

EVD-BASED CHANNEL ESTIMATION IN MULTICELL MULTIUSER MIMO SYSTEMS WITH VERY LARGE ANTENNA ARRAYS

Hien Quoc Ngo Erik G. Larsson

Department of Electrical Engineering (ISY), Linköping University, 581 83 Linköping, Sweden

ABSTRACT

This paper considers multicell multiuser MIMO systems with very large antenna arrays at the base station. We propose an eigenvalue-decomposition-based approach to channel estimation, that estimates the channel blindly from the received data. The approach exploits the asymptotic orthogonality of the channel vectors in very large MIMO systems. We show that the channel to each user can be estimated from the covariance matrix of the received signals, up to a remaining scalar multiplicative ambiguity. A short training sequence is required to resolve this ambiguity. Furthermore, to improve the performance of our approach, we combine it with the iterative least-square with projection (ILSP) algorithm. Numerical results verify the effectiveness of our channel estimation approach.

1. INTRODUCTION

Recently, there has been a great deal of interest in multiuser MIMO (MU-MIMO) systems using very large antenna arrays (a hundred or more antennas). Such systems can provide a remarkable increase in reliability and data rate with simple signal processing [1]. When the number of base station (BS) antennas grows large, the channel vectors between the users and the BS become very long random vectors and under “favorable propagation” conditions, they become pairwise orthogonal. As a consequence, with simple maximum-ratio combining (MRC), assuming that the BS has perfect channel state information (CSI), the interference from the other users can be cancelled without using more time-frequency resources [1]. This dramatically increases the spectral efficiency. Furthermore, by using a very large antenna array at the BS, the transmit power can be drastically reduced [2]. In [2], we showed that, with perfect CSI at the BS, we can reduce the uplink transmit power of each user inversely proportionally to the number of antennas with no reduction in performance. This holds true even with simple linear processing (MRC, or zero-forcing [ZF]) at the base station. These benefits of using large antenna arrays can be reaped if the BS has perfect CSI.

In practice, the BS does not have perfect CSI. Instead, it estimates the channels. The conventional way of doing this is to use uplink pilots. If the channel coherence time is limited, the number of possible orthogonal pilot sequences is limited too and hence, pilot sequences have to be reused in other cells. Therefore, channel estimates obtained in a given cell will be contaminated by pilots transmitted by users in other cells. This causes pilot contamination [3]. As for power efficiency, we showed in [2] that, with CSI estimated from uplink pilots, we can only reduce the uplink transmit power per user inversely proportionally to the square-root of the number of BS

antennas. This is due to the fact that when we reduce the transmit power of each user, channel estimation errors will become significant. We call this effect “noise contamination”. Hence, with channels estimated from pilots, the benefits of using very large antenna arrays are somewhat reduced.

In this paper we investigate whether blind channel estimation techniques could improve the performance of very large MIMO systems. Blind channel estimation techniques have been considered before as a promising approach for increasing the spectral efficiency since they require no or a minimal number of pilot symbols [4]. Generally, blind methods work well when there are unused degrees of freedom in the signal space. This is the case in very large MIMO systems, if the number of users that transmit simultaneously typically is much less than the number of antennas. One particular class of blind methods is based on a subspace partitioning of the received samples. This approach is powerful and can achieve near maximum-likelihood performance when the number of data samples is sufficiently large [5]. This approach requires a particular structure on the transmitted signal or system model, for example that the signals are coded using orthogonal space-time block codes [6, 7]. As shown later, in a system with very large antenna arrays it is possible to apply the subspace estimation technique using eigenvalue decomposition (EVD) on the covariance matrix of the received samples, without requiring any specific structure of the transmitted signals.

The specific contributions of this paper are as follows. We consider multicell MU-MIMO systems where the BS is equipped with a very large antenna array. We propose a simple EVD-based channel estimation scheme for such systems. We show that when the number of BS antennas grows large, CSI can be estimated from the eigenvector of the covariance matrix of the received samples, up to a multiplicative scalar factor ambiguity. By using a short training sequence, this multiplicative factor ambiguity can be resolved. Finally, to improve the performance, we combine our EVD-based channel estimation technique with the iterative least-square with projection (ILSP) algorithm of [8].

