On the Pilot Spacing Constraints for Continuous Time-Varying Fading Channels

Stefano Savazzi, *Student Member, IEEE*, and Umberto Spagnolini, *Senior Member, IEEE*

***Abstract*—In common wireless systems pilot placing can be interpreted as a way to sample the channel with some degree of accuracy. In this letter we investigate the necessary conditions on the pilots spacing for time-varying fading to guarantee a specified average bit error rate. These constraints are evaluated in closed form for large SNR and small fading dynamics and specialized for varying fading correlation and coded/uncoded binary transmissions.**

***Index Terms*—MMSE channel estimation, pilot spacing, time-varying fading channels, training-based channel estimation.**

I. Introduction

In wireless communications, reliable coherent reception is guaranteed as long as the channel estimation exhibits a sufficiently high quality. In most practical systems, the channel gain is acquired by using pilot symbols known to the receiver and multiplexed within the data stream to convey the necessary robustness against channel variations.

Design of training length has been assessed in [1] for the block-fading case to maximize the mutual information. The problem of optimizing both the interval among subsequent training phases and the training length arises when the channel estimate becomes outdated due to time-varying fading. The optimal training length and interval are numerically found in [2]-[4] to maximize the fraction of time spent in transmission and in [5] to minimize the maximum Bit Error Rate (BER) for uncoded transmission and given the number of pilots. It is shown therein that for time-varying fading and high Signal to Noise Ratio (SNR), the optimal training sequences consist of one pilot symbol each so that the fading process is sampled with a given frequency. The interplay between spectral efficiency, transmit power and fading correlation in terms of channel capacity has been recently and independently considered in [6].

In this paper transmission is organized as shown in Fig. 1: each pilot symbol is followed by a block of Δ_s information symbols so that the time-varying faded channel is sampled with a frequency $f_s = \Delta_s^{-1} \leq 1$ [pilots/symbols]. Information on the channel state is conveyed here by MMSE interpolation of the two ($P = 2$) estimates from the pilots that are closest in time[1]. Given that pilots sample the fading variations with a given accuracy (that depends on the pilot symbols energy),

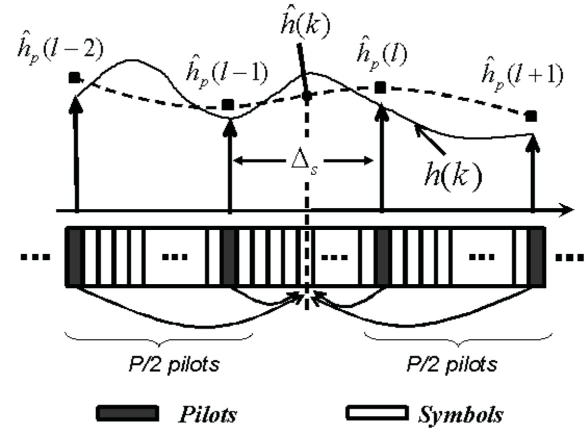

Fig. 1. Training structure and MMSE channel estimation from P pilots ($P/2$ pilots are previous while the remaining $P/2$ are upcoming to information symbols, here $P = 4$ pilots).

we investigate here the necessary conditions on the required (minimum) *sampling frequency* of fading f_s (or pilots spacing) to guarantee a level of accuracy (e.g., an average BER w.r.t. fading) for data detection. Notice that, when the interval between pilots Δ_s is small (for large frequency f_s), the fading is largely oversampled to minimize the out-dating of the channel estimate and improve the performances. In addition to previous works, sampling frequencies f_s are derived here in closed form for high SNR and small fading dynamics to guarantee an average BER at the receiver. Moreover, results are shown for an *arbitrary* fading correlation, for uncoded and bit interleaved coded BPSK transmissions.

This letter is organized as follows, the system and the channel estimation model are introduced in Sect. II. In Sect. III we evaluate the conditions on the required sampling frequency f_s in closed form. Fading sampling frequencies are derived in terms of the fading auto-correlation function (that describes the fading dynamics), the information and pilot symbol energies, the required BER and the coding gain. Next, results are specialized for Gauss-Markov AR-1 fading dynamics [9] and corroborated by numerical evaluation (see also [3] for the BPSK uncoded case).

