Optimum Power and Rate Allocation for Coded V-BLAST: Instantaneous Optimization

Victoria Kostina and Sergey Loyka

***Abstract*—Several instantaneous optimization strategies for rate and/or power allocation in the coded V-BLAST are studied analytically. Outage probabilities and system capacities of these strategies in a spatial multiplexing system are compared under generic settings. The conventional waterfilling algorithm is shown to be suboptimal for the coded V-BLAST and a new algorithm ("fractional water-filling") is proposed, which simultaneously maximizes the system capacity and minimizes the outage probability. Closed-form performance analysis of the considered algorithms is given, and the fractional water-filling algorithm is shown to attain the full MIMO channel diversity, significantly outperforming other strategies. Many of the results also apply to generic multi-stream transmission systems (e.g. spatial multiplexing on the channel eigenmodes, OFDM) or the systems relying on successive interference cancelation (multi-user detection, channel equalization).**

***Index Terms*—Multi-antenna (MIMO) system, spatial multiplexing, coded V-BLAST, power/rate allocation, waterfilling, performance analysis.**

I. INTRODUCTION

IN the first part of this paper [1], we have developed optimum allocations of average power and/or rate to improve the performance of coded V-BLAST, and analyzed the performance improvement. The average optimization, which is based on the channel statistics and the average SNR, is motivated by lower implementation complexity and smaller demand on system resources (i.e. average rather than instantaneous feedback). However, its performance is in general inferior to instantaneous optimization, which is performed for each channel realization. Therefore, in this part we consider an instantaneous optimization of power and rate allocation, develop closed-form analytical solutions and present their performance analysis. The emphasis is on analytical tools and solutions rather than on numerical algorithms and simulations.

Several techniques have been reported to improve the error performance of the uncoded V-BLAST by employing a non-uniform power and/or rate allocation among the transmitters (Tx) [2]-[6]. Uncoded systems, however, are rare and most practical systems are coded. This motivates the study of coded V-BLAST. While the error rate analysis of coded systems is a formidable task hardly possible in a closed-from (except for some special cases) [7], the analysis becomes feasible when powerful capacity-approaching codes are used (e.g. LDPC, turbo-codes or polar codes). Following this philosophy advocated in [8][9], we will assume in this paper that capacity-achieving temporal codes[1] are used for each stream in the V-BLAST, so that the per-stream rates are equal to the corresponding capacities and there are no errors when streams are not in outage, and also no error propagation in-between the streams.

This approach has been used by Zhang et al. [11], who considered an instantaneous optimization of power, rate and antenna mapping for the coded ZF V-BLAST to minimize the total transmit power for given data rate under zero outage constraint, assuming capacity-achieving temporal codes or realistic ones via an SNR gap to capacity. The optimization in [11] is carried out under zero outage constraint, which requires unlimited power investment into particularly bad channel realizations (to support the target rate) and is not feasible when the peak power is constrained (i.e. by an RF power amplifier). Unlike [11], we allow non-zero outage probability and minimize it by proper power/rate allocation, under the total instantaneous power constraint, which automatically constraints the peak power as well. While [11] makes use of the conventional waterfilling (WF) algorithm to assign powers and rates, we show that this algorithm does not achieve the maximum V-BLAST system capacity[2] (because of the successive interference cancelation (SIC)) and propose a new algorithm termed "fractional waterfilling" (FWF) that does maximize the system capacity. While the complexity of the proposed FWF is higher compared to the conventional WF, its incremental complexity is small when the number of transmitters is not too large, as in realistic MIMO systems. It is shown that the FWF is superior to the conventional WF in terms of the outage probability; for a given channel realization, the FWF converges to the conventional WF at high SNR in full-rank channels, and is significantly superior at low SNR.

While efficient numerical algorithms developed in [11] are naturally welcome from the practical perspective, they offer only limited insight and require numerical simulations to evaluate the performance. Therefore, we concentrate on analytical development and performance evaluation in this paper. To the best of our knowledge, it is the first time when the outage performance of a waterfilling algorithm is presented in a closed form.

