Performance Analysis of the V-BLAST Algorithm: An Analytical Approach

Sergey Loyka and Francois Gagnon, *Senior Member, IEEE*

***Abstract*—A geometrically based analytical approach to the performance analysis of the V-BLAST algorithm is presented in this paper, which is based on the analytical model of the Gramm–Schmidt process. This approach presents a new geometrical view of the V-BLAST and explains some of its properties in a complete and rigorous form, including a statistical analysis of postprocessing signal-to-noise ratios for a $2 \times n$ system (where n is the number of receive antennas). Closed-form analytical expressions of the vector signal at ith processing step and its power are presented. A rigorous proof that the diversity order at ith step (without optimal ordering) is $(n - m + i)$ is given (where m is the number of transmit antennas). It is shown that the optimal ordering is based on the least correlation criterion and that the after-processing signal power is determined by the channel correlation matrices in a fashion similar to the channel capacity. Closed-form analytical expressions are derived for outage probabilities and average BER of a $2 \times n$ system. The effect of the optimal ordering is shown to be to increase the first step SNR by 3 dB (rather than to increase the diversity order as one might intuitively expect based on the selection combining argument) and to increase the second step outage probability twice.**

***Index Terms*—Bit-error rate (BER), fading, MIMO, multi-antenna system, outage probability, V-BLAST.**

I. Introduction

INFORMATION-THEORETIC considerations show that the multiple-input multiple-output (MIMO) communication architecture is able to provide extraordinary high spectral efficiencies in rich multipath environments, which are simply unattainable using conventional techniques [1]–[4]. Space-time coding and/or a special signal processing algorithm is to be implemented at the receiver in order to achieve at least part of the MIMO channel capacity. Diagonal Bell Labs Layered Space-Time (D-BLAST) algorithm has been proposed by Foschini for this purpose, which is capable of achieving a substantial part of the MIMO capacity [2]. However, a high complexity of the algorithm implementation is its substantial drawback. A simplified version of the BLAST algorithm is known as V-BLAST (vertical BLAST). It is capable of achieving high spectral efficiency while being relatively simple to implement [5].

Comprehensive evaluation of the system performance is required because the matrix wireless propagation channel may severely degrade the performance of this algorithm [6]–[8]. Some preliminary studies including asymptotic analysis and numerical Monte Carlo simulations have been reported in [9]. While the numerical Monte Carlo approach is useful from many viewpoints, the analytical approach provides deeper insight and comprehensive understanding of the key points in the algorithm operation.

In this paper, we develop a unified analytical approach to the analysis of the V-BLAST algorithm operation based on some general geometrical ideas. This approach is based on the closed-form analytical models of the key V-BLAST and associated system components—interference nulling from yet to be detected symbols (Gram–Schmidt orthogonalization process), interference subtraction from already detected symbols, the optimal ordering procedure [based on the after processing signal-to-noise ratio (SNR)], optimal (maximum ratio or similar) combining, and a statistical (complex Gaussian) model of the matrix wireless propagation channel. In particular, we derive closed-form analytical expressions for the signal and noise vectors at each processing step for wireless channel with the general correlation matrix. Based on these results, we give a rigorous proof that the diversity order (i.e., the asymptotic slope of the outage probability curve) at the ith processing step is $n - m + i$, where n and m are the number of Rx and Tx antennas correspondingly, for uncorrelated Rayleigh channel and if no optimal ordering is used. Closed-form analytical expressions are derived for outage probabilities of a $2 \times n$ system when the optimal ordering is implemented. The effect of the optimal ordering is shown to be to increase the first step SNR by 3 dB (rather than to increase the diversity order as one might intuitively expect based on the selection combining argument) and to increase the second step outage probability twice (this is the "price" to pay for the 3-dB SNR increase at the first step). Numerical Monte Carlo simulations validate the analytical results. The analysis results above allow one to efficiently estimate the average bit-error rate (BER) for a Rayleigh channel since the probability density function (pdf) of instantaneous SNR is expressed analytically in a closed from (as a combination of polynomials and exponents) and, hence, does not require time-consuming numerical evaluation (especially in the low-probability region).

Manuscript received December 13, 2002; revised March 21, 2003; revised May 10, 2003; accepted May 12, 2003. The associate editor coordinating the review of this paper and approving it for publication is Dr. M. Shafi. The paper was presented in part at 2002 International Zurich Seminar on Broadband Communications, Feb. 19–21, Zurich, Switzerland, and at the 2002 IEEE Vehicular Technology Conference, 24–28 September, Vancouver, BC, Canada.

S. Loyka is with the School of Information Technology and Engineering (SITE), University of Ottawa, Ottawa, ON K1N 6N5, Canada (e-mail: sergey.loyka@ieee.org).

F. Gagnon is with the Department of Electrical Engineering, École de Technologie Supérieure 1100, Montreal, QC H3C 1K3, Canada.

Digital Object Identifier 10.1109/TWC.2004.830853

II. V-BLAST Algorithm

The V-BLAST algorithm has been discussed in details elsewhere [5], [9]. Here, we describe its main points for completeness and in order to introduce notations. The main idea of the BLAST architecture is to split the information bit stream into several substreams and transmit them in parallel using a set of Tx antennas (the number of Tx antennas equals the number of substreams) at the same time and frequency. At the Rx side, each Rx antennas "sees" all the transmitted signals, which are mixed due to the nature of the wireless propagation channel. Using appropriate signal processing at the Rx side, these signals can be unmixed so that the matrix wireless channel is transformed into a set of virtual parallel independent channels (provided that mutltipath is rich enough).

The following basic assumptions are employed.

- The channel is random, quasistatic (i.e., fixed for every frame of information bits but varying from frame to frame), frequency independent (i.e., negligible delay spread) and with complex additive white Gaussian noise (AWGN).
- The Tx signal vector is comprised of individual symbol substreams. No space-time coding is employed. However, conventional coding can be used for each substream individually (but no intersubstream coding is allowed).
- The noise vector is comprised of independent AWGN components with equal variance.
- The Tx signals, noise and channel gains are independent of each other.
- Perfect channel knowledge is assumed to be available at the receiver.
- There is no performance degradation due to synchronization and timing errors.

The received signal vector $\mathbf{r}$ can be presented in the following complex baseband vector form [9]:

$$\mathbf{r} = \mathbf{H}\mathbf{q} + \mathbf{v} \tag{1}$$

where $\mathbf{q} = [q_1, \ldots, q_m]^{\mathrm{T}}$ is the transmitted symbol vector, $\mathbf{H}$ is the channel matrix (i.e., the matrix of complex transfer factors from each Tx to each Rx antenna), and $\mathbf{v} = [v_1, \ldots, v_n]^{\mathrm{T}}$ is the noise vector. Presenting the channel matrix in a column-wise way, $\mathbf{H} = [\mathbf{h}_1, \ldots, \mathbf{h}_m]$, where $\mathbf{h}_{\mathrm{i}}$ is a column vector of transfer factors from the ith Tx antenna to all Rx antennas, the received signal can be presented as

$$\mathbf{r} = \sum_i \mathbf{h}_i q_i + \mathbf{v}. \tag{2}$$

The V-BLAST processing begins with the first Tx symbol and proceeds in sequence to the mth symbol. When the optimal ordering procedure is employed, the Tx indexing is changed prior to the processing. The main steps of the V-BLAST processing (detection) algorithm are as follows [5], [9].

1) *The interference cancellation step*: At the ith processing step (i.e., when the signal from the ith transmitter is detected) the interference from the first $i-1$ transmitters, can be subtracted based on the estimations of the Tx symbols (which are actually assumed to be error-free) and the knowledge of $\mathbf{H}$.
2) *The interference nulling step*: Based on the knowledge of the channel matrix, the interference from yet to be detected symbols can be nulled out using the Gram–Schmidt orthogonalization process (applied to the column vectors of $\mathbf{H}$).
3) *The optimal ordering procedure*: the order of symbol processing is organized according to their after-processing SNR's in the decreasing order (i.e., the symbol with highest SNR is detected first).

III. Analysis of the V-BLAST Algorithm

For the sake of notational simplicity, we first describe all the steps without the noise contribution ($\mathbf{v} = \mathbf{0}$), which is added to the analysis later.

The interference cancellation step can be expressed mathematically in a straightforward way [9]. The received signal after the cancellation at the ith step is

$$\mathbf{r}_i' = \mathbf{r} - \sum_{j=1}^{i-1} \mathbf{h}_j \hat{q}_j \tag{3}$$

where $\hat{q}_j$ are the estimations of the already-detected symbols. In the further analysis, we assume for the sake of simplicity that they are error-free. If no substream coding is used, the system performance can degrade considerably in the low SNR regime due to an estimation error, which can further propagate to the next detection step. However, this effect is negligible (second-order in terms of probabilities) in high SNR regime [9]; it is also negligible if appropriate substream coding is used so that BER is low. Hence, we further neglect this effect. It should also be noted that, as the analysis in Section VI demonstrates [especially (36)], in order to estimate the total BER, only the conditional (conditioned on no error at the first step, i.e., no error propagation) outage probability at the second step is required. Hence, our analysis will provide a rigorous result in terms of the total BER.