II. System and Channel Model

Information symbols of duration T_s are sent from a single antenna transmitter over a narrowband channel to a single antenna receiver, the time interval among pilot symbols (see Fig. 1) is $\Delta_t = \Delta_s \times T_s$. The statistical properties of the channel (such as the fading distribution and correlation) are known to the transmitter and the receiver. At time $t_k = kT_s$, where index $k \in 1, ..., \Delta_s$ refers to a symbol position within

Paper approved by G. K. Karagiannidis, the Editor for Fading Channels and Diversity of the IEEE Communications Society. Manuscript received February 15, 2008; revised March 15, 2009.

The authors are with the Dipartimento di Elettronica e Informazione, Politecnico di Milano, I-20133 Milano, Italy (e-mail: {savazzi, spagnoli}@elet.polimi.it).

Digital Object Identifier 10.1109/TCOMM.2009.11.080024

[1]Interpolation improves the MSE compared to sampling and holding estimation [1]-[3] without any significant penalty in estimation complexity and delay.

a block (of Δ_s symbols), the baseband received signal at the output of the matched filter is

$$y(k) = h(k)\sqrt{E_s}x(k) + w(k), \quad (1)$$

E_s is the energy of the information symbol $x(k)$ with $\mathbb{E}[\,|x(k)|^2] = 1$. The baseband received signal over pilot symbol x_p at time $t_\ell = \ell(\Delta_s + 1)T_s$ ($\ell = 0, 1, ...$) is

$$y_p[\ell(\Delta_s + 1)] = h_p[\ell(\Delta_s + 1)]\sqrt{E_p}x_p + w[\ell(\Delta_s + 1)]. \quad (2)$$

with symbol energy E_p. The additive white Gaussian (AWG) noise samples $w \sim CN(0, \sigma^2)$ are uncorrelated, $\sigma^2 = N_0/T_s$ and N_0 is the single sided noise power spectral density. Fading impairments for pilot and information symbols $h, h_p \sim CN(0, 1)$ are modelled as unit power complex-normal distributed (Rayleigh fading) stationary processes. The fading correlation over k signalling intervals

$$\mathcal{R}(k) = E[h^*(n)h(n + k)] \quad (3)$$

depends on the dynamics of the channel (e.g., the Doppler frequency) at hand. As an example, the innovation process of the channel state $h(k)$ can be modelled by a first-order Gauss-Markov (AR-1) process [9] with $h(k) = \rho h(k-1) + u(k)$, where ρ is the fading correlation between two adjacent symbols and driving noise is $u(k) \sim CN(0, 1 - \rho^2)$. In this case the fading correlation over k symbols is $\mathcal{R}(k) = \rho^k$.

The SNR over information symbols is $\mu = E_s/N_0$, while the SNR over the pilot symbols is $\mu_p = E_p/N_0$ and the pilot-to-symbol power ratio is $\beta_p \equiv \mu_p/\mu$, with $\beta_p \geq 1$ to make pilots beneficial in data decoding [10].

A. MMSE channel interpolation in time-varying fading

In this paper the channel gain $h(k)$ at time $k = 1, ..., \Delta_s$ is estimated by combining the channel estimates obtained from $P = 2$ pilots (the previous and the upcoming pilots). For completeness, the MSE of channel estimation is shown in Fig. 2-(b) assuming in general that P pilots ($P/2$ previous and $P/2$ upcoming pilots, $P = 2, 4, 6$ ...) are used to acquire the MMSE channel estimate at time $k = \Delta_s/2$ (at the center of one sub-block of Δ_s symbols where MSE has the maximum value). As far as the fading is time-varying (e.g., for AR-1, $\rho < 1$) and the SNR μ_p over the pilots is large enough ($\mu_p > 15dB$) there is no meaningful advantage in using more than two pilots ($P = 2$) for channel interpolation. In practice, the use of $P > 2$ pilots for channel estimation is not justified from the benefits in BER probability: this motivates the choice of limiting our analysis to the practical case of $P = 2$.