2. MULTI-CELL MULTI-USER MIMO MODEL

Consider a multicell MU-MIMO system with L cells. Each cell contains K single-antenna users and one BS equipped with M antennas. The same frequency band is used for all L cells. We consider the uplink transmission where all users from all cells simultaneously transmit their signals to their desired BSs. Then, the $M \times 1$ received vector at the l th BS is given by¹

$$\mathbf{y}_l(n) = \sqrt{p_u} \sum_{i=1}^L \mathbf{G}_{li} \mathbf{x}_i(n) + \mathbf{n}_l(n) \quad (1)$$

¹When reference to n is unimportant, we will omit this index for simplicity.

This work was supported in part by ELLIIT and the Swedish Research Council (VR). E. G. Larsson is a Royal Swedish Academy of Sciences (KVA) Research Fellow supported by a grant from the Knut and Alice Wallenberg Foundation.

where $\sqrt{p_u}\mathbf{x}_i(n)$ is the $K \times 1$ vector of collectively transmitted symbols by the K users in the i th cell (the average power used by each user is p_u); $\mathbf{n}_l(n)$ is $M \times 1$ additive white noise whose elements are Gaussian with zero mean and unit variance; and \mathbf{G}_{li} is the $M \times K$ channel matrix between the l th BS and the K users in the i th cell. The channel matrix \mathbf{G}_{li} models independent fast fading, geometric attenuation, and log-normal shadow fading. Each element $g_{limk} \triangleq [\mathbf{G}_{li}]_{mk}$ is the channel coefficient between the m th antenna of the l th BS and the k th user in the i th cell, and is given by

$$g_{limk} = h_{limk} \sqrt{\beta_{lik}}, \quad m = 1, 2, \dots, M \quad (2)$$

where h_{limk} is the fast fading coefficient from the k th user in the i th cell to the m th antenna of the l th BS. We assume that h_{limk} is a random variable with zero mean and unit variance. Furthermore, $\sqrt{\beta_{lik}}$ represents the geometric attenuation and shadow fading, which are assumed to be independent of the antenna index m and to be constant and known a priori. These assumptions are reasonable since the distance between the user and the BS is much greater than the distance between the BS antennas, and the value of β_{lik} changes very slowly with time.² Then, the channel matrix \mathbf{G}_{li} can be represented as

$$\mathbf{G}_{li} = \mathbf{H}_{li} \mathbf{D}_{li}^{1/2} \quad (3)$$

where \mathbf{H}_{li} is the $M \times K$ matrix of fast fading coefficients between the K users in the i th cell and the l th BS, i.e., $[\mathbf{H}_{li}]_{mk} = h_{limk}$, and \mathbf{D}_{li} is a $K \times K$ diagonal matrix whose diagonal elements are $[\mathbf{D}_{li}]_{kk} = \beta_{lik}$.

3. EVD-BASED CHANNEL ESTIMATION

For multicell MU-MIMO systems with large antenna arrays at the BS, with conventional LS channel estimation using uplink pilots, the system performance is limited by pilot and noise contamination. Pilot contamination is caused by the interference from other cells during the training phase [1, 3]. Noise contamination occurs when the transmit power is small and the channel estimates are dominated by estimation errors [2]. Another inherent drawback of the pilot-based channel estimation is the spectral efficiency loss which results from the bandwidth consumed by training sequences. To reduce these effects, in this section, we propose an EVD-based channel estimation method.