The interference nulling step is based on the Gram–Schmidt ortogonalization procedure, which builds a set of orthogonal vectors from a set of linearly independent vectors. At this stage, we assume that $\mathbf{h}_i$ are linearly independent (otherwise the V-BLAST algorithm must be modified taking into account all the linearly dependent column vectors and decreasing the number of independent bit substreams). Using the closed-form analytical expression for the Gram–Schmidt process [10, Section IX.6] and after some mathematical development (see Appendix A for details), we arrive to the following expression of the received vector after interference nulling out and cancellation at the ith step

$$\mathbf{r}_i'' = \frac{q_i |\mathbf{h}_i|}{\left|\mathbf{R}^{[i+1,m]}\right|} \begin{vmatrix} \boldsymbol{\eta}_i & \boldsymbol{\eta}_i\boldsymbol{\eta}_{i+1} & \cdots & \boldsymbol{\eta}_i\boldsymbol{\eta}_m \\ \boldsymbol{\eta}_{i+1} & & & \\ \cdots & & \mathbf{R}^{[i+1,m]} & \\ \boldsymbol{\eta}_m & & & \end{vmatrix} \tag{4}$$

where $|\mathbf{h}_i| = \sqrt{\sum_j |h_{\mathrm{j}i}|^2}$, $|\ |$ means determinant when applied to a matrix, $\boldsymbol{\eta}_i = \mathbf{h}_i/|\mathbf{h}_i|, \boldsymbol{\eta}_i\boldsymbol{\eta}_j$ is the inner product ($\mathbf{xy} = \sum_{k=1}^{n} x_k y_k^*$), * denotes complex conjugate, and $\mathbf{R}^{[i+1,m]}$ is the

normalized instantaneous (i.e., for a given channel realization) channel correlation matrix built on $[\boldsymbol{\eta}_{i+1} \ \cdots \ \boldsymbol{\eta}_m]$

$$\mathbf{R}^{[i+1,m]} = \begin{bmatrix} 1 & \boldsymbol{\eta}_{i+1}\boldsymbol{\eta}_{i+2} & \cdots & \boldsymbol{\eta}_{i+1}\boldsymbol{\eta}_m \\ \boldsymbol{\eta}_{i+2}\boldsymbol{\eta}_{i+1} & 1 & \cdots & \boldsymbol{\eta}_{i+2}\boldsymbol{\eta}_m \\ \cdots & \cdots & \cdots & \cdots \\ \boldsymbol{\eta}_m\boldsymbol{\eta}_{i+1} & \boldsymbol{\eta}_m\boldsymbol{\eta}_{i+2} & \cdots & 1 \end{bmatrix}. \quad (5)$$

Determinant in (4) should be understood as the first column Laplace expansion, $\det = \sum_{j=1}^{m-i+1}(-1)^{1+j}M_{j1}\boldsymbol{\eta}_{i+j-1}$, where $\mathrm{M_{j1}}$ is the minor of the determinant obtained by striking out jth row and first column. Note that M_{j1} is a scalar for every j and, hence, the determinant itself is a vector, as it should be. The signal power for a given channel instant $\mathbf{H}$ is simply expressed as

$$|\mathbf{r}_i''|^2 = |q_i|^2|\mathbf{h}_i|^2\frac{\left|\mathbf{R}^{[i,m]}\right|}{\left|\mathbf{R}^{[i+1,m]}\right|}. \quad (6)$$

From this result and using (12), it is straightforward to obtain an instantaneous (i.e., for a given channel $\mathbf{H}$) SNR and, hence, an instantaneous BER for a particular modulation scheme. It is instructive to consider the case of $\mathrm{m} = 2$. At the first processing step, one obtains

$$|\mathbf{r}_1''|^2 = |q_1|^2|\mathbf{h}_1|^2(1 - |R_{12}|^2) \quad (7)$$

where $R_{12} = \boldsymbol{\eta}_1\boldsymbol{\eta}_2$. Hence, the received power and, consequently, SNR, is determined by the total received power from the first Tx antenna (transmitting the first bit stream), e.g., $|q_1|^2|\mathbf{h}_1|^2$, and by the normalized channel correlation coefficient R_{12}. We would like to emphasize the similarity of the results above to the analogous results on the MIMO channel capacity [6], [7], which is also determined by the channel correlation matrix (especially for the case of 2×2 MIMO architecture, i.e., [11]). One could intuitively conclude from (7) that the diversity order is n because of $|\mathbf{h}_1|^2$, which actually means nth-order maximum-ratio combining (MRC). However, as we prove later, the effect of the last factor in (7) is such that the actual diversity order is $n-1$. Let us now consider the optimal ordering procedure. To separate the effect of the transmitted symbol power (i.e., $|q_1|^2$) and of the noise power from the effect of the propagation channel, we assume that all $|q_i|$ are equal (i.e., constant amplitude modulation) and all per-branch noise powers are also equal. If the first Tx symbol is detected first, then the after-processing power is given by (7). If the second Tx symbol is detected first, then (7) should be changed to

$$|\mathbf{r}_1''|^2 = |q_2|^2|\mathbf{h}_2|^2(1 - |R_{12}|^2). \quad (8)$$

When the noise power is equal in all branches, the after-processing SNR is proportional to the received symbol power. Then, the optimal ordering is to detect first the symbol with the highest $|\mathbf{h}_i|^2$, i.e., the same as for the selection combining. However, as we show later, this does not result in the increase of the diversity order.

Let us consider the optimal ordering at the first step for arbitrary m. When the ith Tx symbol is detected first, the signal power after interference nulling out is

$$P_i = |q_i|^2|\mathbf{h}_i|^2\frac{|\mathbf{R}|}{\left|\mathbf{R}^{[i]}\right|} \quad (9)$$

where $\mathbf{R}$ is the full correlation matrix (i.e., built on $[\boldsymbol{\eta}_1 \ \cdots \ \boldsymbol{\eta}_m]$) and $\mathbf{R}^{[i]}$ is the correlation matrix built on all column vectors except for $\boldsymbol{\eta}_i$. Under the assumptions of equal $|q_i|$ and equal $|\mathbf{h}_i|$ (i.e., the same received power from every transmit antenna), the optimal ordering is to detect first the symbol with the smallest $|\mathbf{R}^{[i]}|$ (while it is unlikely to get equal $|\mathbf{h}_i|$ in a fading channel, this assumption allows to isolate the effect of correlation from the effect of unequal received powers). In fact, this means that the overall correlation among $[\boldsymbol{\eta}_1, \ \ldots, \ \boldsymbol{\eta}_{i-1}, \ \boldsymbol{\eta}_{i+1}, \ldots, \ \boldsymbol{\eta}_m]$ must be highest and, consequently, the correlation between $\boldsymbol{\eta}_i$ and $[\boldsymbol{\eta}_1, \ldots, \ \boldsymbol{\eta}_{i-1}, \boldsymbol{\eta}_{i+1}, \ldots, \ \boldsymbol{\eta}_m]$ must be the lowest (this follows from geometrical interpretation of $|\mathbf{R}|$ as a volume in the m-dimensional space [10]). Thus, the best ordering is to detect first that symbol whose column propagation vector has lowest correlation with the other vectors.

Let us now consider the effect of the noise. Equation (4) is generalized as follows:

$$\mathbf{r}_i'' = \mathbf{r}_{0,i}'' + \mathbf{v}_i'' \quad (10)$$

where $\mathbf{r}_{0,i}''$ is given by (4) and

$$\mathbf{v}_i'' = \frac{\mathbf{v}}{\left|\mathbf{R}^{[i+1,m]}\right|}\begin{vmatrix} 1 & \boldsymbol{\eta}_{i+1} & \cdots & \boldsymbol{\eta}_m \\ \boldsymbol{\eta}_{i+1} & & & \\ \cdots & & \mathbf{R}^{[i+1,m]} & \\ \boldsymbol{\eta}_m & & & \end{vmatrix}. \quad (11)$$

Using (11), the after-processing noise power at ith step can be simply expressed as

$$P_{\mathbf{v}_i} = \langle|\mathbf{v}_i''|^2\rangle = (n - m + i)\sigma_1^2 \quad (12)$$

where $\sigma_1^2 = \langle|\mathbf{v}_j|^2\rangle$ is per-branch noise power before processing, and $\langle\ \rangle$ is the expectation over noise voltage (see Appendix A for detailed derivation). Note that the after-processing noise power is less than the total noise power, which is $n\sigma_1^2$. This is the consequence of the orthogonal projection performed by the Gram–Schmidt process (see Fig. 1). One also should note that the after-processing noise power increases with i (step index), being the smallest in the first step and the same as the total noise power in the last step. Geometrical interpretation of the noise transformation during the V-BLAST processing is the same as in Fig. 1.

IV. Fading Outage Curves and Diversity Order

Based on the results above, let us now analyze the signal fading in the V-BLAST system. In particular, we consider the outage probabilities [the probability that the signal level is less than the specified value, which is the same as a cumulative distribution function (cdf)] and diversity order (the asymptotic slope of the outage probability curve).

We assume that the channel gains (i.e., the components of $\mathbf{H}$) are independent identically distributed (i.i.d.) complex Gaussians with zero mean and unit variance (i.e., we consider only the channel variation due to multipath and ignore the absolute propagation loss and large-scale variation due to shadowing). First, we ignore the optimal ordering procedure and prove that the diversity order at the ith step is $(n - m + i)$.

To demonstrate the main idea of the proof, let us consider first the case of $n = m = 2$, i.e., $\mathbf{H} = [\mathbf{h}_1 \ \mathbf{h}_2]$. To be specific, we

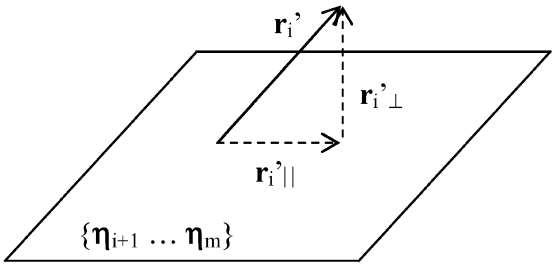

Fig. 1. Geometric illustration of the interference nulling out step: the received vector (after the interference cancellation) is decomposed into orthogonal and parallel components with respect to the space spanned by $\{/eta_{i+1}, \ldots, /eta_m\}$.