The Maximum Likelihood (ML) estimate of the channel from pilot ℓ (by neglecting the timing reference Δ_s) is unbiased as

$$\hat{h}_p(\ell) = h_p(\ell) + \Delta h_p \sim CN(h_p(\ell), \mu_p^{-1}), \quad (4)$$

where estimate error is Δh_p and the quality of the fading sampling is the mean square error (MSE) $\mathbb{E}[|\Delta h_p|^2] = \mu_p^{-1}$.

The estimated channel at the k-th information symbol is modelled as $\hat{h}(k) = h(k) + \Delta h(k)$ with MSE $\sigma_e^2(k) = \mathbb{E}[\left|\hat{h}(k) - h(k)\right|^2]$. For $P = 2$ the channel interpolation $\hat{h}(k)$ is based on the samples $\hat{h}_p(\ell)$ and $\hat{h}_p(\ell + 1)$ from the

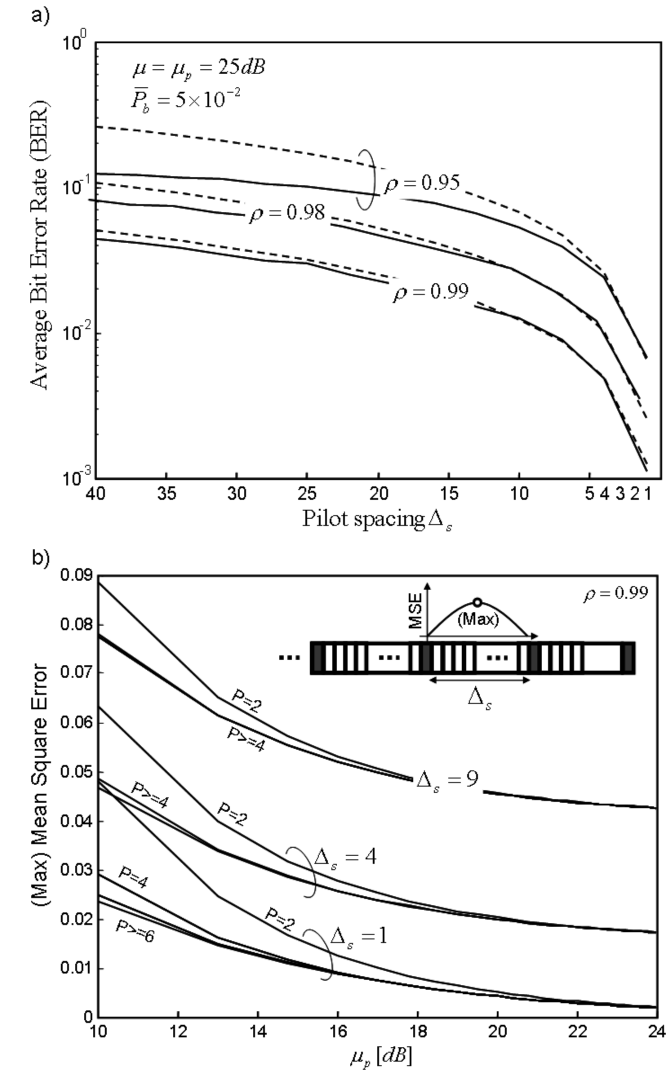

Fig. 2. Average BER (a) for uncoded BPSK versus the pilot spacings Δ_s and for varying AR-1 fading correlation ($\rho = 0.95, 0.98, 0.99$): simulated performances (solid lines), large SNR (and small Doppler) approximation (10) (dashed lines). MSE performances (b) of MMSE estimation for symbols at the center of the sub-block of Δ_s symbols (where the MSE has the maximum value) versus the SNR over the pilots (μ_p), for varying number of pilots ($P = 2, 4, 6...$), pilot spacing ($\Delta_s = 1, 4, 9$) and AR-1 fading correlation $\rho = 0.99$.