Paper approved by N. Al-Dhahir, the Editor for Space-Time, OFDM and Equalization of the IEEE Communications Society. Manuscript received August 9, 2010, revised March 19, 2011.

The paper was presented in part at the IEEE International Conference on Communications, Dresden, Germany, June 2009 [18].

V. Kostina is with the Department of Electrical Engineering, Princeton University, Princeton, NJ 08544 USA (e-mail: vkostina@princeton.edu).

S. Loyka is with the School of Electrical Engineering and Computer Science, University of Ottawa, Ontario, Canada, K1N 6N5 (e-mail: sergey.loyka@ieee.org).

Digital Object Identifier 10.1109/TCOMM.2011.071111.100485

[1]this can also be extended to realistic codes by using the SNR gap to capacity, as in [10][11].

[2]The system capacity is the capacity of an extended channel, which includes the channel itself and also the V-BLAST transmission/processing architecture.

We consider instantaneous (per-stream) rate allocation (IRA), instantaneous (per-stream) power allocation (IPA), and joint instantaneous power/rate allocation (IPRA) to minimize the outage probability or to maximize the system capacity of the coded V-BLAST under the constrained total transmit power and a given target rate. Since the total transmit power is limited, the peak power of RF amplifiers is limited as well. The main contributions of the paper are as follows:

• A comparison of various power/rate optimization strategies for a generic spatial multiplexing system (not only V-BLAST) is given in section III. In particular, the maximization of the instantaneous system capacity (via the IPRA) is shown to also minimize the outage probability and, thus, the two problems are equivalent, under arbitrary fading distribution. The importance of this is due to the fact that the minimization of the outage probability is a challenging, non-convex problem with multiple solutions, while the maximization of the instantaneous capacity is an easier convex problem with a unique solution. While there is a number of strategies to minimize the outage probability, only the FWF algorithm simultaneously minimizes the outage probability and maximizes the system capacity of the coded V-BLAST.

• An optimum instantaneous power allocation to maximize the system capacity of a multi-stream transmission (e.g. V-BLAST, spatial multiplexing on channel eigenmodes, OFDM) under uniform rate allocation is presented in Section V.

• The conventional WF algorithm is shown to be suboptimal for the V-BLAST. To maximize the V-BLAST capacity via the IPRA, the fractional waterfilling algorithm is proposed and its key properties are analyzed (Theorem 4, Corollaries 4-6 in section VI). The FWF is proven to give a solution of the (non-convex) optimization problem of minimizing the outage probability subject to the total power constraint.

• Performance analysis of the IRA, the IPA, the WF and the FWF is presented in section VII. In the low outage regime, the WF achieves the same diversity gain as the IRA, while the FWF brings in an additional diversity gain, achieving the full MIMO channel diversity (Corollary 9) and attaining simultaneously the minimum possible outage probability and the maximum system capacity under the total power constraint. In the low rate (also known as wideband) regime, the outage probabilities of the instantaneous optimization strategies are found in explicit, closed form (Theorem 5).

• Section VIII gives some examples to compare the average and instantaneous optimization, to demonstrate the superiority of the FWF and to validate the analytical results and conclusions via simulations.

Many of these results also apply to generic multi-stream transmission systems (e.g. spatial multiplexing on the channel eigenmodes, OFDM) and also to multiuser detection and inter-symbols equalization systems that use successive interference cancelation.

II. System Model

Motivated by lower complexity and to make the analysis feasible, we consider unordered ZF V-BLAST. The basic system model follows that in [1] and is summarized below for completeness.

The standard baseband discrete-time MIMO system model is,

$$\mathbf{r} = \mathbf{H}\boldsymbol{\Lambda}\mathbf{s} + \boldsymbol{\xi} = \sum\nolimits_{i=1}^{m} \mathbf{h}_i \sqrt{\alpha_i} s_i + \boldsymbol{\xi} \qquad (1)$$