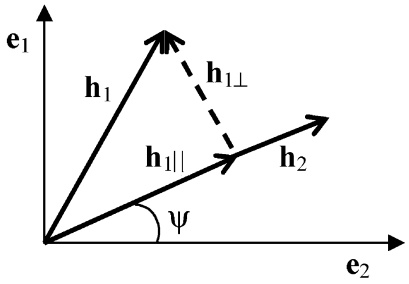

Fig. 2. Geometrical representation of interference nulling out: decomposition of $\mathbf{h}_1$ into $\mathbf{h}_{1\perp}$ and $\mathbf{h}_{1\|}$. $\mathbf{e}_1$ and $\mathbf{e}_2$ are basis vectors of the space spanned by the columns of $\mathbf{H}$ ($\mathbf{h}_1$ and $\mathbf{h}_2$).

assume that the first Tx symbol is detected first. The interference nulling out can be expressed is a general matrix form

$$\mathbf{r}_\perp = \boldsymbol{Q} \cdot \mathbf{r} \tag{13}$$

where $\boldsymbol{Q}$ is an orthogonal projection matrix, which projects $\mathbf{r}$ to the direction orthogonal to $\mathbf{h}_2$. Substituting (13) into (2), one obtains (since we are interested in the received signal power only, we ignore noise in this section)

$$\mathbf{r}_\perp = q_1 \boldsymbol{Q} \cdot \mathbf{h}_1. \tag{14}$$

This means that the signal after interference nulling out is proportional to that part of $\mathbf{h}_1$ which is orthogonal to $\mathbf{h}_2$, see Fig. 2, and the signal power $\sim|\mathbf{h}_{1\perp}|^2$. But the vector magnitude is not affected by rotation on an arbitrary angle. We rotate $[\mathbf{h}_1\ \mathbf{h}_2]$ as a whole on angle ψ so that $\mathbf{h}_2$ is parallel to $\mathbf{e}_2 : h_{1,2} = 0$. This can be expressed as

$$\tilde{\mathbf{h}}_i = \boldsymbol{A} \cdot \mathbf{h}_i \tag{15}$$

where $\boldsymbol{A}$ is the rotation matrix, which satisfies (preservation of length)

$$A \cdot A^+ = A^+ \cdot A = I \tag{16}$$

where "+" denotes conjugate transpose. Using (15), one obtains: $|\mathbf{h}_{1\perp}| = |\tilde{h}_{1,1}|$. It is straightforward to show using (16) that the components of $\tilde{\mathbf{h}}_1$ have the same distribution as the components of $\mathbf{h}_1$ (note that ψ is independent of $\mathbf{h}_1$), i.e., i.i.d. complex Gaussians with unit variance. Hence, $|\mathbf{h}_{1\perp}|^2$ is chi-squared random variable with two degrees of freedom, $|\mathbf{h}_{1\perp}|^2 \sim \chi^2_2$. The same is true for the signal power. Thus, the diversity order in the first step is one. The similar consideration for arbitrary n leads to the conclusion that $|\mathbf{h}_{1\perp}|^2 \sim \chi^2_{2(n-1)}$ (simply because $\mathbf{h}_{1\perp}$ has $n-1$ nonzero components after rotation) and the diversity order is $(n-1)$.

The case of arbitrary m is somewhat more complex, however, straightforward to consider in the similar way. First, we rotate the set $[\mathbf{h}_1, \ldots, \mathbf{h}_m]$ as a whole so that $\mathbf{h}_m$ becomes parallel to $\mathbf{e}_m$, where $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_m$ are basis vectors of the space spanned by $\mathbf{h}_1, \mathbf{h}_2, \ldots, \mathbf{h}_m$. In the second rotation we keep $\mathbf{h}_m$ fixed (i.e., a rotation around the $\mathbf{e}_m$ axis) and position $\mathbf{h}_{m-1}$ into the $[\mathbf{e}_{m-1}\ \mathbf{e}_m]$ plane. The rotations are continued until $\mathbf{h}_2$ is positioned into the $[\mathbf{e}_2\ \mathbf{e}_3, \ldots, \mathbf{e}_m]$ hyper plane. After the rotation, $\mathbf{h}_{1\perp}$ has $(n-m+1)$ nonzero components. Every such rotation preserves the distribution of the components. Hence, $|\mathbf{h}_{1\perp}|^2 \sim \chi^2_{2(n-m+1)}$ and the diversity order is $(n-m+1)$. Similar consideration for the ith step leads to the conclusion that $|\mathbf{h}_{i\perp}|^2 \sim \chi^2_{2(n-m+i)}$ and the diversity order is $(n-m+i)$. Note that the lowest diversity order is at the first step and the highest is at the last (i.e., n). When $n = m$, no diversity is obtained at the first step.

V. The Effect of Optimal Ordering

For the sake of simplicity, we assume that all transmitted symbols have the same unit power and that the channel coefficients are i.i.d. complex Gausians with unit variance and zero mean (i.e., Rayleigh fading with $\langle|h_{ij}|^2\rangle = 1$). Since we are interested in outage probability (i.e., the probability that the received signal power drops below a given value), we consider the received powers only (in each Rx antenna from each Tx antenna). Normally, outage probability is defined in terms of SNR. However, since the noise powers in all the receive branches are assumed to be equal and the optimal ordering does not affect noise, it is equivalent to the definition in terms of signal power.

Under the assumptions above, the vector signal received by the Rx antennas from the ith Tx antenna is

$$\mathbf{r}_i = \mathbf{h}_i. \tag{17}$$

We consider first 2×2 V-BLAST and further it is generalized to the $2 \times \mathrm{n}$ case.

A. 2×2 Case

Consider the following transformation of the received vectors (as a whole) illustrated in Fig. 3. Since it is the rotation, the vector lengths (i.e., the signal magnitudes) as well as the angle φ are not changed. This rotation is similar to the one in the previous section. Hence, the primed vector lengths have the same distribution as the unprimed ones, namely

$$|\mathbf{h}_{1\perp}|^2 \sim \chi^2_2, \quad \mathbf{h}_{1\|} \sim \chi^2_2, \quad |\mathbf{h}_1|^2 \sim \chi^2_4. \tag{18}$$

Thus, we further use unprimed notations. The components of $\mathbf{h}_2$ have the same distributions as in (18). The outage probability (i.e., cdf) for $|\mathbf{h}_1|^2$ (or $|\mathbf{h}_2|^2$) is

$$\Pr[|\mathbf{h}_1|^2 < x] = \Pr[|\mathbf{h}_2|^2 < x] = F_h(x) = 1 - e^{-x}(1+x) \tag{19}$$

i.e., the second-order MRC. The optimal ordering procedure (after the interference nulling) can be described as follows:

$$s_1 = \max[|\mathbf{h}_{1\perp}|^2, |\mathbf{h}_{2\perp}|^2] = (\sin\varphi)^2 \max[|\mathbf{h}_1|^2, |\mathbf{h}_2|^2] \tag{20}$$

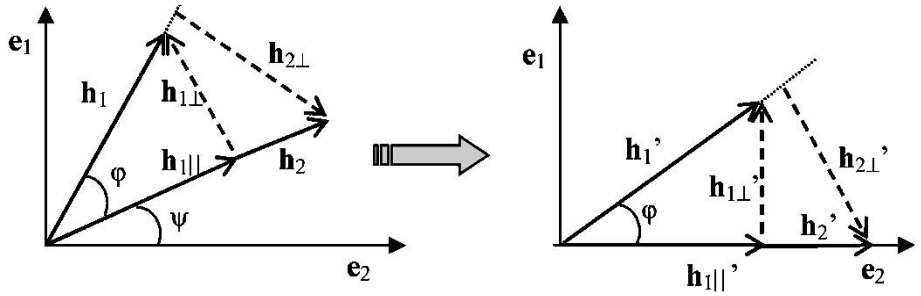

Fig. 3. Rotation of the received vectors by angle Ψ.

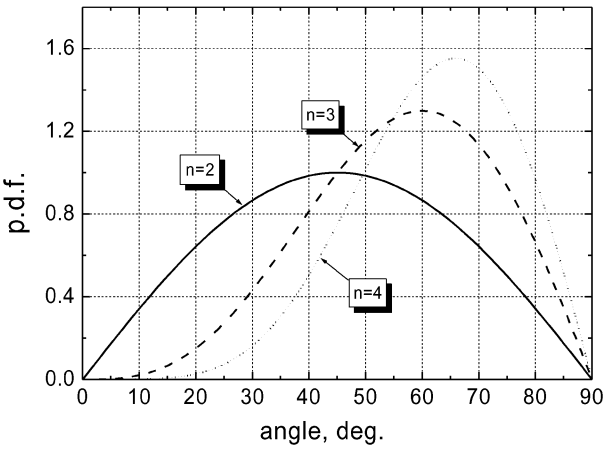

Fig. 4. Probability density function (pdf) of φ for various numbers of receive antennas (n).

where s_1 is the signal power after the optimal ordering, i.e., we compare $|\mathbf{h}_{1\perp}|$ and $|\mathbf{h}_{2\perp}|$ and take the maximum. In fact, (20) tells us that the optimal ordering 2×2 for system is to detect first the substream with the largest before-detection power [i.e., the max at the right-hand side (RHS) of (20)]. While $|\mathbf{h}_{1\perp}|^2$ and $|\mathbf{h}_{2\perp}|^2$ are χ_2^2 (i.e., Rayleigh distributed), taking the maximum does not result in second-order diversity because they are not independent, as $\sin\varphi$ at the RHS of (20) indicates. Using the fact that the distribution of $\max[|\mathbf{h}_1|^2, |\mathbf{h}_2|^2]$ is $F_h^2(x)$, the distribution $F_1(x)$ of s_1 may be presented in the following form:

$$F_1(x) = \Pr[s_1 < x] = \Pr\left[\max[|\mathbf{h}_1|^2, |\mathbf{h}_2|^2] < \frac{x}{\sin^2\varphi}\right] = \int_0^{\pi/2} F_h^2\left(\frac{x}{\sin^2\varphi}\right) f_\varphi(\varphi)\, d\varphi \tag{21}$$

where $f_\varphi(\varphi)$ is the pdf of φ (note that it is not uniform). It can be shown that $f_\varphi(\varphi) = \sin 2\varphi$ (see Appendix B). This pdf is illustrated in Fig. 4. The most probable direction is 45° as one would intuitively expect (because $|\mathbf{h}_{1\perp}|^2 \sim \chi_2^2, |\mathbf{h}_{1||}|^2 \sim \chi_2^2$ and, hence, the most probable values of the abscissa ($|\mathbf{h}_{1||}|$) and ordinate ($|\mathbf{h}_{1\perp}|$) are the same (see Fig. 3). The probability to get the angle close to 0° (90°) is very small because it means that $|\mathbf{h}_{1||}|$ ($|\mathbf{h}_{1\perp}|$) can be any and $|\mathbf{h}_{1\perp}|$ ($|\mathbf{h}_{1||}|$) must be close to 0, and this probability is small.

