16th European Signal Processing Conference (EUSIPCO 2008), Lausanne, Switzerland, August 25-29, 2008, copyright by EURASIP

A ROBUST CHANNEL ESTIMATOR AT THE HIGH DOPPLER FREQUENCY VIA MATCHING PURSUIT TECHNIQUE

Lei Chen and Bernard Mulgrew

Institute for Digital Communications, School of Engineerings and Electronics, The University of Edinburgh
Edinburgh, EH9 3JL, UK
{L.Chen, B.Mulgrew}@ed.ac.uk

ABSTRACT

In this paper, we propose a robust channel estimator for fast time-varying channels. For this technique, a basis expansion model (BEM) is used in order to approximate the time-variant channel. Pilot symbol assisted modulation (PSAM) is employed as the training scheme, and the coefficients of the BEM are obtained via the matching pursuit (MP) algorithm. Through computer simulations, it is demonstrated that MP is able to offer a reliable channel estimate even when it selects an inappropriate number of basis functions, whilst it was previously reported that more basis functions than necessary could cause performance degradation by enhancing the noise. Thus, MP is robust with respect to mismatch in the number of basis functions in the BEM. Furthermore, the simulation results show that MP can operate well even without knowledge of the Signal-to-Noise Ratio (SNR).

1. INTRODUCTION

In wireless communication systems, the high mobility of terminals or the carrier frequency offset between transmitter and receiver gives rise to a fast time-varying channel. For this channel having a high Doppler frequency, a basis expansion model (BEM) [1] has been widely used to approximate it. The coefficients of the BEM could be obtained by the least squares (LS) algorithm when pilot symbol assisted modulation (PSAM) [2] is employed for the time-multiplexed training [3]. However, such a channel characterization scheme has been reported to be sensitive to noise, which, when it is high, results in performance degradation [4].

Both Zemen *et al.* and Tang *et al.* have realized this problem, and suggested their solutions individually. The first produced the optimal number of basis functions in the BEM, i.e. the optimum model order, via mathematical analysis [5], whilst the latter proposed the optimal training pattern, which is the appropriate pilot positions of PSAM in the time domain [6]. But, both solutions have their own problems when applied in practical situations.

Concerning the first solution, the optimal model order highly depends on the instant Signal-to-Noise Ratio (SNR) of a single transmission block [5]. For the fast time-varying channel, the instantaneous value of the SNR for a block is hard to accurately estimate, and shifts considerably from block to block. Hence, it is difficult to exactly determine the proper number of basis functions for any specific block. As a consequence, the slight mismatch of the number of basis functions would cause a performance degradation of the channel estimation. With respect to the second solution, the pilots in the optimal training pattern are no longer uniformly distributed and equispaced [6]. For PSAM, however, the periodic and uniform placement of the pilots is the preferred training pattern as stated in [7].

In this paper, we propose a channel estimator which is robust to mismatches in the number of basis functions in the BEM for fast time-varying channels. In other words, the technique is non-sensitive to noise when employing more basis functions than necessary at a high Doppler frequency, say for instance the normalized Doppler frequency to be $f_d T_s = 0.05$. Moreover, the proposed channel estimation scheme can still provide a reliable estimate without knowledge of the SNR. This scheme consists of the channel model and the estimation algorithm. The model exploits the concept of BEM, and the basis is discrete Karhunen-Loève BEM (DKL-BEM) [8, 9], which is the discrete Karhunen-Loève decomposition of the bathtub-shaped Doppler power spectrum, i.e. Jakes Doppler spectrum. The matching pursuit (MP) algorithm [10] is used to produce the coefficients of the basis functions. Due to the orthogonality of the DKL-BEM, MP has good convergence performance and does not need to re-select the basis functions as in [11]. For PSAM, the pilots are uniformly placed in the time domain.

The MP technique has been applied to produce the coefficients of the basis functions in [11]. However, its channel estimator is quite different from ours. Firstly, the previous study used the polynomial basis, which has been shown to be outperformed by DKL-BEM at high Doppler frequencies [8], while the proposed scheme employs DKL-BEM. Secondly, its pilot placement, which clusters at the center and each end of the block, is entirely different from ours, which arranges the pilots evenly throughout the block. Thirdly, its normalized Doppler frequency, which is $f_d T_s = 0.0008$, is very low compared with the value used here.