where $\mathbf{s} = [s_1, s_2, ...s_m]^T$ and $\mathbf{r} = [r_1, r_2, ...r_n]^T$ are the vectors representing the Tx and Rx symbols respectively, "T" denotes transposition, $\mathbf{H} = [\mathbf{h}_1, \mathbf{h}_2, ...\mathbf{h}_m]$ is the $n \times m$ matrix of the complex channel gains between each Tx and each Rx antenna, where $\mathbf{h}_i$ denotes i-th column of $\mathbf{H}$, n and m are the numbers of Rx and Tx antennas respectively, $n \geq m$, $\boldsymbol{\xi}$ is the vector of circularly-symmetric additive white Gaussian noise (AWGN), which is independent and identically distributed (i.i.d.) in each receiver, $\boldsymbol{\Lambda} = diag\left(\sqrt{\alpha_1}, \ldots, \sqrt{\alpha_m}\right)$, where α_i is the power allocated to the i-th transmitter. For the regular V-BLAST, the total power is distributed uniformly among the transmitters, $\alpha_1 = \alpha_2 = ... = \alpha_m = 1$. The channel will be assumed to be either ergodic ("fast fading"), in which case the key performance measure is the ergodic system capacity, or non-ergodic ("slow fading"), in which case the key performance measures are the outage probability and the outage capacity and also the instantaneous system capacity (for given channel realization) [8]. Details of a mathematical model of the uncoded V-BLAST, on which our model of the coded V-BLAST is based, and its analysis can be found in [2][12][6] and are omitted here.

III. Instantaneous vs. Average Optimization

Let us consider a generic multi-stream transmission system (e.g. spatial multiplexing, OFDM, multi-user) operating in a fading channel of generic statistics (not only i.i.d. Rayleigh), which is quasi-static (non-ergodic or "slow fading"). The main performance indicator in this setting is the system outage probability [3], i.e. the probability that the system cannot support a target total rate mR [8],

$$P_{out} = \Pr\{C < mR\} \qquad (2)$$

where C is the instantaneous (i.e. for given channel realization) system capacity (i.e. the sum of per-stream capacities), and an optimization strategy should target this measure. On the other hand, when the channel is ergodic, the mean (ergodic) capacity $\overline{C}$ is an appropriate performance measure and its optimization is of interest. In both scenarios, the optimization can be instantaneous (i.e. for each channel realization) or average (i.e. based on the channel statistics), and may include per-stream power, rate or joint power/rate allocation.

In this section, we compare the performance of four different optimization strategies:

• average power and/or rate allocation to maximize the mean capacity,

$$\overline{\boldsymbol{\alpha}}_C = \arg\max_{\boldsymbol{\alpha}(\gamma_0)} \overline{C}(\boldsymbol{\alpha}), \qquad (3)$$

• average power and/or rate allocation to minimize the outage probability,

$$\overline{\boldsymbol{\alpha}}_{out} = \arg\min_{\boldsymbol{\alpha}(\gamma_0)} P_{out}(\boldsymbol{\alpha}), \qquad (4)$$

[3] which is also the best achievable codeword error probability for sufficiently long codewords [16][17].

• instantaneous power and/or rate allocation to maximize the instantaneous system capacity (this will automatically maximize the ergodic capacity as well),

$$\boldsymbol{\alpha}_C = \arg\max_{\boldsymbol{\alpha}(\gamma_0,\mathbf{H})} C(\boldsymbol{\alpha}), \tag{5}$$

• instantaneous power and/or rate allocation to minimize the outage probability,

$$\boldsymbol{\alpha}_{out} = \arg\min_{\boldsymbol{\alpha}(\gamma_0,\mathbf{H})} P_{out}(\boldsymbol{\alpha}), \tag{6}$$

and C, $\overline{C}$ and P_{out} are considered as functions of the power and/or rate allocation $\boldsymbol{\alpha}$ [4], all subject to the total power $\sum_{i=1}^{m} \alpha_i = m$ constraint, where $\boldsymbol{\alpha}(\gamma_0)$ and $\boldsymbol{\alpha}(\gamma_0, \mathbf{H})$ indicate average and instantaneous optimizations respectively, and γ_0 is the average SNR.

The following Lemma will be instrumental.