pilots. Given an arbitrary correlation function $\mathcal{R}(\cdot)$, the MMSE estimator is

$$\hat{h}(\alpha\Delta_s) = \\ = q(\alpha, \Delta_s|\mathcal{R})\hat{h}_p(\ell) + q(1-\alpha, \Delta_s|\mathcal{R})\hat{h}_p(\ell + 1) \quad (5)$$

where, for analytical convenience, $k = \alpha\Delta_s$ with $\alpha \in (0, 1)$. Weighting $q(\cdot)$ is

$$q(\alpha, \Delta_s|\mathcal{R}) = \frac{\mathcal{R}(\alpha\Delta_s) - \mathcal{R}(\Delta_s)\mathcal{R}[(1-\alpha)\Delta_s]}{1 - \mathcal{R}^2(\Delta_s)} \quad (6)$$

The MSE $\sigma_e^2(\alpha\Delta_s)$ profile is influenced by channel outdating $var[\hat{h}(\alpha\Delta_s)|h_p(\ell), h_p(\ell + 1)]$ and by estimation errors

$E[|\Delta h_p)|^2] = \mu_p^{-1}$:

$$\sigma_e^2(\alpha\Delta_s|\mathcal{R}) = \\ = 1 - \frac{\mathcal{R}^2(\alpha\Delta_s) + \mathcal{R}^2((1-\alpha)\,\Delta_s) - 2\mathcal{S}(\alpha)}{1-\mathcal{R}^2(\Delta_s)} + \frac{\Psi(\alpha)}{\mu_p}. \quad (7)$$

where $\mathcal{S}(\alpha) = \mathcal{R}(\alpha\Delta_s)\mathcal{R}(\Delta_s)\mathcal{R}((1-\alpha)\,\Delta_s)$ and $\Psi(\alpha) = q(\alpha,\Delta_s|\mathcal{R})^2 + q(1-\alpha,\Delta_s|\mathcal{R})^2$. Notice that pilot symbol energy (or SNR μ_p) rules the quality of the sampling.

III. Minimum Sampling Frequency of Fading

The minimum required sampling frequency of fading (or pilots spacing) is found to guarantee an average (w.r.t. fading statistic) BER at the receiver $\bar{P}_b$. Probability $\bar{P}_b$ is related to the particular application at hand. The average BER is computed for uncoded and bit interleaved coded BPSK transmission. The impact of channel estimation error on performance is assessed for large enough SNR and small fading dynamics (e.g., low Doppler).

The baseband received signal (1) at symbol k can be restated as

$$y(k) = \hat{h}(k)\sqrt{E_s}x(k) + \bar{w}(k), \quad (8)$$

where the impairments $\bar{w}(k) = \Delta h(k)\sqrt{E_s}x(k) + w(k)$ are taken from independent Gaussian random variables so that the conditional error probabilities can be conveniently expressed by using the $Q(\cdot)$ function as in [2] and [11]. Conditions for the minimum sampling frequency of fading are derived for arbitrary fading correlation $\mathcal{R}(\cdot)$ and for uncoded and coded transmissions: to ease system design, results are shown for varying SNRs (μ) over information symbols and over pilot symbols ($\mu_p = \beta_p\mu$). We show how requirements on the minimum sampling frequencies can influence the energy allocation (e.g., the SNRs ratio β_p) between training and data. The reader might refer to [3] and [7] for a more detailed discussion on the optimal allocation of energy between pilot and data symbols.

A. Uncoded BPSK

The following Proposition defines the minimum sampling frequency of the fading $f_s = \Delta_s^{-1}$ to guarantee the target BER $\bar{P}_b$ given the fading correlation $\mathcal{R}(\cdot)$.

Proposition 1: Detection of uncoded bit stream from its received symbols with probability at least $1 - \bar{P}_b$ can be achieved if the sampling *frequency* of fading $f_s = \Delta_s^{-1}$ is such that

$$2\mathcal{R}^2(1/2f_s)\left[1-\mu_p^{-1}\Lambda^{-1}(f_s)\right] \geq \\ \geq \left(1-4\bar{P}_b+\mu^{-1}\right)\Lambda(f_s) \quad (9)$$

where $\Lambda(f_s) = \left[1+\mathcal{R}(f_s^{-1})\right]$, equality holds for $\max_\alpha \sigma_e^2(\alpha f_s^{-1}|\mathcal{R}) \ll 1$ and asymptotically large SNR. The *quality* of the sampling shall be at least $\beta_p = \mu_p/\mu > \left(8\mu\bar{P}_b\right)^{-1}$.