It should be noted that (21) holds true in the $2\times\mathrm{n}$ case as well, provided that the appropriate expressions are used for $f_\varphi(\varphi)$ and $F_h(x)$. Evaluating the integral in (21), one obtains, after some manipulations, the following simple expression for the outage probability at the first detection step with the optimal ordering

$$F_1(x) = 1 - 2e^{-x} + \left(1 + \frac{x}{2}\right) e^{-2x}. \tag{22}$$

The detailed proof is given in Appendix B. The asymptotic behavior of this outage probability (i.e., in the small outage probability region) is

$$F_1(x) \approx \frac{x}{2}, \quad x \to 0. \tag{23}$$

Comparing this with the asymptotic behavior of the Rayleigh distribution ($F_R(x) \approx x,\ x \to 0$), we conclude that the effect of optimal ordering is to increase SNR (or decrease outage probability) by 3 dB rather than to increase the diversity order, as one might intuitively expect based on the selection combining argument [see (20)]. The reason for this is that $|\mathbf{h}_{1\perp}|$ and $|\mathbf{h}_{2\perp}|$ are correlated.

The outage probability at the second detection step can be derived using an expression similar to (20). In particular, we note that at the second step we have to use the received vector with the minimum length $\min[|\mathbf{h}_1|, |\mathbf{h}_2|]$ because the vector with the maximum length was used in the first step. We also note that there is no need for interference nulling at this step. Hence

$$s_2 = \min[|\mathbf{h}_1|^2, |\mathbf{h}_2|^2] \tag{24}$$

and the outage probability $F_2(x)$ is

$$F_2(x) = \Pr[s_2 < x] = 1 - [1 - F_h(x)]^2 = F_h(x)[2 - F_h(x)]. \tag{25}$$

Its asymptotic behavior is

$$F_2(x) \approx 2F_h(x) \approx x^2, \quad x \to 0. \tag{26}$$

Comparing it with the second order MRC outage probability, $F_{\mathrm{MRC}}(x) = F_h(x) \approx x^2/2$, we conclude that the effect of optimal ordering at the second detection step is to increase outage probability twice (1.5-dB loss in SNR). This is the "price" one has to pay for the 3-dB increase in SNR at the first step. It should be noted that (25) holds true in the $2\times\mathrm{n}$ case as well, provided that the appropriate expression is used for $F_h(x)$.

B. $2\times\mathrm{n}$ *Case*

In this case

$$|\mathbf{h}_{1\perp}|^2 \sim \chi_{2n-2}^2, \quad \mathbf{h}_{1||} \sim \chi_2^2, \quad |\mathbf{h}_1|^2 \sim \chi_{2n}^2 \tag{27}$$

$$F_h(x) = \Pr[|\mathbf{h}_1|^2 < x] = 1 - e^{-x}\sum_{k=0}^{n-1}\frac{x^k}{k!}. \tag{28}$$

The same distributions hold true for $\mathbf{h}_2$ as well. Using a techniques similar to that in ([12], p. 191), it can be shown that in this case $f_\varphi(\varphi) = 2(n-1)(\sin\varphi)^{2n-3}\cos\varphi$ (see Appendix B for details). This pdf is illustrated in Fig. 4 for $n = 2, 3, 4$. The most probable direction shifts to the right (90°) when n increases, starting from 45° for $n = 2$. This is what should be intuitively expected because $\mathbf{h}_{1\|} \sim \chi^2_2$ and $|\mathbf{h}_{1\perp}|^2 \sim \chi^2_{2n-2}$, and, hence, the most probable value of the ordinate ($|\mathbf{h}_{1\perp}|$) is greater than the most probable value of the abscissa ($|\mathbf{h}_{1\|}|$) (for $n > 2$), and the former increases with n while the later is fixed. Using (21) and the pdf above, one obtains

$$F_1(x) = \int_0^{\pi/2} F_h^2\left(\frac{x}{\sin^2\varphi}\right) f_\varphi(\varphi)\, d\varphi = (n-1)\int_0^1 F_h^2\left(\frac{x}{t}\right)\cdot t^{n-2} dt. \tag{29}$$

After some manipulations, (29) reduces to

$$F_1(x) = 1 - (n-1)e^{-x}\left[\sum_{i=0}^{n-2}(2a_i - e^{-x}[b_i + (2x)^{n-1}c_i])x^i\right] \tag{30}$$

where

$$a_i = \sum_{m=0}^{i}\frac{(-1)^{i-m}(n-i-2)!}{(n-m-1)!m!}$$

$$b_i = (-2)^i(n-i-2)!\sum_{k=0}^{i}\frac{(-1)^k}{k!(n-1-k)!}$$

$$c_i = \frac{2^i}{i!}\sum_{j=i+n}^{2n-2}\sum_{k=j-n+1}^{n-2}\frac{(j-n)!2^{-j}}{k!(j-k)!}. \tag{31}$$

The details of the proof are given in Appendix B. Hence, a general form of $F_1(x)$ is

$$F_1(x) = 1 - p_1(x)e^{-x} + p_2(x)e^{-2x} \tag{32}$$

where $p_1(x)$ and $p_2(x)$ are polynomials of degree at most $(n-2)$ and $(2n-3)$ correspondingly. The asymptotic behavior of the outage probability is

$$F_1(x) \approx \frac{1}{(n-1)!}\left(\frac{x}{2}\right)^{n-1}, \quad x \to 0. \tag{33}$$

Comparing it with $(n-1)$-order MRC asymptotic behavior, $F_{\mathrm{MRC}}(x) \approx (1/(n-1)!)x^{n-1}$, we conclude that the effect of the optimal ordering at the first detection step is to increase SNR by 3 dB rather than to increase the diversity order. It is interesting to note that the conclusion proved to be true for 2×2 system, is also true in the general $2\times n$ case. The outage probability at the second detection step is given by (25), where $F_h(x)$ is that in (28). We do the same conclusion as in the 2×2 case: the effect of optimal ordering at the second detection step is to increase the outage probability twice. This is the "price" to pay for the increased SNR at the first step.

It should be noted that F_2 represents a conditional (conditioned on no error at the first step, i.e., no error propagation) outage probability at the second step. Assuming that an error at the first step automatically results at an error at the second step, unconditional outage probability can be estimated as follows: $F_2' = F_2(1-P_1) + P_1$, where P_1 is the BER at the first step. Clearly, F_2' is lower limited by P_1. Note that error propagation does not affect the first step outage probability. As the analysis in the next section shows, in order to evaluate the total BER one needs the conditional outage probability only.

VI. Average BER Analysis

Using the results above we can analyze now average BER performance for a particular modulation format in the same way as it is done for diversity combining systems (see, for example, [14] for detailed discussion). In doing so, we have to use the instantaneous SNR distributions derived above. Hence, the average BER will be different at different processing steps since the SNR distribution varies from step to step and since the noise power also varies from step to step. It should be noted that, while the V-BLAST is a coherent algorithm (and, hence, an MRC combining can be used), we carried out the analysis above in terms of total post-processing signal and noise powers and this implicitly corresponds to noncoherent equal gain combining (EGC). Thus, we further assume that this type of combining is employed. While practical applications of noncoherent EGC may be limited, the mathematical analysis becomes tractable analytically and some important insights may be obtained in this way. It will also provide a lower bound on the MRC performance, which is known to be superior to EGC.

The purpose of this section is to illustrate the analytical BER analysis, which is possible using the results above, rather than to give a comprehensive exposition of the topic.

We consider first the case of 2×2 system. Generalization to $2\times \mathrm{n}$ systems is straightforward (however, it involves more lengthy mathematics). The average BER at ith processing step can be expressed as

$$\overline{P_{e,i}} = \int_0^\infty \rho_i(\gamma)P_e(\gamma)\, d\gamma \tag{34}$$

where $P_e(\gamma)$ is the instantaneous BER or probability of error (i.e., for a given instantaneous SNR γ), and $\rho_i(\gamma)$ is the SNR pdf at the ith step, which can be obtained from the outage probability expressions derived above

$$\rho_1(\gamma) = \frac{d}{d\gamma}F_1\left(\frac{\gamma}{\overline{\gamma_0}}\right), \quad \rho_2(\gamma) = \frac{d}{d\gamma}F_2\left(\frac{2\gamma}{\overline{\gamma_0}}\right) \tag{35}$$

where $\overline{\gamma_0}$ is the average before-processing SNR per branch per transmitter (i.e., when only one transmitter is active), which is assumed to be identical for all branches and for both transmitters. Since the analysis above was carrier out for a normalized signal power, the outage probabilities (cdfs) in terms of instantaneous SNR are obtained by substitution $x \to \gamma/\overline{\gamma_0}$ for the first step and by $x \to 2\gamma/\overline{\gamma_0}$ for the second one (the factor 2 is due to the fact that the total post-processing noise power at the second step is twice of the branch noise power—see (12)). In fact, the average BER at the second step as defined by (34) ignores the error propagation from the first step, i.e., it is a conditional BER (conditioned on having no error at the first step). This is exactly

what one needs to evaluate the total probability of error, i.e., the probability that there is at least one error (either at first step, or at second, or at both). The total instantaneous probability of error is

$$P_{e,\text{tot}} = P_{e,1} + P_{e,2}(1 - P_{e,1}) \tag{36}$$

where $P_{e,1} = P_e(\gamma_1)$ is the instantaneous BER and γ_1 is the instantaneous SNR, both at first step, $P_{e,2} = P_e(\gamma_2)$ is the conditional instantaneous BER and γ_2 is the instantaneous SNR, both at second step. Note that (36) is rigorous and it accounts for the propagation of error (through the first term in (36), which describes 2 events: (i) error at first step and no error at second; (ii) error at first step and error at the second step). The total average BER is

$$\begin{aligned}\overline{P_{e,\text{tot}}} &= \langle P_{e,1}\rangle_{\gamma_1} + \langle P_{e,2}\rangle_{\gamma_2} - \langle P_{e,1}P_{e,2}\rangle_{\gamma_1,\gamma_2} \\ &\le \overline{P_{e,1}} + \overline{P_{e,2}}\end{aligned} \tag{37}$$

where $\langle x\rangle$ is the expectation of x (over channel fading), $\overline{P_{e,1}} = \langle P_{e,1}\rangle_{\gamma_1}$ and $\overline{P_{e,2}} = \langle P_{e,2}\rangle_{\gamma_2}$. The third term in (37) cannot be easily evaluated since γ_1 and γ_2 are not statistically independent. Hence, we use the upper bound at the right and, using Monte Carlo simulations, show that this bound is very tight (the difference in BER is less than 10% for $\bar{\gamma}_0 \ge 5$ dB). An intuitive explanation for this is that the third term in (37) is a "second-order" one and, hence, for reasonably large average SNR it should not affect significantly the total average BER.