3.1. Mathematical Preliminaries

We first consider the properties of the covariance matrix of the received vector \mathbf{y}_l . From (1) and (3), this covariance matrix is given by

$$\mathbf{R}_y \triangleq \mathbb{E} \left\{ \mathbf{y}_l \mathbf{y}_l^H \right\} = p_u \sum_{i=1}^L \mathbf{H}_{li} \mathbf{D}_{li} \mathbf{H}_{li}^H + \mathbf{I}_M. \quad (4)$$

From the law of large numbers, it follows that when the number of BS antennas is large, if the fast channel coefficients are i.i.d. then the channel vectors between the users and the BS become pairwise orthogonal, i.e.,

$$\frac{1}{M} \mathbf{H}_{li}^H \mathbf{H}_{lj} \rightarrow \delta_{ij} \mathbf{I}_K, \quad \text{as } M \rightarrow \infty \quad (5)$$

²This is true assuming that the base station antennas are located, not distributed. For example, at 3 GHz, a cylindrical array comprising 4 rings of 16 dual polarized antennas elements spaced half a wavelength apart, hence having a total of 128 antennas, occupies only a physical size of 0.3×0.35 meters.

This is a key property of large MIMO systems which facilitates a simple EVD-based channel estimation that does not require any specific structure of the transmitted signals. Multiplying (4) from the right by \mathbf{H}_{ll} , and using (5), we obtain

$$\begin{aligned} \mathbf{R}_y \mathbf{H}_{ll} &\approx M p_u \mathbf{H}_{ll} \mathbf{D}_{ll} + \mathbf{H}_{ll}, \quad \text{as } M \text{ large} \\ &= \mathbf{H}_{ll} (M p_u \mathbf{D}_{ll} + \mathbf{I}_K). \end{aligned} \quad (6)$$

For large M , the columns of \mathbf{H}_{ll} are pairwise orthogonal, and $M p_u \mathbf{D}_{ll} + \mathbf{I}_K$ is a diagonal matrix. Therefore, Equation (6) can be considered as a characteristic equation for the covariance matrix \mathbf{R}_y . As a consequence, the k th column of \mathbf{H}_{ll} is the eigenvector corresponding to the eigenvalue $M p_u \beta_{llk} + 1$ of \mathbf{R}_y .

Remark 1 Since $M p_u \beta_{llk} + 1$, $k = 1, 2, \dots, K$, are distinct and can be known a priori, the ordering of the eigenvectors can be determined. Each column of \mathbf{H}_{ll} can be estimated from a corresponding eigenvector of \mathbf{R}_y up to a scalar multiplicative ambiguity.³ This is due to the fact that if \mathbf{u}_k is an eigenvector of \mathbf{R}_y corresponding to the eigenvalue $M p_u \beta_{llk} + 1$, then $c_k \mathbf{u}_k$ is also an eigenvector corresponding to that eigenvalue, for any $c_k \in \mathbb{C}$.

Let \mathbf{U}_{ll} be the $M \times K$ matrix whose k th column is the eigenvector of \mathbf{R}_y corresponding to the eigenvalue $M p_u \beta_{llk} + 1$. Then, the channel estimate of \mathbf{H}_{ll} can be found via

$$\hat{\mathbf{H}}_{ll} = \mathbf{U}_{ll} \mathbf{\Xi} \quad (7)$$

where $\mathbf{\Xi} \triangleq \text{diag} \{c_1, c_2, \dots, c_K\}$. The multiplicative matrix ambiguity $\mathbf{\Xi}$ can be solved by using a short pilot sequence (see Section 3.2).

3.2. Resolving the Multiplicative Factor Ambiguity

In each cell, a short training sequence of length ν symbols is used for uplink training. We assume that the training sequences of different cells are pairwise orthogonal. Then, the $M \times \nu$ received training matrix at the l th BS is

$$\mathbf{Y}_{t,l} = \sqrt{p_t} \mathbf{H}_{ll} \mathbf{D}_{ll}^{1/2} \mathbf{X}_{t,l} + \mathbf{N}_{t,l} \quad (8)$$

where $\sqrt{p_t} \mathbf{X}_{t,l}$ is the $K \times \nu$ training matrix (p_t is the power used by each user for each training symbol), and $\mathbf{N}_{t,l}$ is the noise matrix. From (7) and (8), the multiplicative matrix $\mathbf{\Xi}$ can be estimated as

$$\hat{\mathbf{\Xi}} = \arg \min_{\mathbf{\Xi} \in \mathbf{\Lambda}} \left\| \mathbf{Y}_{t,l} - \sqrt{p_t} \mathbf{U}_{ll} \mathbf{\Xi} \mathbf{D}_{ll}^{1/2} \mathbf{X}_{t,l} \right\|_F^2 \quad (9)$$