Proof: The conditional BER for uncoded BPSK transmission over symbol k can be approximated as done in [11] by $\Pr(E|h(k),\sigma_e^2(k|\mathcal{R})) \approx \mathcal{Q}\left(\sqrt{2\cdot SNR_k}\right)$ with SNR at the decision variable (from model (8)) $SNR_k \approx \mu\cdot|h(k)|^2\left(1+\mu\cdot\sigma_e^2(k|\mathcal{R})\right)^{-1}$ and MSE $\sigma_e^2(k|\mathcal{R}$ in (7) with $k = \alpha\Delta_s$. The approximation is tight for large SNR and small enough Doppler so that *i)* all the impairments from channel estimation errors are approximated as white Gaussian noise; *ii)* the conditional error probabilities can be taken from conventional $Q()$ function; *iii)* bias of channel power estimate can be neglected $\mathbb{E}[|\hat{h}(k)|^2] = \left(1-\sigma_e^2(k|\mathcal{R})\right) \approx \mathbb{E}[\,|h(k))|^2]$ as the most significant source of error is the MSE of channel estimation that adds to AWG noise term [14]. At large SNR the average BER $P_e = \mathbb{E}_{h(k)}[P(E|h(k))]$ scales as [12][2]

$$P_e(\sigma_e^2(k|\mathcal{R})) \approx (4\mu)^{-1}\left(1+\mu\times\sigma_e^2(k|\mathcal{R})\right), \quad (10)$$

tightness of (10) for large SNR and small enough Doppler is also numerically proved in Fig. 2-(a) for AR-1 model of fading. Condition $\beta_p > \left(4\mu\bar{P}_b\right)^{-1}$ is necessary as it constrains the channel estimation error over pilots $\mathbb{E}[\,|\Delta h_p|^2] = \mu_p^{-1}$ so that the average BER (10) is below $\bar{P}_b$ in case of block fading ($\sigma_e^2(k|\mathcal{R})) \equiv (2\mu_p)^{-1}$). For $\max_\alpha \sigma_e^2(\alpha f_s^{-1}|\mathcal{R}) \ll 1$ the largest MSE happens at the center of the block[3] ($\alpha = 1/2$) with

$$\max_\alpha \sigma_e^2(\alpha f_s^{-1}|\mathcal{R}) \approx \\ \approx 1 - 2\mathcal{R}^2(f_s^{-1}/2)\frac{\left(1-\mu_p^{-1}\Lambda^{-1}(f_s)\right)}{\Lambda(f_s)}. \quad (11)$$

Condition (9) follows as $P_e \leq P_e(\max_\alpha \sigma_e^2) \leq \bar{P}_b$, thus $\max_\alpha \sigma_e^2(\alpha f_s^{-1}|\mathcal{R}) \leq 4\bar{P}_b - \mu^{-1}$. ■

Corollary 1: For Gauss-Markov fading dynamics, the condition (9) becomes

$$f_s \geq f_{\min}^{(u)} = \frac{\varepsilon}{8\bar{P}_b}\left(1+\frac{1}{4\mu\bar{P}_b}\left[1+\frac{1}{2\beta_p}\right]\right), \quad (12)$$

where $\varepsilon = 1-\rho \ll 1$ is the fading decorrelation. In the limit for $\mu \to \infty$ and $\mu_p \geq \mu$ the minimum sampling frequency of fading scales as

$$\lim_{\mu\to\infty} f_{\min}^{(u)} = \varepsilon \times \bar{P}_b^{-1}/8 \quad (13)$$