The DKL-BEM is optimal for the wireless channel of a bathtub-shaped Doppler power spectrum, and the discrete prolate spheroidal BEM (DPS-BEM) is optimal for that of a rectangular power spectrum [9]. Hence, the proposed channel estimator could be simply extended to the DPS-BEM case. Furthermore, this technique based on a flat-fading channel could be easily extended to the orthogonal frequency division multiplexing (OFDM) case, since OFDM can be viewed as several parallel flat-fading channels in the frequency domain.

2. PROBLEM FORMULATION

We consider a baseband-equivalent discrete-time representation of a wireless communication system over a flat-fading channel with a single transmit and receive antenna. The received signal sample $y(n)$ at time n could be formulated as

$$y(n) = s(n)h(n) + \eta(n) \tag{1}$$

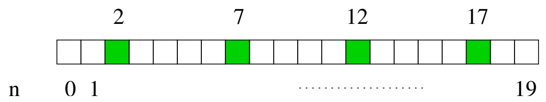

Figure 1: Example of the training pattern $\mathsf{P} = \{2,7,12,17\}$ defined by (2) for $N = 20$ and $J = 4$.

where $s(n)$, $h(n)$, and $\eta(n)$ denotes the transmitted symbol, the channel impulse response, and the complex zero-mean Gaussian noise with variance σ_η^2, respectively, at time instance n. Let us also define f_d as the maximum Doppler frequency of the channel, T_s as the symbol duration, N as the number of the symbols in a transmission block, J as the number of pilots within a block. Thus, there are $N-J$ data symbols in a block.

In this study, the pilots of the PSAM are uniformly distributed throughout the transmission block, which is specified by the index set

$$\mathsf{P} = \left\{ \left\lfloor i\frac{N}{J} + \frac{N}{2J} \right\rfloor \mid i \in \{0,\ldots,J-1\} \right\} \tag{2}$$

Fig. 1 shows an example for the pilot placement defined in (2).

2.1 Basis Expansion Model

Recently, the BEM technique has been widely used to approximate the time-varying channel since it is able to translate a time-varying sequence into a few constant coefficients with the aid of basis functions.

$$h_\infty(n) = \sum_{d=0}^{\infty} c_d\, q_d(n) \tag{3}$$

where $h_\infty(n)$ is the time-variant sequence; c_d and $q_d(n)$ represents the d-th constant coefficient and basis function at time n, respectively.

By truncating $h_\infty(n)$ into $h(n) = \sum_{d=0}^{D-1} c_d\, q_d(n)$ and inserting it into (1):

$$\begin{aligned} y(n) &= s(n)\left\{\sum_{d=0}^{D-1} c_d\, q_d(n)\right\} + \eta(n) \\ &= \sum_{d=0}^{D-1} s(n)\, c_d\, q_d(n) + \eta(n) \end{aligned} \tag{4}$$

Hence, the transmission over the entire block can be represented in the form of the matrix as

$$\begin{aligned} \mathbf{y} &= \mathbf{S}\,\mathbf{h} + \boldsymbol{\eta} \\ &= \mathbf{S}\,\mathbf{Q}\,\mathbf{c} + \boldsymbol{\eta} \end{aligned} \tag{5}$$

where $\mathbf{y} = [y(0), y(1), \cdots, y(N-1)]^T$ is the $N \times 1$ vector consisting of the received symbols; $\mathbf{S}$ is the $N \times N$ diagonal matrix with $[s(0), s(1), \cdots, s(N-1)]$ as its diagonal; $\mathbf{h} = [h(0), h(1), \cdots, h(N-1)]^T$ is the vector of the channel impulse response; $\boldsymbol{\eta}$ is the $N \times 1$ vector of the noise; $\mathbf{Q}$ is the $N \times D$ matrix of the basis functions, with $\mathbf{Q}$ being:

$$\mathbf{Q} = \begin{bmatrix} q_0(0) & q_1(0) & \cdots & q_{D-1}(0) \\ q_0(1) & q_1(1) & \cdots & q_{D-1}(1) \\ \vdots & \vdots & \ddots & \vdots \\ q_0(N-1) & q_1(N-1) & \cdots & q_{D-1}(N-1) \end{bmatrix}$$

and $\mathbf{c} = [c_0, c_1, \cdots, c_{D-1}]^T$ is the $D \times 1$ vector of the constant coefficients in the BEM.

