 5. Simulation of a vehicle-to-vehicle channel with causal power allocation using a fake random process with adaptive λ; upper plot: channel coefficients with indication which coefficients are used for transmission, lower plot: achieved rates; channel lifetime is 200 blocks, the sum transmit SNR is limited to 20dB. The channel statistics used at the transmitter are incorrect: the path loss is assumed to be 10dB too low.

loss in rate and achieve only 32 bits per channel-use (instead of almost 49 bits per channel use in the acausal optimal case), the non-adaptive scheme (from Section IV-C) would, under the same conditions, achieve only 20.26 bits per channel use and use a sum transmit SNR of only 13.52dB instead of 20dB. This is illustrated in Fig. 6.

Of course the simulation results dicussed above are only snapshots for specific realisations of the channel coefficient sequences. They are useful to understand the operation of the schemes and to explain effects observed but for a fair performance comparison averaging over many channel realisations is required. The results are given in Fig. 7 for a range of sum SNRs of 0....30dB with $\Delta = 500$ (i.e., channel lifetime is 1000) and averaging is carried out over 1000 realisations of a full channel sequence. We observe that both causal schemes (adaptive and non-adaptive) operate close to the theoretical performance limits as long as the channel statistics are known exactly at the transmitter. Moreover, for the simulation in Fig. 7 a relatively long channel lifetime of 1000 blocks is assumed, so a realisation (sequence of coefficients) will reveal most of the statistical properties of the random process. This

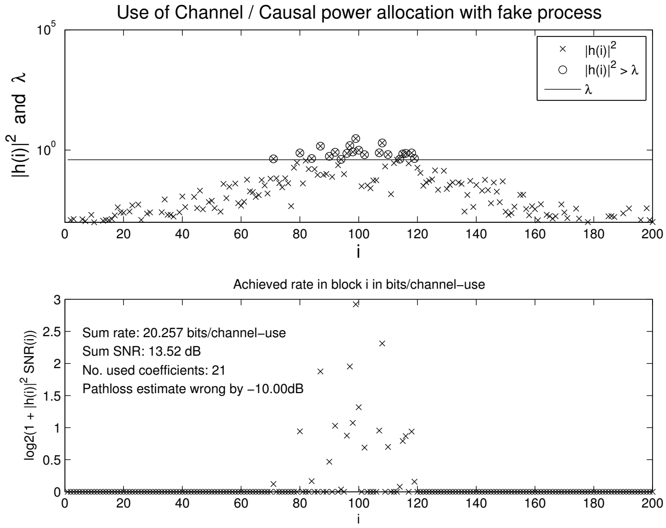

Fig. 6. Simulation of a vehicle-to-vehicle channel with causal power allocation using a fake random process with non-adaptive λ; upper plot: channel coefficients with indication which coefficients are used for transmission, lower plot: achieved rates; channel lifetime is 200 blocks, the sum transmit SNR is limited to 20dB. The channel statistics used at the transmitter are incorrect: the path loss is assumed to be 10dB too low.

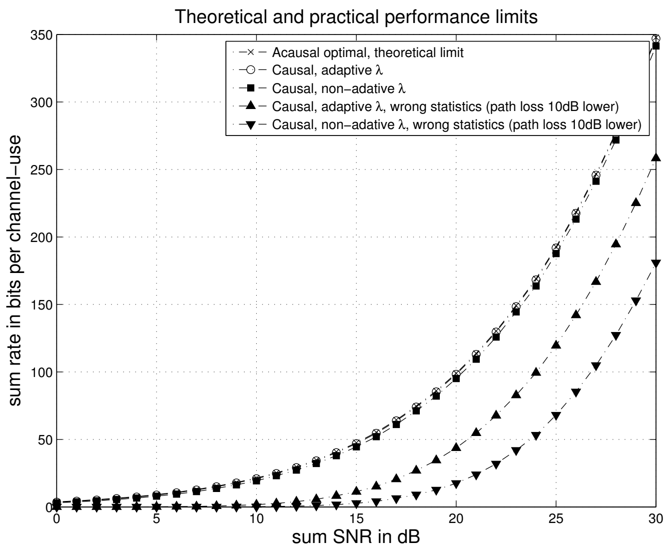

Fig. 7. Comaprison of acausal and causal power allocation schemes with and without adaptation of λ. Simulation results are for a channel lifetime of 1000 blocks and they are averaged over 1000 simulated channels. Results for incorrect statistical knowledge at the transmitter are also included.

explains, why both causal schemes perform very well. The picture changes, however, when the statistical properties are not known exactly. This is demonstrated by the two lower curves in Fig. 7 where we have assumed that the path loss estimate at the transmitter is 10dB too low (i.e., the real signal is much weaker than expected). This means that the initial λ-value used in the causal schemes is too large, so channel coefficients that are actually large relative to the true channel statistics are not used for transmission. Hence, in the non-adaptive scheme the power budget can not be fully used leading to rather bad performance. The adaptive scheme, which learns the process parameters stepwise with each new incoming channel coefficient, reduces the λ value and, thereby, will spend the given power budget in full and eventually will achieve a much better sum rate than the non-adaptive scheme.

VI. Conclusions

We have stated theoretical limits on the total amount of information that can be transmitted through a vehicle-to-vehicle communication channel. We have stated an optimal algorithm for transmit power allocation for the case of acausal channel knowledge at the transmitter and we introduced pragmatic schemes for power allocation for the more realistic case that only causal channel knowledge is available at the transmitter. The block-adaptive causal scheme was demonstrated to have superior performance, particularly in the case of incorrect knowledge of the statistics of the long-term path loss.

References

[1] M. Effros, A. Goldsmith, and Y. Liang, "Generalizing capacity: New definitions and capacity theorems for composite channels," *IEEE Transactions on Information Theory*, vol. 56, no. 7, pp. 3069–3087, July 2010.

[2] E. Biglieri, J. Proakis, and S. Shamai (Shitz), "Fading channels: Information-theoretic and communications aspects," *IEEE Transactions on Information Theory*, vol. 44, no. 6, pp. 2619–2692, Oct. 1998.

[3] J. C. Kieffer, "ϵ-Capacity of binary symmetric averaged channels," *IEEE Transactions on Information Theory*, vol. 53, no. 1, pp. 288–303, Jan. 2007.

[4] S. Verdu and T. S. Han, "A general formula for channel capacity," *IEEE Transactions on Information Theory,*, vol. 40, no. 4, pp. 1147–1157, July 1994.

[5] L. H. Ozarow, S. Shamai, and A. D. Wyner, "Information theoretic considerations for cellular mobile radio," *IEEE Transactions on Vehicular Technology*, vol. 43, no. 2, pp. 359–378, May 1994.

[6] T. Cover and J. Thomas, *Elements of Information Theory*, John Wiley & Sons, Inc., 1991.

[7] D. Tse and P. Viswanath, *Fundamentals of Wireless Communications*, Cambridge University Press, 2005.

[8] H. W. Kuhn and A. W. Tucker, "Nonlinear programming," in *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability (available online at http://projecteuclid.org)*, Berkeley, CA, USA, July/August 1950, pp. 481–492.