where $\mathbf{\Lambda}$ is a set of $K \times K$ diagonal matrices. Denote by $\bar{\mathbf{y}}_n \triangleq \left[(\mathbf{y}_{t,l}^R(n))^T (\mathbf{y}_{t,l}^I(n))^T \right]^T$, where $\mathbf{y}_{t,l}(n)$ is the n th column of $\mathbf{Y}_{t,l}$, \mathbf{B}^R and \mathbf{B}^I denote the real and imaginary parts of matrix \mathbf{B} ; and

$$\bar{\mathbf{A}}_n \triangleq \begin{bmatrix} \mathbf{A}_n^R & -\mathbf{A}_n^I \\ \mathbf{A}_n^I & \mathbf{A}_n^R \end{bmatrix} \quad (10)$$

where $\mathbf{A}_n \triangleq \sqrt{p_t} \mathbf{U}_{ll} \mathbf{D}_{ll}^{1/2} \bar{\mathbf{X}}_n$, $\bar{\mathbf{X}}_n \triangleq \text{diag}(\mathbf{x}_{t,l}(n))$. Then, we obtain (the proof is omitted due to space constraints)

$$\hat{\mathbf{\Xi}} = \text{diag}(\hat{\boldsymbol{\xi}}) \quad (11)$$

where $\hat{\boldsymbol{\xi}} = [\mathbf{I}_K \ j \mathbf{I}_K] \hat{\boldsymbol{\xi}}$, where

$$\hat{\boldsymbol{\xi}} = \left(\sum_{n=1}^{\nu} \bar{\mathbf{A}}_n^T \bar{\mathbf{A}}_n \right)^{-1} \sum_{n=1}^{\nu} \bar{\mathbf{A}}_n^T \bar{\mathbf{y}}_n. \quad (12)$$

³Note that the channel matrices from other cells \mathbf{H}_{li} , for $i \neq l$, can be estimated in the same way if the large-scale fading \mathbf{D}_{li} is known a priori.

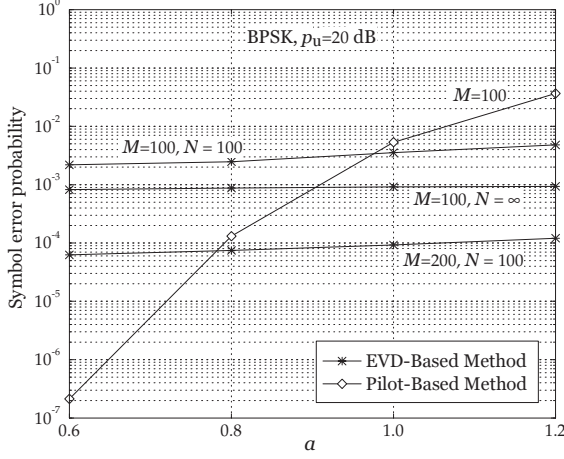


Fig. 1. Symbol error probability versus α at $p_u = 20$ dB, and BPSK modulation.

3.3. Implementation of the EVD-based Channel Estimation

As discussed, when M is large the channel matrix \mathbf{H}_{ll} can be determined by using the EVD of the covariance matrix \mathbf{R}_y . In practice, this covariance matrix is unavailable. Instead, we use the sample data covariance matrix $\hat{\mathbf{R}}_y$:

$$\hat{\mathbf{R}}_y \triangleq \frac{1}{N} \sum_{n=1}^N \mathbf{y}_l(n) \mathbf{y}_l(n)^H \quad (13)$$

where N is the number of samples. Here, we assume that the channel is still constant over at least N samples.

We summarize our proposed algorithm for estimating \mathbf{H}_{ll} as the following algorithm. This channel estimate is performed once for each coherence time period.