In this study, the basis function in $\mathbf{Q}$ is the DKL-BEM [9], which is the discrete Karhunen-Loève (DKL) decomposition of the bathtub-shaped Doppler power spectrum. The total number of necessary basis functions is determined by $D = \lceil 2N f_d T_s \rceil + 1$, and doubling it would be beneficial [3].

With the help of the BEM, the channel impulse response has been expanded into a sum of basis functions, i. e. $\mathbf{h} = \mathbf{Q}\,\mathbf{c}$. And (5) can be further developed to

$$\mathbf{y} = \mathbf{G}\,\mathbf{c} + \boldsymbol{\eta} \tag{6}$$

where $\mathbf{G}$ is the product of the transmitted symbols and the basis functions, defined as $\mathbf{G} = \mathbf{S}\,\mathbf{Q}$.

At this point, the channel characterization problem has been converted to the identification of the coefficients in BEM, i.e. the estimation of $\mathbf{c}$, if $\mathbf{y}$ and $\mathbf{G}$ are known. In order to obtain parameters $\mathbf{c}$, one of the possible solutions is to use PSAM.

2.2 Subsampled at the Pilot Position

According to PSAM, the transmitted pilot symbols are perfectly known at the receiver. So, all the matrices in (5) and (6) can be subsampled at the pilot positions.

$$\begin{aligned} \mathbf{y}_{\mathsf{p}} &= \mathbf{S}_{\mathsf{p}}\,\mathbf{h}_{\mathsf{p}} + \boldsymbol{\eta}_{\mathsf{p}} \\ &= \mathbf{S}_{\mathsf{p}}\,\mathbf{Q}_{\mathsf{p}}\,\mathbf{c} + \boldsymbol{\eta}_{\mathsf{p}} \\ &= \mathbf{G}_{\mathsf{p}}\,\mathbf{c} + \boldsymbol{\eta}_{\mathsf{p}} \end{aligned} \tag{7}$$

where $\mathbf{y}_{\mathsf{p}} = [y(\mathsf{p}_0), y(\mathsf{p}_1), \cdots, y(\mathsf{p}_{J-1})]^T$ is the $J \times 1$ vector collecting the received symbols at the pilot position; $\mathbf{S}$ is the $J \times J$ diagonal matrix with $[s(\mathsf{p}_0), s(\mathsf{p}_1), \cdots, s(\mathsf{p}_{J-1})]$ as its diagonal; $\mathbf{h} = [h(\mathsf{p}_0), h(\mathsf{p}_1), \cdots, h(\mathsf{p}_{J-1})]^T$ is the $J \times 1$ vector of the channel impulse response at the pilot structure; $\boldsymbol{\eta}_{\mathsf{p}}$, $\mathbf{Q}_{\mathsf{p}}$, and $\mathbf{G}_{\mathsf{p}}$ are the subsampled version of $\boldsymbol{\eta}$, $\mathbf{Q}$, and$\mathbf{G}$, respectively, according to the training pattern.

It is interesting to notice that the coefficient vector $\mathbf{c}$ is not affected by the subsampling of the transmission block. At this moment, if the constant parameters $\mathbf{c}$ could be accurately determined by some method, the entire channel state information $\mathbf{h}$ would be reproduced successfully. In contrast to the work in [4,5,9], where the LS algorithm is used for the estimation of $\mathbf{c}$, here the MP algorithm is adopted. Finally, the estimated coefficients $\hat{\mathbf{c}}$ along with the basis functions serve to recover the channel impulse response $\hat{\mathbf{h}}$ by

$$\hat{\mathbf{h}} = \mathbf{Q}\,\hat{\mathbf{c}} \tag{8}$$

Since the coefficient vector $\mathbf{c}$ might be sparse having some very small quantities, it is not necessary to estimate all the entries accurately. In contrast, to identify and estimate the significant elements, which provide the predominant contribution to the channel state information, in the vector $\mathbf{c}$, would be a reasonable idea.