Lemma 1. *Consider two optimization strategies $\boldsymbol{\alpha}^1$ and $\boldsymbol{\alpha}^2$ (power and/or rate) such that $C^1 = C(\boldsymbol{\alpha}^1) \geq C(\boldsymbol{\alpha}^2) = C^2\ \forall \mathbf{H}$. Then the corresponding outage probabilities are also ordered likewise,*

$$P_{out}^1 = \Pr\{C^1 < mR\} \leq \Pr\{C^2 < mR\} = P_{out}^2 \tag{7}$$

Proof: Define the outage sets $\mathcal{O}^i = \{\mathbf{H} : C^i < R\},\ i = 1,2$. The outage probabilities can then be expressed as $P_{out}^i = \Pr\{\mathbf{H} \in \mathcal{O}^i\}$. From $C^1 \geq C^2$, it follows that $\mathcal{O}^1 \subseteq \mathcal{O}^2$, and thus $P_{out}^1 \leq P_{out}^2$. ■

We are now in a position to compare the optimization strategies in (3)-(6).

Theorem 1. *The outage probabilities of the optimization strategies in* (3)-(6) *are ordered as follows,*

$$\begin{aligned} \Pr\{C(\overline{\boldsymbol{\alpha}}_C) < mR\} &\geq \Pr\{C(\overline{\boldsymbol{\alpha}}_{out}) < mR\} && (8)\\ &\geq \Pr\{C(\boldsymbol{\alpha}_{out}) < mR\} = P_{out}^{\star} && (9)\\ &= \Pr\{C(\boldsymbol{\alpha}_C) < mR\} && (10) \end{aligned}$$

i.e. the instantaneous optimizations of the capacity and outage probability achieve the same lowest outage probability $P_{out}^{\star}$, the average optimization of the outage probability gives an intermediate result, and the average optimization of the ergodic capacity is the worst.

Proof: The inequality in (8) is by the definition of $\overline{\boldsymbol{\alpha}}_{out}$ (i.e. $\overline{\boldsymbol{\alpha}}_{out}$ is the best average power/rate allocation that minimizes P_{out}). The inequality in (9) is because the instantaneous optimization of P_{out} cannot be worse than the average one. To prove the equality in (10) note that $P_{out}^{\star} \leq \Pr\{C(\boldsymbol{\alpha}_C) < mR\}$ (by the definition of $\boldsymbol{\alpha}_{out}$) and also that $C(\boldsymbol{\alpha}_C) \geq C(\boldsymbol{\alpha}_{out})$ (by the definition of $\boldsymbol{\alpha}_C$). Using the last inequality and Lemma 1, $P_{out}^{\star} \geq \Pr\{C(\boldsymbol{\alpha}_C) < mR\}$. Combining this with the opposite inequality above, (10) follows. It can be shown (by examples) that none of the inequalities in Theorem 1 can be strengthened to equalities, in general. ■

Another important performance metric is the outage capacity C_ϵ defined as the maximum rate supported by the system with constrained outage probability [8], $C_\epsilon = \max_R \{P_{out}(R) \leq \epsilon\}$. The following corollary is immediate from Theorem 1.

[4] to simplify the notations, rate allocation is also included in $\boldsymbol{\alpha}$ here.

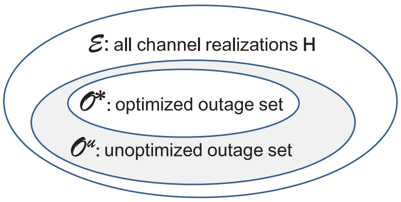


Fig. 1. Possible channel realizations are divided into different sets for optimization of the outage probability: if $\mathbf{H}$ falls into the no-outage set $\{\mathcal{E} - \mathcal{O}^u\}$, no optimization is necessary; the goal of any optimization strategy is to shrink the unoptimized outage set $\mathcal{O}^u$.