We further assume that noncoherent binary orthogonal FSK is used (other modulation schemes can be analyzed in a similar way; however, analytical analysis may not be always possible), whose instantaneous BER is [14]

$$P_e(\gamma) = \frac{1}{2}\exp\left(-\frac{\gamma}{2}\right). \tag{38}$$

Using (34), (35), and (38) one obtains

$$\overline{P_{e,1}} = \frac{1}{2} - \frac{\overline{\gamma_0}}{2+\overline{\gamma_0}} + \frac{\overline{\gamma_0}}{2}\left(\frac{1}{4+\overline{\gamma_0}} + \frac{1}{(4+\overline{\gamma_0})^2}\right) \tag{39}$$

$$\overline{P_{e,2}} = \frac{16}{(8+\overline{\gamma_0})^2} + \frac{128}{(8+\overline{\gamma_0})^3}. \tag{40}$$

Using (37), (39), and (40), the total average BER can be easily estimated using the tight upper bound in (37). Asymptotically, at high average SNR regime ($\overline{\gamma_0} \to \infty$), one obtains

$$\overline{P_{e,1}} \approx \frac{1}{2\overline{\gamma_0}}, \quad \overline{P_{e,2}} \approx \frac{16}{\overline{\gamma_0}^2}. \tag{41}$$

As it follows from (39) and (40), (41) provides good approximations provided that $\overline{\gamma_0} \gg 4$ (first step) and $\overline{\gamma_0} \gg 8$ (second step). Note that in high SNR regime the total average BER is dominated by the first-order term

$$\overline{P_{e,\text{tot}}} \approx \overline{P_{e,1}} + \overline{P_{e,2}} \approx \overline{P_{e,1}} \approx \frac{1}{2\overline{\gamma_0}} \tag{42}$$

which agrees well with intuitive expectation. These analytical results were validated using extensive Monte Carlo simulations as discussed in the next section. Comparing (41) and (42) with (23) and (26) one may conclude that the average BER at high SNR regime behaves in the same way as the outage probability (up to a constant), as it is the case for conventional diversity combining systems. Comparing (41) to the average BER of a conventional noncoherent FSK, we conclude that the effect of the optimal ordering is to increase SNR asymptotically by 3 dB.

The generalization of the results above for $2 \times \mathrm{n}$ system is straightforward (however, it involves more lengthy mathematics). We give here the final results for the average BER

$$\begin{aligned}\overline{P_{e,1}} = \frac{1}{2} &- \frac{\overline{\gamma_0}}{2}\sum_{i=0}^{n-2} i! a_i \left(\frac{2(n-1)}{2(n-1)+\overline{\gamma_0}}\right)^{i+1} \\ &+ \frac{\overline{\gamma_0}}{4}\sum_{i=0}^{2n-3} i!\beta_i \left(\frac{2(n-1)}{4(n-1)+\overline{\gamma_0}}\right)^{i+1}\end{aligned} \tag{43}$$

$$\begin{aligned}\overline{P_{e,2}} = &\left(\frac{2n}{4n+\overline{\gamma_0}}\right)^n + (2^n-2)\left(\frac{2n}{4n+\overline{\gamma_0}}\right)^{n+1} \\ &- \frac{\overline{\gamma_0}}{4n}\sum_{i=n+2}^{2n-1}\xi_i\left(\frac{2n}{4n+\overline{\gamma_0}}\right)^i\end{aligned} \tag{44}$$

where

$$\begin{aligned}\beta_i &= \begin{cases} b_i, & i \le n-2 \\ 2^{n-1}c_{i-n+1}, & i \ge n-1 \end{cases} \\ \xi_i &= \sum_{k=i-n}^{n-1}\frac{(i-1)!}{k!(i-1-k)!}.\end{aligned} \tag{45}$$

The asymptotic behavior at high SNR regime is

$$\overline{P_{e,1}} \approx \frac{1}{2}\left(\frac{n-1}{\overline{\gamma_0}}\right)^{n-1}, \quad \overline{P_{e,2}} \approx \left(\frac{2n}{\overline{\gamma_0}}\right)^n. \tag{46}$$

Similarly to the case of 2×2 system, these approximations are accurate provided that that $\overline{\gamma_0} \gg 4(n-1)$ (first step) and $\overline{\gamma_0} \gg 4n$ (second step). For $n = 2$, these equations reduce to (39)–(41), as it must be.

Similarly to diversity combining systems [14], binary DPSK will perform better by 3 dB in terms of average SNR. Coherent modulation/demodulation formats can be analyzed in a similar way. Consider, for example, coherent binary phase-shift keying (BPSK) whose instantaneous BER is [14]

$$P_e(\gamma) = Q(\sqrt{2\gamma}), \quad Q(x) = \frac{1}{\sqrt{2\pi}}\int_x^\infty \exp\left(-\frac{t^2}{2}\right)dt. \tag{47}$$

We further use a modified form of (34) to simplify the development

$$\overline{P_{e,i}} = \int_0^\infty \rho_i(\gamma)P_e(\gamma)\,d\gamma = -\int_0^\infty F_i'(\gamma)\frac{dP_e(\gamma)}{d\gamma}\,d\gamma \tag{48}$$

where $F_1'(\gamma) = F_1(\gamma/\overline{\gamma_0}), F_2'(\gamma) = F_2(2\gamma/\overline{\gamma_0})$. Equation (48) allows for a closed-form solution. In the case of a 2×2 system, it is especially simple

$$\overline{P_{e,1}} = \frac{1}{2} - \sqrt{\frac{\overline{\gamma_0}}{1+\overline{\gamma_0}}} + \frac{1}{2}\sqrt{\frac{\overline{\gamma_0}}{2+\overline{\gamma_0}}}\left(1 + \frac{1}{4(2+\overline{\gamma_0})}\right) \tag{49}$$

$$\overline{P_{e,2}} = \frac{1}{2} - \sqrt{\frac{\overline{\gamma_0}}{4+\overline{\gamma_0}}}\left(\frac{1}{2} + \frac{1}{4+\overline{\gamma_0}} + \frac{3}{2(4+\overline{\gamma_0})^2}\right). \tag{50}$$

The asymptotic behavior of BER for large average SNR is

$$\overline{P_{e,1}} \approx \frac{1}{8\overline{\gamma_0}}, \quad \overline{P_{e,2}} \approx \frac{3}{2\overline{\gamma_0}^2}. \tag{51}$$

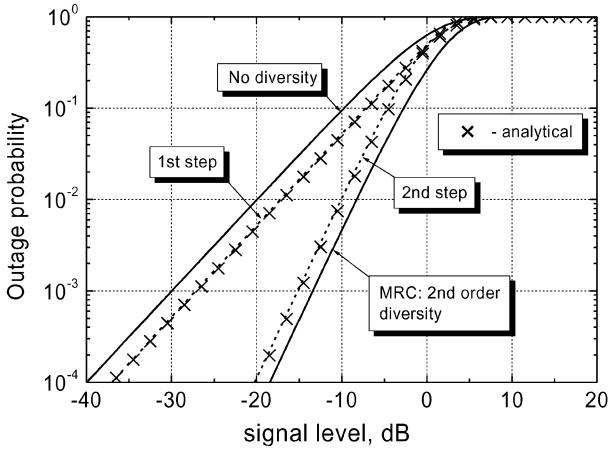

Fig. 5. Outage probability curves of the V-BLAST algorithm for $\mathrm{m} = \mathrm{n} = 2$.

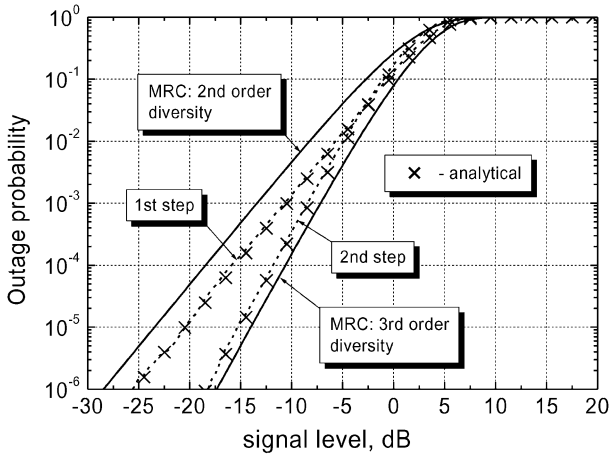

Fig. 6. Outage probability curves of the V-BLAST algorithm for$\mathrm{n} = 3$, $\mathrm{m} = 2$.

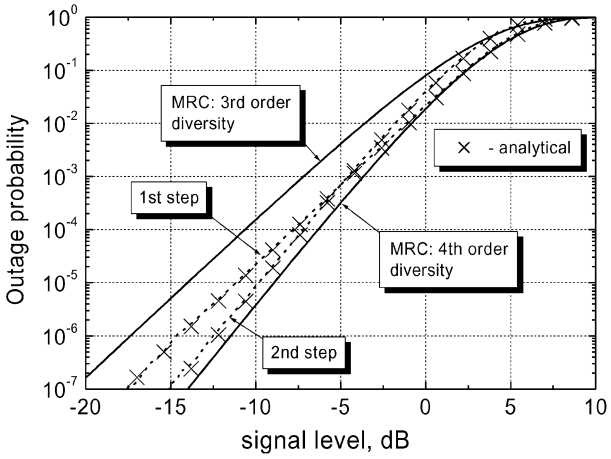

Fig. 7. Outage probability curves of the V-BLAST algorithm for $\mathrm{n} = 4$, $\mathrm{m} = 2$.