Proof: For AR-1 with $\gamma = \rho^2$, the MSE at the center of the block reduces, for $\gamma \approx 1 - 2\Delta_s\varepsilon$, to $\sigma_e^2(\alpha = 1/2|\mathcal{R}) \approx (\varepsilon/2f_s)\left[1-\mu_p^{-1}\right] + (2\mu_p)^{-1}$. ■

B. Bit Interleaved coded BPSK

The information bit stream is now encoded (e.g., by using a convolutional encoder) into a number of codewords. The code rate R_c constrains the symbol energy to E_sR_c so that the energy of *information* symbol is the same for both coded and uncoded systems. Performances are ruled at high SNR by the error event of minimum length d_f. Codewords are then fed back into an ideal interleaver that converts the channel and the estimation noise into memoryless processes thus constraining the elements in the received sequence (after de-interleaving) to be *all* independent.

Proposition 2: Detection of interleaved and coded bit stream from its received symbols with probability at least $1-\bar{P}_b$ can

[2]Similar BER expressions can be derived for Gray-coded M-PSK with AR-1 model of fading [7].

[3]This holds true for monotonic correlation functions $\mathcal{R}(\cdot)$ within the range of interest.

be achieved if the sampling *frequency* of fading $f_s = \Delta_s^{-1}$ is such that

$$\int_0^1 \sigma_e^2(\alpha f_s^{-1}|\mathcal{R})d\alpha \leq \frac{4R_c}{\psi(\bar{P}_b, d_f)} - \frac{1}{\mu} \quad (14)$$

with $\psi(\bar{P}_b, d_f) = (w_f \binom{2d_f-1}{d_f}/\bar{P}_b)^{1/d_f}$ where equality holds for $\max_\alpha \sigma_e^2(\alpha f_s^{-1}|\mathcal{R}) \ll 1$ and asymptotically large SNR. The *quality* of the sampling shall be at least $\beta_p = \mu_p/\mu > \psi(\bar{P}_b, d_f)(4R_c\mu)^{-1}$.

Proof: The conditional BER using MLSE decoding at the receiver is ruled at large SNR by the probability of an error event of minimum length d_f [12]

$$\Pr(E|\mathbf{h}, \sigma^2) \approx w_f \times Q(\sqrt{2R_c \times \sum\nolimits_{k\in\xi(d_f)} SNR_k}) \quad (15)$$

with vectors $\mathbf{h} = \{h(k)\}_{k\in\xi(d_f)}$, $\sigma^2 = \left\{\sigma_e^2(k|\mathcal{R})\right\}_{k\in\xi(d_f)}$and $\xi(d_f)$ refers to the set of d_f codeword bit positions that cause the error event with distance d_f. w_f is the number of error bits in the error event. The average BER $P_e\left(\sigma_e^2\right) = \mathbb{E}_{\mathbf{h},\sigma^2}[\Pr(E|\mathbf{h}, \sigma^2)]$ over the fading process $\mathbf{h}$ [12] *and* the set of estimation errors σ^2 scales, for large SNR and small fading dynamics as

$$P_e\left(\sigma_e^2\right) \approx$$
$$\approx w_f \binom{2d_f-1}{d_f} \left(1 + \mu \int_0^1 \sigma_e^2(\alpha f_s^{-1}|\mathcal{R})d\alpha\right)^{d_f} (4R_c\mu)^{-d_f}. \quad (16)$$

The (ideal) interleaving can alleviate the impairments that are caused by channel out-dating by associating independent and uncorrelated estimation errors to the bits that belong to the same error event (for the block fading case, the reader might also refer to [13]). It follows (14) by constraining $P_e\left(\sigma_e^2\right) \leq \bar{P}_b$. Notice that $\int_0^1 \sigma_e^2(\alpha/f_s|\mathcal{R})d\alpha$ can be written as a function of $\int_0^1 \mathcal{R}^2(\alpha/f_s)d\alpha$ and $\int_0^{1/2} \mathcal{R}(\alpha/f_s)\mathcal{R}((1-\alpha)/f_s)d\alpha$, while last two integrals can be found by straightforward numerical integration. ∎