Algorithm 1 Proposed EVD-based channel estimation method

1. Using a data block of N samples, compute $\hat{\mathbf{R}}_y$ as (13).
2. Perform the EVD of $\hat{\mathbf{R}}_y$. Then obtain an $M \times K$ matrix \mathbf{U}_N whose k th column is the eigenvector corresponding to the eigenvalue which is closest to $M p_u \beta_{llk} + 1$.⁴
3. Compute the estimate $\hat{\Xi}$ of the multiplicative matrix Ξ from ν pilot symbols using (11).⁵
4. The channel estimate of \mathbf{H}_{ll} is determined as $\tilde{\mathbf{H}}_{ll} = \mathbf{U}_N \hat{\Xi}$.

Treating the above channel estimate as the true channel, we then use a linear detector (e.g., MRC, ZF) to detect the transmitted signals. Since the columns of the channel estimate $\tilde{\mathbf{H}}_{ll}$ are pairwise orthogonal for large M , the performances of MRC and ZF detectors are the same [2].

Remark 2 There are two main sources of errors in the channel estimate: (i) The covariance matrix error: this error is due to the use of the sample covariance matrix instead of the true covariance matrix. This error will decrease as the number of samples N increases (this requires that the coherence time is large); (ii) The error due

⁴Since the eigenvalue is obtained from the sample data covariance matrix, the corresponding eigenvalue is only approximately equal to $M p_u \beta_{llk} + 1$.

⁵When using (11) replace the true covariance matrix by the sample covariance matrix.

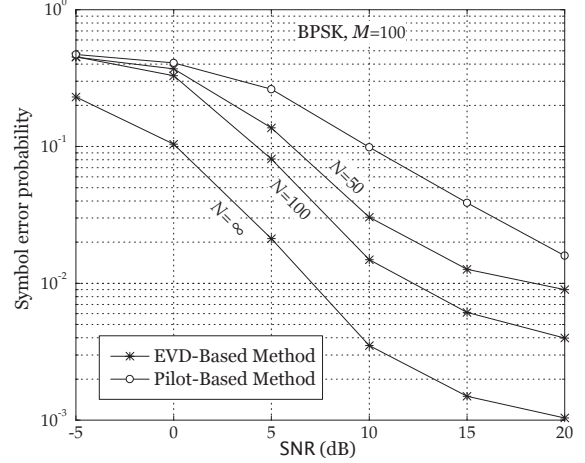


Fig. 2. Symbol error probability versus SNR for $M = 100$, $a = 1$, $p_u = \text{SNR}/M$, and BPSK modulation.

to the channel vectors not being perfectly orthogonal as assumed in (5). Our method exploits the asymptotic orthogonality of the channel vectors. This property is true only in the asymptotic regime, i.e., when $M \rightarrow \infty$. In practice, M is large but finite and hence, an error results.

4. JOINT EVD-BASED METHOD AND ILSP ALGORITHM

As discussed above (see Remark 2), the EVD-based channel estimates will suffer from errors owing to a finite coherence time and a finite M . To reduce this error, in this section, we consider combining our EVD algorithm with the ILSP algorithm of [8].

Define the $K \times N$ matrix of transmitted signals from the K users in the i th cell and the $M \times N$ matrix of received signals at the l th BS respectively as

$$\mathbf{X}_i \triangleq [\mathbf{x}_i(1) \ \mathbf{x}_i(2) \ \dots \ \mathbf{x}_i(N)], \quad i = 1, 2, \dots, L \quad (14)$$

$$\mathbf{Y}_l \triangleq [\mathbf{y}_l(1) \ \mathbf{y}_l(2) \ \dots \ \mathbf{y}_l(N)]. \quad (15)$$