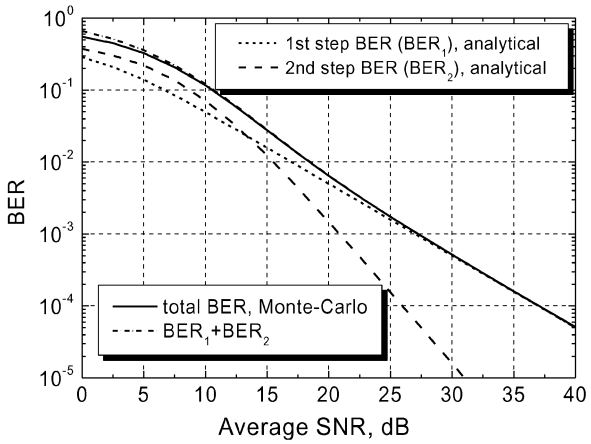

Fig. 8. Average BER versus average per-branch SNR for 2×2 system.

Comparing it with (41), we note that, similarly to conventional systems, there is 6-dB; improvement at first step with respect to noncoherent FSK. Comparing it with a conventional BPSK, the 3-dB improvement due to the optimal ordering is also obvious. The second step improvement is a bit less than 6 dB. second-order diversity is also apparent at this step. Similar analysis can be done for a $2 \times \mathrm{n}$ systems. However, the expressions become more complicated.

VII. Numerical Monte-Carlo Simulations

In order to validate the analytical results above, we use numerical Monte Carlo simulations for $m = 2$. First, the V-BLAST algorithm outage curves have been simulated without the optimal ordering. No difference has been observed between the analytical results above, which are shown as MRC curves on the graphs, and the Monte Carlo simulations (thus, the analytical results are not shown on the graphs), which validates the analytical results. Second, the V-BLAST outage curves have been simulated with the optimal ordering procedure. Some of the results are presented in Figs. 5–7. No significant difference has been found between analytical outage probability results and Monte Carlo simulations. This validates our conclusion that the effect of the optimal ordering is to increase signal power (and SNR) rather than to increase the diversity order. However, as shown on Fig. 7, it is difficult to observe this tendency using Monte Carlo simulations for reasonable outage probabilities when $n \geq 4$. The small difference between analytical and Monte Carlo results observed in these graphs may be attributed to three main reasons. First, we used histograms of 0.2-dB width to build the curves and this contributes to the small shift observed. This width is a compromise between the number of Monte Carlo simulations (trials) required (10^8 for Fig. 7) and time/computing power available. Decreasing the width results in smaller shift but it increases the "noise" from interval to interval for a given number of simulations. On the other hand, the number of simulations required is limited by the computer power available. Second, since we used only 10^8 trials for Fig. 7, some inaccuracies are produced at small probabilities (like 10^{-7}). Third, performing 10^8 Monte Carlo simulations may reveal some limitations of the random number generator we used (in particular—its uniformity). Of course, this effect is much more pronounced for a really large number of simulations and is negligible otherwise.

Fig. 8 shows the average BER performance of 2×2 V-BLAST as discussed in the previous section. Since no difference has been found between numerical Monte Carlo simulations and analytical results in (39) and (40), numerical

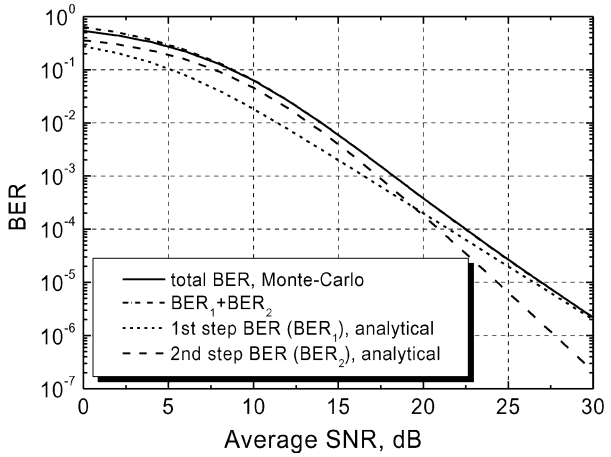

Fig. 9. Average BER versus average per-branch SNR for 2×3 system.

results are not shown. As Fig. 8 indicates, the upper bound in (37) is indeed a tight one for approximately $\bar{\gamma} \geq 5$ dB. The second step BER dominates the first step BER at $\bar{\gamma} < 13$ dB and when $\bar{\gamma} \geq 13$ dB the first step BER is dominant. The contribution of $\overline{P_{e,2}}$ to the total average BER is negligible and, hence, (42) holds true at approximately $\bar{\gamma} \geq 20$ dB. In this region, the total average BER is totally dominated by the first step BER and the second step BER can be ignored in estimating the total BER. Fig. 9 demonstrates that similar conclusions hold true for $n = 3$ as well.

VIII. Conclusion

Using a closed-from model of the Gramm–Schmidt process, we have developed an analytical approach to the performance analysis of the V-BLAST algorithm. In particular, closed-form analytical expressions have been presented for the signal and noise vectors at ith processing step, as well as for the outage probabilities. The after-processing signal power is determined by the channel instantaneous correlation matrices (in the same fashion as the channel capacity is). The optimal ordering is proved to be equivalent to the least correlation criterion.

Performing the statistical analysis analytically for a Rayleigh uncorrelated channel, we have proved that the diversity order at ith processing step is $(n - m + i)$, provided that no optimal ordering is used. For the $2 \times \mathrm{n}$ system, the effect of the optimal ordering at the first detection step is to increase SNR by 3 dB rather then to increase the diversity order (as one might intuitively expect based on the selection combining argument). For the second detection step, the effect of the optimal ordering is to increase the outage probability twice. This is the "price" to pay for increased SNR at the first step. However, the diversity order at the second step is n. Thus, 3-dB increase in outage probability will not degrade the overall performance since the original outage probability is low (for reasonably large SNR). On the contrary, it is important to improve the first step SNR since the diversity order is $(n-1)$, less than at the second step. Numerical Monte Carlo simulations validate the analytical results above. While Monte Carlo simulations are very lengthy for small outage probabilities, the numerical evaluation of the analytical expressions above is computationally efficient.

Based on the distributions of instantaneous SNR derived above, we further analytically analyze the average BER performance assuming noncoherent EGC and noncoherent orthogonal FSK. Rigorous and approximate (at high SNR regime) closed-form expressions are given. Similar to the case of conventional diversity combing systems, the asymptotic behavior of the average BER resembles that of the corresponding outage probability (up to a constant). Other modulation/combing format can be analyzed in a similar way (however, analytical analysis may not be always feasible, at least in a reasonably simple form).

Appendix A

A closed-form analytical model of the Gramm–Schmidt process is given in [10, Section IX.6]. Let us consider the ist processing step. Assuming that the first $(i-1)$ symbols are detected without errors and the interference cancellation is accomplished, the received vector at this step is

$$\mathbf{r}'_i = \mathbf{r} - \sum_{k=1}^{i-1} \mathbf{h}_k q_k = \sum_{k=i}^{m} \mathbf{h}_k q_k. \tag{A1}$$

The orthogonal projection of $\mathbf{r}'_i$ into the space spanned by $\{/eta_{i+1}, \ldots, /eta_m\}$ is given by [10]

$$\mathbf{r}'_{i,\|} = \frac{-1}{\left|\mathbf{R}^{[i+1,m]}\right|} \begin{vmatrix} & \mathbf{R}^{[i+1,m]} & & \boldsymbol{\eta}_{i+1} \\ & & & \cdots \\ & & & \boldsymbol{\eta}_m \\ \mathbf{r}'_i\boldsymbol{\eta}_{i+1} & \cdots & \mathbf{r}'_i\boldsymbol{\eta}_m & 0 \end{vmatrix}. \tag{A2}$$

The component of $\mathbf{r}'_i$ orthogonal to $\{\boldsymbol{\eta}_{i+1}, \ldots, \boldsymbol{\eta}_m\}$ is

$$\begin{aligned} \mathbf{r}'_{i,\perp} &= \mathbf{r}'_i - \mathbf{r}'_{i,\|} \\ &= \frac{1}{\left|\mathbf{R}^{[i+1,m]}\right|} \begin{vmatrix} & \mathbf{R}^{[i+1,m]} & & \boldsymbol{\eta}_{i+1} \\ & & & \cdots \\ & & & \boldsymbol{\eta}_m \\ \mathbf{r}'_i\boldsymbol{\eta}_{i+1} & \cdots & \mathbf{r}'_i\boldsymbol{\eta}_m & \mathbf{r}'_i \end{vmatrix}. \end{aligned} \tag{A3}$$

Taking into account (A1), we note that the last row in the numerator determinant in (A3) includes components proportional to first to $(m-i)$th rows, which can be dropped out because they do no affect the determinant value. Thus, the only component of $\mathbf{r}'_i$ that gives contribution to the determinant is $\mathbf{h}_i q_i$. Consequently, (A3) reduces to

$$\mathbf{r}'_{i,\perp} = \frac{q_i|\mathbf{h}_i|}{\left|\mathbf{R}^{[i+1,m]}\right|} \begin{vmatrix} & \mathbf{R}^{[i+1,m]} & & \boldsymbol{\eta}_{i+1} \\ & & & \cdots \\ & & & \boldsymbol{\eta}_m \\ \boldsymbol{\eta}_i\boldsymbol{\eta}_{i+1} & \cdots & \boldsymbol{\eta}_i\boldsymbol{\eta}_m & \boldsymbol{\eta}_i \end{vmatrix}. \tag{A4}$$