3. MATCHING PURSUIT

The basic MP algorithm was first proposed in [10]. Its goal is to approximate the observation signal via a linear combination of the elements over a large redundant dictionary. These elements are selected one after another according to which one best matches the signal structure at each iteration.

$$\mathbf{y}_\mathsf{p} \approx \mathbf{G}_\mathsf{p}\,\hat{\mathbf{c}} \tag{9}$$

where $\mathbf{G}_\mathsf{p} = [\mathbf{g}_\mathsf{p}[0], \mathbf{g}_\mathsf{p}[1], \cdots, \mathbf{g}_\mathsf{p}[D-1]]$ is the dictionary, and $\mathbf{g}_\mathsf{p}[i] = [s(\mathsf{p}_0)q_i(\mathsf{p}_0), s(\mathsf{p}_1)q_i(\mathsf{p}_1), \cdots, s(\mathsf{p}_{J-1})q_i(\mathsf{p}_{J-1})]^T$ is the element.

Let the residual of the observation vector after the m-th iteration be represented by ${\mathbf{y}_\mathsf{p}}^{(m)}$, with ${\mathbf{y}_\mathsf{p}}^{(0)} = \mathbf{y}_\mathsf{p}$. The indices of the m selected vectors are stored in the index set $\mathsf{K}_m = [k_1, k_2, \ldots, k_m]$.

At the m-th iteration, the MP algorithm selects k_m by finding the candidate vector which is best aligned with the residual ${\mathbf{y}_\mathsf{p}}^{(m-1)}$, as well as having the longest projection of ${\mathbf{y}_\mathsf{p}}^{(m-1)}$, that is

$$k_m = \arg\max_{l} \frac{\left|\left\langle {\mathbf{y}_\mathsf{p}}^{(m-1)}, \mathbf{g}_\mathsf{p}[l]\right\rangle\right|^2}{\left\|\mathbf{g}_\mathsf{p}[l]\right\|^2}, \quad \text{where } l \notin \mathsf{K}_{m-1} \tag{10}$$

Consequently, the corresponding coefficient is produced by

$$\hat{c}_{k_m} = \frac{\left\langle {\mathbf{y}_\mathsf{p}}^{(m-1)}, \mathbf{g}_\mathsf{p}[k_m]\right\rangle}{\left\|\mathbf{g}_\mathsf{p}[k_m]\right\|^2} \tag{11}$$

In our case, the MP algorithm has been modified in order to guarantee that the picked elements are never re-selected and that the produced coefficients are never re-estimated. This is because the elements are approximately orthogonal, and the modification can speed up the convergence and avoid the re-selection problem. In order to do so, the index of the selected vectors, after being stored in K_m, are removed from the redundant dictionary. As a result, at each iteration a new vector is picked up. A similar modification has been adopted in [11].

The residual vector ${\mathbf{y}_\mathsf{p}}^{(m)}$ is then updated by

$${\mathbf{y}_\mathsf{p}}^{(m)} = {\mathbf{y}_\mathsf{p}}^{(m-1)} - \hat{c}_{k_m}\,\mathbf{g}_\mathsf{p}[k_m] \tag{12}$$

In our study, the process terminates when m is larger than the given maximum iteration, N_a, i.e. the number of identified basis functions also known as the model order. Furthermore, there is no necessity for the usual stopping criterion associated with MP, because the total number of basis functions can be approximately determined by the Doppler power spectrum [3].