From (1), we have

$$\mathbf{Y}_l = \sqrt{p_u} \mathbf{G}_{ll} \mathbf{X}_i + \sqrt{p_u} \sum_{i \neq l} \mathbf{G}_{li} \mathbf{X}_i + \mathbf{N}_l \quad (16)$$

where $\mathbf{N}_l \triangleq [\mathbf{n}_l(1) \ \mathbf{n}_l(2) \ \dots \ \mathbf{n}_l(N)]$. Treating the last two terms of (16) as noise, and applying the ILSP algorithm in [8], we obtain an iterative algorithm that jointly estimates the channel and the transmitted data. The principle of operation of the ILSP algorithm is as follows. Firstly, we assume that the channel \mathbf{G}_{ll} is known, from an initial channel estimation procedure. The data are then detected via least-squares, projecting the solution onto the symbol constellation \mathcal{X} as

$$\hat{\mathbf{X}}_l = \arg \min_{\mathbf{X}_l \in \mathcal{X}} \left\| \frac{1}{\sqrt{p_u}} \mathbf{G}_{ll}^\dagger \mathbf{Y}_l - \mathbf{X}_l \right\|_F^2 \quad (17)$$

where the superscript $(\cdot)^\dagger$ denotes the pseudo-inverse. Next, the detected data $\hat{\mathbf{X}}_l$ are used as if they were equal to the true transmitted signal and the channel is re-estimated using least-squares,

$$\hat{\mathbf{G}}_{ll} = \frac{1}{\sqrt{p_u}} \mathbf{Y}_l \hat{\mathbf{X}}_l^\dagger. \quad (18)$$

Equations (17) and (18), yield the ILSP algorithm for our problem. Applying the ILSP algorithm, and using the channel estimate based on EVD method discussed in Section 3 as the initial channel estimate, we obtain the joint EVD method and ILSP algorithm (EVD-ILSP).

Algorithm 2 *The EVD-ILSP algorithm*

1. Initialize $\hat{\mathbf{G}}_{ll,0} = \tilde{\mathbf{H}}_{ll}\mathbf{D}_{ll}^{1/2}$ (obtained by using the EVD-based method). Choose number of iterations K_{step} . Set $k = 0$.
2. $k := k + 1$
 - $\hat{\mathbf{X}}_{l,k} = \arg \min_{\mathbf{X}_l \in \mathcal{X}} \left\| \frac{1}{\sqrt{p_u}} \hat{\mathbf{G}}_{ll,k-1}^\dagger \mathbf{Y}_l - \mathbf{X}_l \right\|_F^2$
 - $\hat{\mathbf{G}}_{ll,k} = \frac{1}{\sqrt{p_u}} \mathbf{Y}_l \hat{\mathbf{X}}_{l,k}^\dagger$
3. Repeat 2 until $k = K_{\text{step}}$.

5. NUMERICAL RESULTS

We simulate a system with $L = 3$ cells, each containing 3 users. We consider the SEP for the uplink of the 1st user in 1st cell, assuming BPSK modulation and ZF receivers. We choose $\mathbf{D}_{11} = \text{diag}\{0.98, 0.63, 0.47\}$, $\mathbf{D}_{12} = a \times \text{diag}\{0.36, 0.29, 0.05\}$, and $\mathbf{D}_{13} = a \times \text{diag}\{0.32, 0.14, 0.11\}$. For the EVD-based method, we use $\nu = 1$ (one) training symbol to resolve the multiplicative factor ambiguity. For the pilot-based method, we perform the least-square estimation scheme using 3 symbols for pilots.

Fig. 1 shows the the SEP versus a of the EVD-based and the conventional pilot-based channel estimation methods with different N and M at $p_u = 20$ dB. ($N = \infty$ implies that we have perfect knowledge of the true covariance matrix.) We can see that when a increases (the effect of pilot contamination increases), the system performance degrades dramatically when using the pilot-based method. This is due to the fact that the pilot-based method suffers from pilot contamination. In particular, the EVD-based method is not affected much by the pilot contamination, and it can significantly improve the system performance when the effect of pilot contamination is large. It can be also seen from the figure that the effectiveness of our EVD-based method increases when the number of samples N and the number of BS antennas M increase.

To ascertain the effectiveness of the EVD-based channel estimation method under noise-limited conditions, we consider the SEP when the transmit power of each user is chosen to be proportional to $1/M$. We choose $M = 100$, and $a = 1$. Fig. 2 shows the comparisons between the SEPs versus SNR of the EVD-based method and the pilot-aided method for different N . Here, with each SNR, we set $p_u = \text{SNR}/M$. We can see that by using the EVD-based method, the system performance significantly improves compared with the conventional pilot-based method. When N increases, the sample covariance matrix tends to the true covariance matrix and hence, as we can see from the figure, the SEP decreases.