Relocating the last row to the top position and the right column to the left position, one obtains

$$\mathbf{r}'_{i,\perp} = \frac{q_i|\mathbf{h}_i|}{\left|\mathbf{R}^{[i+1,m]}\right|} \begin{vmatrix} \boldsymbol{\eta}_i & \boldsymbol{\eta}_i\boldsymbol{\eta}_{i+1} & \cdots & \boldsymbol{\eta}_i\boldsymbol{\eta}_m \\ \boldsymbol{\eta}_{i+1} & & & \\ \cdots & & \mathbf{R}^{[i+1,m]} & \\ \boldsymbol{\eta}_m & & & \end{vmatrix} \tag{A5}$$

which is the same as (4). The signal power is:

$$|\mathbf{r}'_{i,\perp}|^2 = \mathbf{r}'_{i,\perp}\mathbf{r}'_i. \tag{A6}$$

Substituting (A6) into (A5), we obtain

$$|\mathbf{r}'_{i,\perp}|^2 = \frac{q_i|\mathbf{h}_i|}{\left|\mathbf{R}^{[i+1,m]}\right|}\begin{vmatrix} \boldsymbol{\eta}_i\mathbf{r}'_i & \boldsymbol{\eta}_i\boldsymbol{\eta}_{i+1} & \cdots & \boldsymbol{\eta}_i\boldsymbol{\eta}_m \\ \boldsymbol{\eta}_{i+1}\mathbf{r}'_i & & & \\ \cdots & & \mathbf{R}^{[i+1,m]} & \\ \boldsymbol{\eta}_m\mathbf{r}'_i & & & \end{vmatrix}. \tag{A7}$$

We note that the first column in the numerator determinant in (A7) includes components proportional to second to $(m-i+1)$th columns, which can be dropped out. Consequently, (A7) reduces to

$$\begin{aligned}|\mathbf{r}'_{i,\perp}|^2 &= \frac{|q_i|^2|\mathbf{h}_i|^2}{\left|\mathbf{R}^{[i+1,m]}\right|}\begin{vmatrix} 1 & \boldsymbol{\eta}_i\boldsymbol{\eta}_{i+1} & \cdots & \boldsymbol{\eta}_i\boldsymbol{\eta}_m \\ \boldsymbol{\eta}_{i+1}\boldsymbol{\eta}_i & & & \\ \cdots & & \mathbf{R}^{[i+1,m]} & \\ \boldsymbol{\eta}_m\boldsymbol{\eta}_i & & & \end{vmatrix} \\ &= |q_i|^2|\mathbf{h}_i|^2\frac{\left|\mathbf{R}^{[i,m]}\right|}{\left|\mathbf{R}^{[i+1,m]}\right|}.\end{aligned} \tag{A8}$$

Similar to (A3) and (A5), the noise vector after the interference nulling out can be presented as

$$\mathbf{v}''_i = \frac{1}{\left|\mathbf{R}^{[i+1,m]}\right|}\begin{vmatrix} \mathbf{v} & \mathbf{v}\boldsymbol{\eta}_{i+1} & \cdots & \mathbf{v}\boldsymbol{\eta}_m \\ \boldsymbol{\eta}_{i+1} & & & \\ \cdots & & \mathbf{R}^{[i+1,m]} & \\ \boldsymbol{\eta}_m & & & \end{vmatrix} \tag{A9}$$

which is equivalent to (11). The instantaneous noise power, $|\mathbf{v}''_i|^2$, is

$$\begin{aligned}|\mathbf{v}''_i|^2 &= \mathbf{v}\mathbf{v}''_i \\ &= \frac{1}{\left|\mathbf{R}^{[i+1,m]}\right|}\begin{vmatrix} |\mathbf{v}|^2 & \mathbf{v}\boldsymbol{\eta}_{i+1} & \cdots & \mathbf{v}\boldsymbol{\eta}_m \\ \mathbf{v}\boldsymbol{\eta}_{i+1} & & & \\ \cdots & & \mathbf{R}^{[i+1,m]} & \\ \mathbf{v}\boldsymbol{\eta}_m & & & \end{vmatrix} \\ &= |\mathbf{v}|^2 - \frac{\mathbf{v}\boldsymbol{\eta}_{i+1}}{\left|\mathbf{R}^{[i+1,m]}\right|} \\ &\quad\times\begin{vmatrix} \mathbf{v}\boldsymbol{\eta}_{i+1} & \mathbf{v}\boldsymbol{\eta}_{i+2} & \cdots & \mathbf{v}\boldsymbol{\eta}_m \\ \boldsymbol{\eta}_{i+2}\boldsymbol{\eta}_{i+1} & 1 & \cdots & \boldsymbol{\eta}_{i+2}\boldsymbol{\eta}_m \\ \cdots & \cdots & \cdots & \cdots \\ \boldsymbol{\eta}_m\boldsymbol{\eta}_{i+1} & \boldsymbol{\eta}_m\boldsymbol{\eta}_{i+2} & \cdots & 1 \end{vmatrix} \\ &\quad+\cdots+\frac{(-1)^{m-i}\mathbf{v}\boldsymbol{\eta}_m}{\left|\mathbf{R}^{[i+1,m]}\right|} \\ &\quad\times\begin{vmatrix} \mathbf{v}\boldsymbol{\eta}_{i+1} & \mathbf{v}\boldsymbol{\eta}_{i+2} & \cdots & \mathbf{v}\boldsymbol{\eta}_m \\ 1 & \boldsymbol{\eta}_{i+1}\boldsymbol{\eta}_{i+2} & \cdots & \boldsymbol{\eta}_{i+1}\boldsymbol{\eta}_m \\ \cdots & \cdots & \cdots & \cdots \\ \boldsymbol{\eta}_{m-1}\boldsymbol{\eta}_{i+1} & \boldsymbol{\eta}_{m-1}\boldsymbol{\eta}_{i+2} & \cdots & \boldsymbol{\eta}_{m-1}\boldsymbol{\eta}_m \end{vmatrix}\end{aligned} \tag{A10}$$

where we expanded the determinant along the first column. Using the following identity

$$\langle(\mathbf{v}\boldsymbol{\eta}_i)(\mathbf{v}\boldsymbol{\eta}_j)\rangle = \sigma_1^2\boldsymbol{\eta}_i\boldsymbol{\eta}_j \tag{A11}$$

where the expectation is taken over noise, and $\sigma_1^2 = \langle|\mathbf{v_j}|^2\rangle$ is per-branch noise power before processing (note that we assume that all branches have equal and uncorrelated noise), we obtain the following:

$$\begin{aligned}P_{\mathbf{v}_i} &= \langle|\mathbf{v}''_\mathbf{i}|^2\rangle \\ &= n\sigma_1^2 - \frac{\sigma_1^2}{\left|\mathbf{R}^{[i+1,m]}\right|} \\ &\quad\times\begin{vmatrix} 1 & \boldsymbol{\eta}_{i+1}\boldsymbol{\eta}_{i+2} & \cdots & \boldsymbol{\eta}_{i+1}\boldsymbol{\eta}_m \\ \boldsymbol{\eta}_{i+2}\boldsymbol{\eta}_{i+1} & 1 & \cdots & \boldsymbol{\eta}_{i+2}\boldsymbol{\eta}_m \\ \cdots & \cdots & \cdots & \cdots \\ \boldsymbol{\eta}_m\boldsymbol{\eta}_{i+1} & \boldsymbol{\eta}_m\boldsymbol{\eta}_{i+2} & \cdots & 1 \end{vmatrix} \\ &\quad+\cdots+\frac{(-1)^{m-i}\sigma_1^2}{\left|\mathbf{R}^{[i+1,m]}\right|} \\ &\quad\times\begin{vmatrix} \boldsymbol{\eta}_m\boldsymbol{\eta}_{i+1} & \boldsymbol{\eta}_m\boldsymbol{\eta}_{i+2} & \cdots & 1 \\ 1 & \boldsymbol{\eta}_{i+1}\boldsymbol{\eta}_{i+2} & \cdots & \boldsymbol{\eta}_{i+1}\boldsymbol{\eta}_m \\ \cdots & \cdots & \cdots & \cdots \\ \boldsymbol{\eta}_{m-1}\boldsymbol{\eta}_{i+1} & \boldsymbol{\eta}_{m-1}\boldsymbol{\eta}_{i+2} & \cdots & \boldsymbol{\eta}_{m-1}\boldsymbol{\eta}_m \end{vmatrix} \\ &= n\sigma_1^2 - \sigma_1^2 - \cdots - \sigma_1^2 = (n-m+i)\sigma_1^2.\end{aligned} \tag{A12}$$

In order to obtain the last two equalities in (A12), we exchanged the position of some rows in the determinants to get $\mathbf{R}^{[i+1,m]}$ in each term of the sum. This proves (12).

APPENDIX B

To find the pdf of φ, we follow the same approach as in [12, p. 191]. The variable $u = \tan\varphi = |\mathbf{h}_{1\perp}|/|\mathbf{h}_{1||}|$ has the following distribution [12, p. 237]:

$$f_u(u) = \frac{2u}{(1+u^2)^2}, \quad u \geq 0. \tag{B1}$$

Hence, the pdf of φ is

$$f_\varphi(\varphi) = \frac{1}{|d\varphi/du|}f_u(\tan\varphi) = \sin 2\varphi, \quad \varphi \in [0, \pi/2]. \tag{B2}$$

To prove (22), we start with (21)

$$\Pr[s_1 < x] = F_1(x) = \int_0^{\pi/2} F_h^2\left(\frac{x}{\sin^2\varphi}\right) f_\varphi(\varphi)\, d\varphi. \tag{B3}$$

First, we transform the integral in (B3) to the following form:

$$F_1(x) = \int_0^1 F_h^2\left(\frac{x}{t}\right)dt = \int_1^\infty \frac{F_h^2(xt)}{t^2}dt. \tag{B4}$$

The transformations are performed using sequentially the substitutions $\sin^2\varphi \to t$ and $t \to 1/t$. Further, using (B2) and (19), the last integral in (B4) is transformed to

$$\begin{aligned}F_1(x) = 1 - 2E_2(x) + E_2(2x) \\ - 2xE_1(x) + 2xE_1(2x) + xe^{-x}\end{aligned} \tag{B5}$$

where

$$E_k(x) = \int_1^\infty \frac{e^{-xt}}{t^k}dt \tag{B6}$$

is the integral exponential function [13]. We further use the following recursive rule:

$$E_{k+1}(x)=\frac{1}{k}\left(e^{-x}-xE_k(x)\right),\quad k=1,2,3,\ldots \tag{B7}$$

to express E_2 using E_1 and obtain

$$F_1(x)=1-2e^{-x}+\left(1+\frac{x}{2}\right)e^{-2x} \tag{B8}$$

which is the same as (22).