Corollary 2: For Gauss-Markov fading dynamics condition (14) becomes

$$f_s \geq f_{\min}^{(c)} = \varepsilon \frac{\psi(\bar{P}_b, d_f)}{16R_c}\left(1 + \frac{\psi(\bar{P}_b, d_f)}{\mu R_c}\left[1 + \frac{3}{8\beta_p}\right]\right) \quad (17)$$

for $\varepsilon \ll 1$. In the limit for $\mu \to \infty$ and for $\mu_p \geq \mu$ the minimum frequency of fading sampling scales as

$$\lim_{\mu\to\infty} f_{\min}^{(c)} = \frac{\varepsilon}{16R_c\bar{P}_b^{1/d_f}} \times \left[w_f\binom{2d_f-1}{d_f}\right]^{1/d_f} \quad (18)$$

Proof: For AR-1 model of fading $\int_0^1 \sigma_e^2(\alpha f_s^{-1}|\mathcal{R})d\alpha \approx (\varepsilon/4f_s)\left[1-\mu_p^{-1}\right] + (4\mu_p/3)^{-1}$. ∎

A numerical evaluation of the minimum sampling frequencies for uncoded (Corollary 1, (13)) and bit interleaved coded (Corollary 2, (18)) BPSK is shown in Fig. 3 for AR-1 fading dynamics, $\bar{P}_b = 5 \times 10^{-2}$ (uncoded case) and $\bar{P}_b = 10^{-2}$ (coded case). Solid lines are obtained by simulating the BER for varying pilot spacings and show the required fading sampling frequencies for varying SNRs μ and for $\beta_p = \mu_p/\mu = 1$. Convolutional code has $R_c = 1/2$ and $d_f = 5$. Asymptotic bounds $\lim_{\mu\to\infty} f_{\min}^{(c)}$ and $\lim_{\mu\to\infty} f_{\min}^{(u)}$ (in dashed lines) reveal as effective in estimating the minimum sampling interval when $\mu > 20dB$ and $\varepsilon < 0.035$.

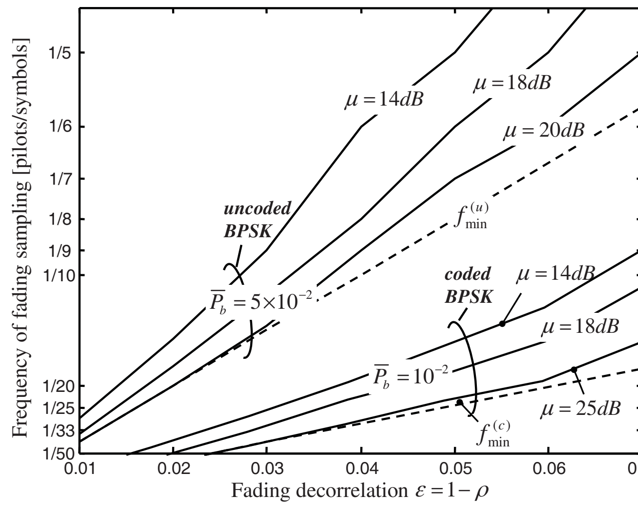

Fig. 3. Minimum sampling frequencies for asymptotically high SNR (dashed lines) for uncoded (Corollary 1) and bit interleaved convolutional coded (from Corollary 2 with $R_c = 1/2$, $d_f = 5$ and $w_{d_f} = 1$) BPSK for AR-1 model of fading ($\bar{P}_b = 10^{-2}$). Simulated curves (solid lines) shows the required fading sampling frequencies for varying SNRs and for $\beta_p = 1$. Convolutional code has $R_c = 1/2$ and generators (in octal) $[5, 7]$. Log-likelihood ratios (LLRs) for Viterbi decoding are computed by assuming the noisy terms as independent Gaussian RVs.