4. SIMULATION STUDY

In this section, the proposed channel estimator is compared with the LS technique in terms of the estimation accuracy, i.e. mean square error (MSE),

$$\mathrm{MSE} = \frac{1}{N}\sum_{n=0}^{N-1}|h(n) - \hat{h}(n)|^2 \tag{13}$$

as well as the bit error rate (BER) when PSAM is employed. In PSAM, the estimated channel is used to scale and rotate the decision boundary in order to detect the data symbols [2]. The comparison is conducted under the condition that the BEM of both MP and LS is given the same number of basis

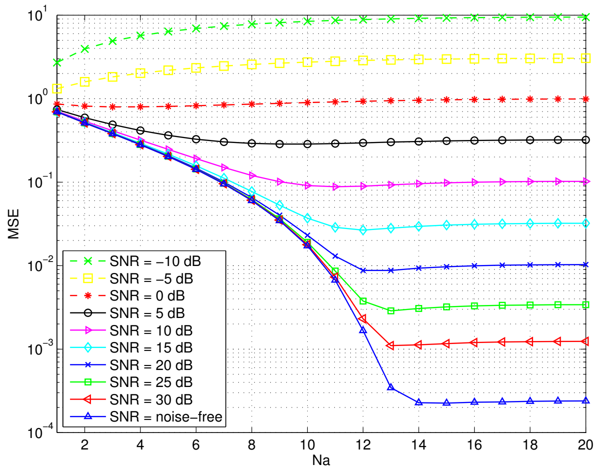

Figure 2: MSE vs N_a under the different SNR, when MP is used.

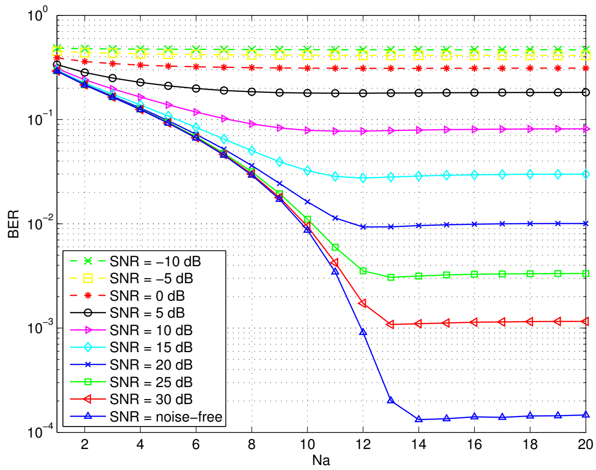

Figure 3: BER vs N_a under the different SNR, when MP is used.

functions, N_a. However, the MP and LS may select individual basis functions differently, because the MP adaptively selects N_a significant basis from the dictionary of D basis functions according to their individual projection contribution, while the LS uses the fixed N_a basis which have the first N_a largest eigenvalues in the DKL-BEM.

The modulation scheme used is quadrature phase shift keying (QPSK), and SNR $= 1/\sigma_\eta^2$. The block consists of 80 data symbols and 20 pilot symbols, which are uniformly placed according to (2). The normalized maximum Doppler frequency of the flat-fading channel is $f_dT_s = 0.05$. As far as the BEM is concerned, the total number of basis functions is $\lceil 2Nf_dT_s\rceil + 1 = 11$, and it would be beneficial to double the total number of the candidate basis so as to obtain a more precise approximation. Hence, $D = 2\lceil 2Nf_dT_s\rceil + 1 = 21$ basis are employed in this paper. In order to prevent (7) from being underdetermined, N_a varies from 1 to 20 since there are only 20 pilots. The stochastic channel for the simulations is generated based on the classical Jakes Doppler spectrum. The simulation results are averaged over 10^6 transmission blocks. Moreover, we assume that the receiver side has the perfect

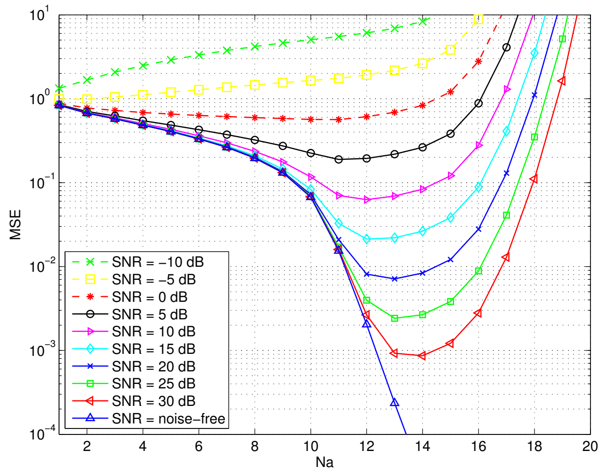

Figure 4: MSE vs N_a under the different SNR, when LS is used.