The general $2\times n$ case is considered in a similar way. The pdf of the variable u can be shown to be the following [12, p. 237]:

$$f_u(u)=\frac{2(n-1)u^{2n-3}}{(1+u^2)^n},\quad u\geq 0 \tag{B9}$$

and, hence, the pdf of φ is

$$\begin{aligned} f_\varphi(\varphi)&=\frac{1}{|d\varphi/du|}f_u(\tan\varphi)\\ &=2(n-1)\sin^{2n-3}\varphi\cdot\cos\varphi,\quad \varphi\in[0,\pi/2]. \end{aligned} \tag{B10}$$

To prove (30), we start with (29)

$$F_1(x)=(n-1)\int_0^1 F_h^2\left(\frac{x}{t}\right)\cdot t^{n-2}dt. \tag{B11}$$

Using (28), we transform it to

$$F_1(x)=1-2(n-1)I_2+(n-1)I_3 \tag{B12}$$

where

$$I_2=\sum_{m=0}^{n-1}\frac{x^m}{m!}\int_0^1 e^{\frac{-x}{t}}t^{n-2-m}dt=\sum_{m=0}^{n-1}\frac{x^m}{m!}E_{n-m}(x) \tag{B13}$$

$$\begin{aligned} I_3&=\sum_{k=0}^{n-1}\sum_{m=0}^{n-1}\frac{x^{k+m}}{k!m!}\int_0^1 e^{\frac{-2x}{t}}t^{n-2-k-m}dt\\ &=\sum_{k=0}^{n-1}\sum_{m=0}^{n-1}\frac{x^{k+m}}{k!m!}E_{n-k-m}(2x). \end{aligned} \tag{B14}$$

Using the recursive rule (B7), we obtain

$$\begin{aligned} E_{k+1}(x)=e^{-x}\sum_{i=0}^{k-1}\frac{(-x)^i(k-i-1)!}{k!}\\ +\frac{(-x)^k}{k!}E_1(x),\quad k=1,2,3,\ldots. \end{aligned} \tag{B15}$$

When $n-k-m<1$ in (B14), we use the following relation [13]:

$$\begin{aligned} E_{-k}(x)=\alpha_k(x)&=\int_1^\infty t^k e^{-xt}dt\\ &=k!x^{-k-1}e^{-x}\sum_{i=0}^{k}\frac{x^i}{i!},\quad k=1,2,3,\ldots. \end{aligned} \tag{B16}$$

Combining (B15) and (B13), one obtains

$$\begin{aligned} I_2=e^{-x}\sum_{m=0}^{n-1}\sum_{k=0}^{n-m-2}\frac{(-x)^k(n-m-k-2)!}{(n-m-1)!}\frac{x^m}{m!}\\ +E_1(x)x^{n-1}\sum_{m=0}^{n-1}\frac{(-1)^{n-m-1}}{(n-m-1)!m!}. \end{aligned} \tag{B17}$$

It is straightforward to show that the last sum in (B17) equals to zero. Hence

$$I_2=e^{-x}\sum_{i=0}^{n-2}a_ix^i,\quad a_i=\sum_{m=0}^{i}\frac{(-1)^{i-m}(n-i-2)!}{(n-m-1)!m!} \tag{B18}$$

which is in agreement with (30) and (31).

We further present I_3 as follows:

$$\begin{aligned} I_3&=I_{31}+I_{32},\quad I_{31}=\sum_{i=0}^{n-1}d_ix^iE_{n-i}(2x), I_{32}\\ &=\sum_{i=n}^{2n-2}d_ix^i\alpha_{n-i}(2x),\\ d_i&=\begin{cases}\dfrac{2^i}{i!}, & i<n\\ \displaystyle\sum_{k=i-n+1}^{n-1}\frac{1}{k!(i-k)!}, & i\geq n\end{cases}. \end{aligned} \tag{B19}$$

Using (B15), one obtains

$$\begin{aligned} I_{31}&=e^{-2x}\sum_{m=0}^{n-1}\sum_{k=0}^{n-m-2}b_{mk}x^{m+k}\\ &\quad+E_1(2x)(-2x)^{n-1}\sum_{m=0}^{n-1}\frac{(-2)^{-m}d_m}{(n-m-1)!}\\ b_{mk}&=\frac{(-1)^k2^{k+m}(n-k-m-2)!}{(n-m-1)!m!}. \end{aligned} \tag{B20}$$

It is straightforward to show that the second sum in (B20) equals to zero. Hence, (B20) reduces to

$$\begin{aligned} I_{31}&=e^{-2x}\sum_{m=0}^{n-2}b_mx^m,\\ b_m&=\sum_{k=0}^{m}b_{k,m-k}\\ &=(-2)^m(n-m-2)!\sum_{k=0}^{m}\frac{(-1)^k}{(n-k-1)!k!}. \end{aligned} \tag{B21}$$

In a similar way, combining (B16) and (B19), one obtains

$$\begin{aligned} I_{32}&=e^{-2x}(2x)^{n-1}\sum_{m=0}^{n-2}c_mx^m,\\ c_m&=\frac{2^m}{m!}\sum_{j=m+n}^{2n-2}\sum_{k=j-n+1}^{n-1}\frac{(j-n)!2^{-j}}{(j-k)!k!}. \end{aligned} \tag{B22}$$

Combining (B18), (B21), and (B22), one obtains (30) and (31). This concludes the proof.

REFERENCES

[1] G. J. Foschini and M. J. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," *Wireless Personal Commun.*, vol. 6, no. 3, pp. 311–335, Mar. 1998.

[2] G. J. Foschini, "Layered space-time architecture for wireless communication in a fading environment when using multiple antennas," *Bell Lab. Tech. J.*, vol. 1, no. 2, pp. 41–59, 1996.

[3] I. E. Telatar, "Capacity of multi-antenna Gaussian channels," *Eur. Trans. Telecom., v.10, N.6, Dec. 1999*, June 1995.

[4] G. G. Rayleigh and J. M. Cioffi, "Spatio-temporal coding for wireless communications," *IEEE Trans. Commun.*, vol. 46, pp. 357–366, Mar. 1998.

[5] G. D. Golden, G. J. Foschini, R. A. Valenzuela, and P. W. Wolniansky, "Detection algorithm and initial laboratory results using V-BLAST space-time communication architecture," *Electron. Lett.*, vol. 35, no. 1, pp. 14–16, Jan. 7, 1999.

[6] S. L. Loyka and J. R. Mosig, "Channel capacity of N-antenna BLAST architecture," *Electron. Lett.*, vol. 36, no. 7, pp. 660–661, Mar. 2000.

[7] S. L. Loyka, "Channel capacity of MIMO architecture using the exponential correlation matrix," *IEEE Commun. Lett.*, vol. 5, pp. 369–371, Sept. 2001.

[8] S. Loyka and G. Tsoulos, "Estimating MIMO system performance using the correlation matrix approach," *IEEE Commun. Lett.*, vol. 6, pp. 19–21, Jan. 2002.

[9] G. J. Foschini *et al.*, "Simplified processing for high spectral efficiency wireless communication employing multi-element arrays," *IEEE J. Select. Areas Commun.*, vol. 17, pp. 1841–1852, Nov. 1999.

[10] F. R. Gantmaher, *The Theory of Matrices*. New York: Chelsea Pub. Co., 1959.

[11] S. L. Loyka, "Channel capacity of two-antenna BLAST architecture," *Electron. Lett.*, vol. 35, no. 17, pp. 1421–1422, Aug. 19, 1999.

[12] A. Papoulis and S. U. Pillai, *Probability, Random Variables, and Stochastic Processes*. Boston, MA: McGraw Hill, 2002.

[13] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions*. New York: Dover, 1974.

[14] J. G. Proakis, *Digital Communications*. Boston, MA: McGraw Hill, 2001.

Sergey Loyka was born in Minsk, Belarus. He received the Ph.D. degree in radio engineering from the Belorussian State University of Informatics and Radioelectronics, Minsk, in 1995 and the M.S. degree with honors from Minsk Radioengineering Institute, Minsk, in 1992.

Since 2001, he has been a Faculty Member at the School of Information Technology and Engineering (SITE), University of Ottawa, Ottawa, ON, Canada. Prior to that, he was a Research Fellow in the Laboratory of Communications and Integrated Microelectronics (LACIME) of École de Technologie Supérieure, Montreal, QC, Canada; a Senior Scientist at the Electromagnetic Compatibility Laboratory, Belorussian State University of Informatics and Radioelectronics; an Invited Scientist at the Laboratory of Electromagnetism and Acoustic, Swiss Federal Institute of Technology, Lausanne, Switzerland. His research areas include wireless communications, smart antennas, RF/microwave circuit and system modeling and simulation, and electromagnetic compatibility.

Mr. Loyka received a number of awards from the URSI, the IEEE, the Swiss, Belarus and former USSR governments, and the Soros Foundation.

Francois Gagnon (S'87–M'87–SM'99) was born in Québec City, QC, Canada. He received the B.Eng. and Ph.D. degrees in electrical engineering from École Polytechnique de Montreal, Montreal, QC, Canada.

Since 1991, he has been a Professor in the Department of Electrical Engineering, École de Technologie Supérieure, Montreal. He has chaired the department from 1999 to 2001 and is now the holder of the Ultra Electronics (TCS) Chair in Wireless Communication at the same university. His research interest covers wireless high-speed communications, modulation, coding, high-speed DSP implementations, and military point-to-point communications. He has been very involved in the creation of the new generation of High-Capacity Line-Of-Sight military radios offered by the Canadian Marconi Corporation, which is now Ultra Electronics Tactical Communication Systems.