IV. Concluding Remarks

The minimum sampling frequency of the fading (or pilots spacing) is evaluated here to guarantee an average BER at the receiver for bit interleaved coded and uncoded binary systems. MMSE channel interpolation serves as an effective (and low complexity) method to acquire the channel estimate in time-varying fading by combining the estimates from the previous and the upcoming pilot symbols. The minimum sampling frequency of fading is shown to depend in general on the (arbitrary) fading correlation function $\mathcal{R}(\cdot)$, the SNRs for pilot (μ_p) and information (μ) symbols and the required performances ($\bar{P}_b$). The large fading oversampling that is required for uncoded transmission can be traded for longer codewords with smaller rates (and symbol energy) to improve the robustness against estimation errors (and noise). The joint exploitation of (convolutional) coding and (ideal) interleaving is shown to reduce the impact of errors due to channel out-dating by associating independent estimation errors to each symbol belonging to the same error event. The conditions on the required minimum sampling frequencies hold for generic code design and are shown to be valid for high SNR and small fading dynamics.

References

[1] B. Hassibi and B. M. Hochwald, "How much training is needed in multiple-antenna wireless links?" *IEEE Trans. Inform. Theroy*, vol. 49, no. 4, pp. 951-963, Apr. 2003.

[2] V. Pohl, P. H. Nguyen, V. Jungnickel, and C von Helmolt, "Continuous flat-fading MIMO channels: achievable rate and optimal length of the training and data phases," *IEEE Trans. Wireless Commun.*, vol. 4, no 4, pp. 1889-1900, July 2005.

[3] S. Akin and M. C. Gursoy, "Training optimization for Gaussian-Markov Rayleigh fading channels," in *Proc. IEEE Int. Conf. on Commun.*, pp. 5999-6004, Glasgow, June 2007.

[4] S. Savazzi and U. Spagnolini, "Optimizing training lengths and training intervals in continuous time-varying fading channels," *IEEE Trans. Signal Processing*, vol. 57, no. 3, Mar. 2009.

[5] M. Dong, L. Tong, and B. M. Sadler, "Optimal insertion of pilot symbols for transmissions over time-varying flat fading channels," *IEEE Trans. Signal Processing*, vol. 52, no. 5, pp. 1403-1418, May 2004.

[6] A. Lonzano, "Interplay of spectral efficiency, power and Doppler spectrum for reference-signal-assisted wireless communication," *IEEE Trans. Wireless Commun.*, vol. 7, no. 12, pp. 5020-5029, Dec. 2008.

[7] S. Misra, A. Swami, and T. Moore, "Bit-error rate optimized training for time-varying fading channels," in *Proc. IEEE 6th Workshop on Signal Proc. Advances in Wireless Commun.*, pp. 358-362, June 2005.

[8] S. Misra, A. Swami, and L. Tong, "Optimal training over the Gauss-Markov fading channel: a cutoff rate analysis," in *Proc. IEEE Int. Conf. on Acoustics, Speech, and Signal Proc.*, vol. 3, pp. 809-812, May 2004.

[9] K. E. Baddour and N. C. Beaulieu, "Autoregressive modeling for fading channel simulation," *IEEE Trans. Wireless Commun.*, vol. 4, no. 4, pp. 1650, July 2005.

[10] A. Bdeir, I. Abou-Faycal, and M. Medard, "Power allocation schemes for pilot symbol assisted modulation over Rayleigh fading channels with no feedback," in *Proc. IEEE Int. Conf. on Commun.*, vol. 2, pp. 737-741, June 2004.

[11] M. A. R. Baissas and A. M. Sayeed, "Pilot-based estimation of time-varying multipath channels for coherent CDMA receivers," *IEEE Trans. Signal Processing*, vol. 50, no. 8, pp 2037-2049, Aug. 2002.

[12] J. Proakis, *Digital Communications*, 4th ed. McGraw Hill, 2001.

[13] A. Gorokhov, "On the performance of the Viterbi decoder with trained and semi-blind channel estimators," in *Proc. IEEE Int. Conf. on Acoustics, Speech, and Signal Proc.*, vol. 3, pp. 2631-2634, May 1999.

[14] J. Baltersee, G. Fock, and H. Meyr, "Achievable rate of MIMO channels with data-aided channel estimation and perfect interleaving," *IEEE J. Select. Areas Commun.*, vol. 19, no. 12, pp. 2358-2368, Dec. 2001.