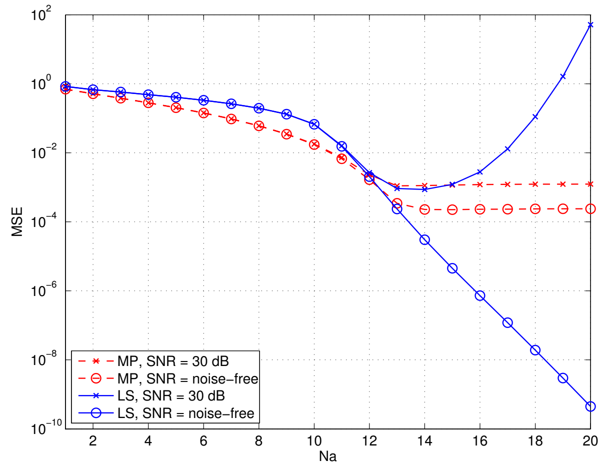

Figure 6: MSE vs N_a. The performance comparison between MP and LS for SNR = 30 dB and noise-free case.

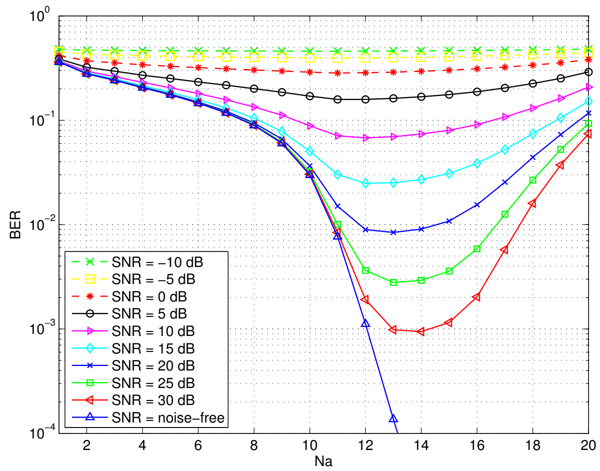

Figure 5: BER vs N_a under the different SNR, when LS is used.

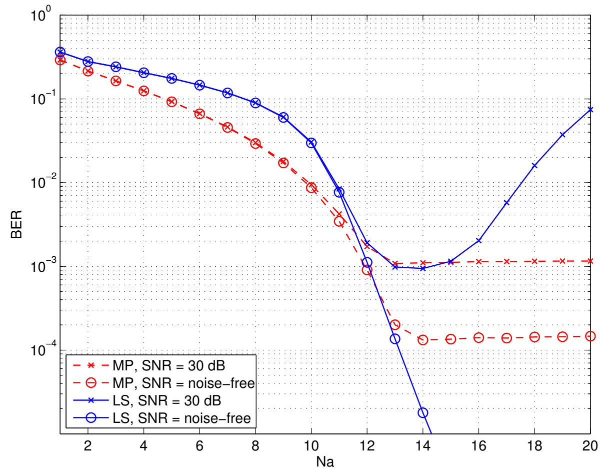

Figure 7: BER vs N_a. The performance comparison between MP and LS for SNR = 30 dB and noise-free case.

knowledge of the actual normalized Doppler frequency.