Fig. 3 shows the SEP of the EVD-based method versus the number of BS antennas at $p_u = 20$ dB and $a = 1$, for different N , with and without using the ILSP algorithm. With the ILSP algorithm, we choose $K_{\text{step}} = 5$. As expected, comparing with the EVD-based method, the joint EVD-based and ILSP algorithm offers a performance improvement. Also here, the system performance improves significantly when M and N increase.

6. CONCLUDING REMARKS

Very large MIMO systems with $M \gg K \gg 1$ offer many unused degrees of freedom. We proposed a channel estimation method

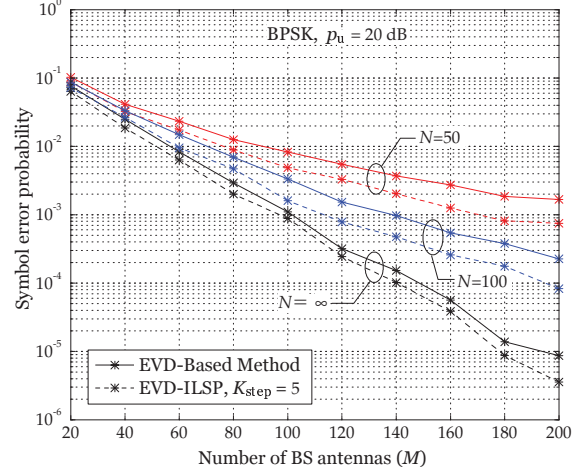


Fig. 3. Symbol error probability versus M for $p_u = 20$ dB, and $a = 1$.

that exploits these excess degrees of freedom, together with the asymptotic orthogonality between the channel vectors that occurs under “favorable propagation” conditions. Combining the proposed method with the ILSP algorithm of [8] further enhances performance.

7. REFERENCES

- [1] T. L. Marzetta, “Noncooperative cellular wireless with unlimited numbers of base station antennas,” *IEEE Trans. Wireless Commun.*, vol. 9, no. 11, pp. 3590–3600, Nov. 2010.
- [2] H. Q. Ngo, E. G. Larsson, and T. L. Marzetta, “Uplink power efficiency of multiuser MIMO with very large antenna arrays,” in *Proc. 49th Allerton Conference on Communication, Control, and Computing*, 2011. [Online]. Available: <http://liu.diva-portal.org/smash/record.jsf?pid=diva2:436517>
- [3] J. Jose, A. Ashikhmin, T. Marzetta, and S. Vishwanath, “Pilot contamination and precoding in multi-cell TDD systems,” *IEEE Trans. Wireless Commun.*, vol. 10, no. 8, pp. 2640–2651, Aug. 2011.
- [4] A.-J. van der Veen, S. Talwar, and A. Paulraj, “Blind estimation of multiple digital signals transmitted over FIR channels,” *IEEE Signal Process. Lett.*, vol. 2, no. 5, pp. 99–102, May 1995.
- [5] A.-J. van der Veen, S. Talwar, and A. Paulraj, “A subspace approach to blind space-time signal processing for wireless communication systems,” *IEEE Signal Process. Lett.*, vol. 45, no. 1, pp. 173–189, Jan. 1997.
- [6] E. Beres and R. Adve, “Blind channel estimation for orthogonal STBC in MISO systems,” *IEEE Trans. Veh. Technol.*, vol. 56, no. 4, pp. 2042–2050, 2007.
- [7] B. Muquet, M. de Courville, and P. Duhamel, “Subspace-based blind and semi-blind channel estimation for OFDM systems,” *IEEE Trans. Signal Process.*, vol. 50, no. 7, pp. 1699–1712, 2002.
- [8] S. Talwar, M. Viberg, and A. Paulraj, “Blind separation of synchronous co-channel digital signals using an antenna array—part I: Algorithms,” *IEEE Trans. Signal Process.*, vol. 44, no. 5, pp. 1184–1197, May 1996.