Fig. 2 shows the performance of MP with respect to the model order, N_a, under different SNRs in terms of MSE. The label "noise-free" means noise is not present, i.e. $\sigma_\eta^2 = 0$. It is observed that, as more basis functions are selected by MP, the performance of the channel estimator improves and then, after a turning point, it levels out. Furthermore, it is noticed that the performance is always convergent as the model order increases, even for different SNRs. The corresponding BER performance has a similar behaviour as it can be seen in Fig. 3. In Fig. 4, the MSE performance of LS against N_a is shown. We observe that in the LS case the performance does not stay at any level even though there exists a turning point as well. Before the turning point the performance becomes better, and afterwards it gets worse. Thus, the performance of the LS technique is affected by the model order, and it is sensitive to the order mismatch. In other words, when the number of basis functions is different from the optimal model order, which is the N_a at the turning point, the performance degrades. The specific value of the optimal model order is dependent on the SNR, as reported in [5]. For example, as seen in Fig. 4, when SNR = 30 dB, the optimal model order is 14, and when SNR = 20 dB, the optimal model order is 13. As shown in Fig. 5 the performance in terms of BER behaves similarly to that in terms of MSE.

To aid comparison, the performance of MP and LS is presented together in a single figure as Fig. 6. For the noise-free case, the LS performance always decreases, while for the MP it goes down initially but then stays constant. In the presence of noise, corresponding to SNR equal to 30 dB, the MP method outperforms LS before the turning point. On the turning point, LS has a more accurate result than MP. However, after the turning point, the performance of LS becomes significantly worse, while that of the MP holds at some level, thus showing its robustness. This phenomenon indicates that MP is robust to mismatches in the number of basis functions, while LS is sensitive to the mismatch. Furthermore, for both noise-free and noisy cases, MP outperforms LS when the model order is smaller than 12. The reason is that MP adaptively selects the predominant basis, whilst LS only uses the fixed basis set. Moreover, the observation that MP under all SNRs can level off means it could still produce reliable es-

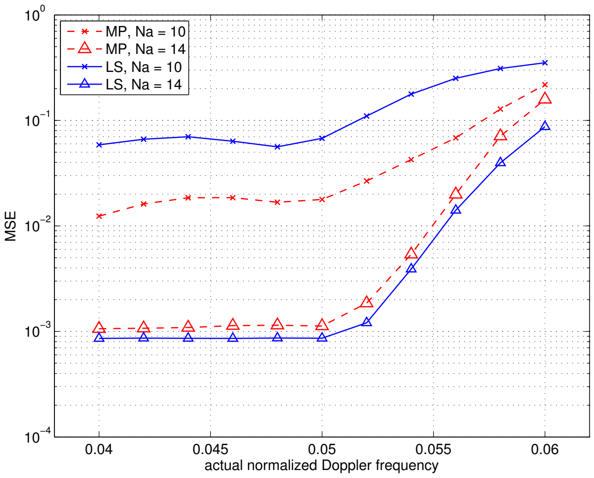

Figure 8: MSE vs actual normalized Doppler frequency. The performance comparison between MP and LS, when the actual Doppler frequency varies within $[0.040, 0.060]$, SNR $= 30$dB. N_a is $\{10, 14\}$, and $N_a = 14$ is optimal for LS under SNR $= 30$dB.

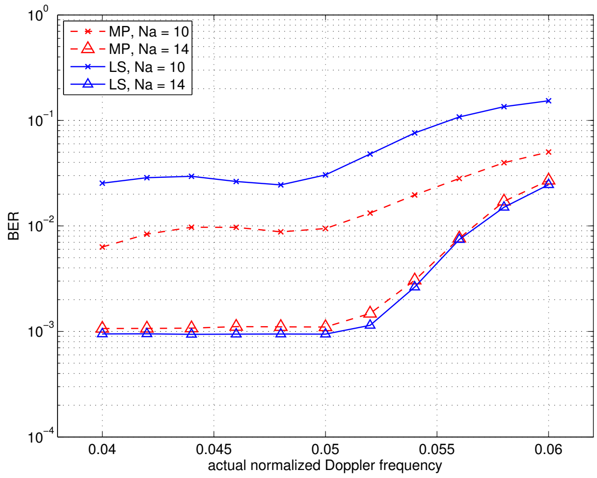

Figure 9: BER vs actual normalized Doppler frequency. The performance comparison between MP and LS, when the actual Doppler frequency varies within $[0.040, 0.060]$, SNR $= 30$dB,. N_a is $\{10, 14\}$, and $N_a = 14$ is optimal for LS under SNR $= 30$dB.

timates without knowledge of the SNR, whilst LS requires the exact knowledge of the SNR in order to perform effectively. In Fig. 7, the BER curves of both MP and LS follow the similar behaviours as the MSE curves.

In the simulations above, the perfect knowledge of the Doppler frequency is assumed. In the real world, however, the Doppler frequency can not be estimated precisely. Still, the actual Doppler frequency could be approximately obtained in a controllable manner. Hence, in this study, we assume a 20% mismatch of the target frequency, which means the actual normalized Doppler frequency varies from 0.04 to 0.06 while the estimator is fixed at 0.05.

Fig. 8 and Fig. 9 presents the performance comparison of MP and LS with respect to the Doppler frequency mismatches in terms of MSE and BER when the model order is $\{10, 14\}$. It is noticed that both MP and LS have similar performance degradations under mismatches in the Doppler frequency. When the actual Doppler frequency is smaller than the one assumed in the BEM, the performance of the channel estimator is close to that without frequency mismatches. When the actual Doppler frequency is larger than the assumption, both of the two suffer from performance degradations.

5. CONCLUSION

A channel estimator has been proposed which is robust to mismatches in the number of basis functions in BEM for fast time-varying channels. In other words, the technique is non-sensitive to noise at high Doppler frequencies, for example, the normalized Doppler frequency of $f_d T_s = 0.05$. Moreover, the proposed channel estimator could provide reliable estimates even without knowledge of the SNR. All of these points have been verified by simulation results.

REFERENCES

[1] M. K. Tsatsanis and G. B. Giannakis, "Modelling and equalization of rapidly fading channels," *Int. J. Adapt. Control Signal Process.*, vol. 10, no. 2/3, pp. 159–176, Mar. 1996.

[2] J. K. Cavers, "An analysis of pilot symbol assisted modulation for rayleigh fading channels," *IEEE Trans. Vehicular Technology*, vol. 40, no. 4, pp. 686–693, Nov. 1991.

[3] T. Zemen and C. F. Mecklenbräuker, "Time-variant channel estimation using discrete prolate spheroidal sequences," *IEEE Trans. Signal Processing*, vol. 53, no. 9, pp. 3597–3607, Sept. 2005.

[4] I. Barhumi, G. Leus, and M. Moonen, "MMSE estimation of basis expansion models for rapidly time-varying channels," in *Proc. EUSIPCO*, Sept. 2005.

[5] T. Zemen, C. F. Mecklenbräuker, F. Kaltenberger, and B. H. Fleury, "Minimum-energy band-limited predictor with dynamic subspace selection for time-variant flat-fading channels," *IEEE Trans. Signal Processing*, vol. 55, no. 9, pp. 4534 – 4548, Sept. 2007.

[6] Z. Tang and G. Leus, "Time-multiplexed training for time-selective channels," *IEEE Signal Processing Letters*, vol. 14, no. 9, pp. 585–588, Sept. 2007.

[7] L. Tong, B. M. Sadler, and M. Dong, "Pilot-assisted wireless transmissions: general model, design criteria, and signal processing," *IEEE Signal Processing Magazine*, vol. 21, no. 6, pp. 12–25, Nov. 2004.

[8] K. A. D. Teo and S. Ohno, "Optimal MMSE finite parameter model for doubly-selective channels," in *Proc. IEEE Global Telecommun. Conf. (GLOBECOM)*, Nov. 2005, vol. 6, pp. 3503–3507.

[9] Z. Tang, R. C. Cannizzaro, G. Leus, and P. Banelli, "Pilot-assisted time-varying channel estimation for OFDM systems," *IEEE Trans. Signal Processing*, vol. 55, no. 5, pp. 2226–2238, May 2007.

[10] S. G. Mallat and Z. Zhang, "Matching pursuits with time-frequency dictionaries," *IEEE Trans. Signal Processing*, vol. 41, pp. 3397–3415, Dec. 1993.

[11] Y. Liu and D. K. Borah, "Estimation of time-varying frequency-selective channels using a matching pursuit technique," in *Proc. Wireless Communications and Networking*, Mar. 2003, vol. 2, pp. 941–946.