Corollary 1. *The outage capacities C_ϵ of the optimization strategies above are ordered as follows:*

$$C_\epsilon(\overline{\boldsymbol{\alpha}}_C) \leq C_\epsilon(\overline{\boldsymbol{\alpha}}_{out}) \leq C_\epsilon(\boldsymbol{\alpha}_C) = C_\epsilon(\boldsymbol{\alpha}_{out})$$

As a side remark, we note that even though $C_\epsilon(\boldsymbol{\alpha}_C) = C_\epsilon(\boldsymbol{\alpha}_{out})$, the corresponding ergodic capacities are not necessarily the same: while $\boldsymbol{\alpha}_C$ maximizes the ergodic capacity, $\boldsymbol{\alpha}_{out}$ does not, in general (see Corollary 2). The importance of (10) in Theorem 1 is due to the fact that while the problem in (6) is non-convex (as we show below, it has multiple solutions) and very difficult to deal with in general, either numerically or analytically, the problem in (5) has a well-known solution via the waterfilling (when no successive interference cancelation is used at the receiver) or via the fractional waterfilling for the coded V-BLAST (see Section VI). Since the two problems in (5) and (6) are equivalent (achieve the same minimum P_{out}), the solution of (5) also applies to (6). Let us now consider the problem in (6) in more detail.

Proposition 1. *Instantaneous optimization of the outage probability in* (6) *is in general a non-convex problem with infinite number of solutions, one of which is the solution to the problem in* (5)*, and all of them achieve the same $P_{out}^{\star}$.*

Proof: Let $C(\boldsymbol{\alpha})$ be the instantaneous capacity for given power allocation $\boldsymbol{\alpha}$ and $\mathcal{O}(\boldsymbol{\alpha}) = \{\mathbf{H} : C(\boldsymbol{\alpha}) < mR\}$ be the corresponding outage set. To minimize $P_{out} = \Pr\{\mathbf{H} \in \mathcal{O}(\boldsymbol{\alpha})\}$ is to minimize the set $\mathcal{O}(\boldsymbol{\alpha})$, which is obviously accomplished by maximizing $C(\boldsymbol{\alpha})$ for each $\mathbf{H}$, i.e. via (5). To demonstrate that this is not the only solution, we note that no optimization is necessary for all $\mathbf{H}$ such that $C(1, ..., 1) \geq mR$ (i.e. if the unoptimized instantaneous capacity is not less than the target rate mR) since such optimization, while increasing the capacity, will not shrink the outage set and, thus, will not reduce the outage probability (see Fig. 1). Thus, any power allocation can be used in such a case provided that the resulting capacity does not drop below mR. ■

The Corollary below gives some examples.

Corollary 2. *Examples of several strategies to minimize P_{out}:*

1) *$\boldsymbol{\alpha} = \boldsymbol{\alpha}_C$ in* (5) *for any $\mathbf{H}$ optimizes both C and P_{out}.*
2) *$\boldsymbol{\alpha} = \boldsymbol{\alpha}_C$ in* (5) *for $\mathbf{H} \in \mathcal{O}^u$, where $\mathcal{O}^u = \{\mathbf{H} : C(1, ..., 1) < mR\}$ is the unoptimized outage set, and uniform otherwise, optimizes P_{out} but not necessarily C.*
3) *In an iterative numerical algorithm to find $\boldsymbol{\alpha}$, stop optimization as soon as $C(\boldsymbol{\alpha}) \geq mR$.*

4) $\boldsymbol{\alpha} = \boldsymbol{\alpha}_C$ *for* $\mathbf{H} \in \{\mathcal{O}^u - \mathcal{O}^\star\}$, *where* $\mathcal{O}^\star = \{\mathbf{H} : C(\boldsymbol{\alpha}_C) < mR\}$ *is the optimized outage set (outage takes place even if the full optimum power allocation is performed), and uniform otherwise.*

Note that these strategies are not the same, yet they achieve the same $P^\star_{out}$, which implies that P_{out} should not be relied on as the only performance/optimization criterion of a communication system. It should be used in conjunction with other performance measures, such as C. In general, a good optimization strategy should optimize both P_{out} and C, as strategy #1 in Corollary 2.

IV. Instantaneous Rate Allocation (IRA)

Here, we consider the optimum instantaneous rate allocation (IRA) with the uniform power allocation, $\alpha_i = 1$, across all streams of the coded V-BLAST.

With the uniform rate allocation, i.e. when the per-stream target rate is R, the system outage takes place if any of the streams is not able to support this rate, i.e. if the capacity of at least one stream is lower than the target rate,

$$P^u_{out} = \Pr\{C_i < R \ \exists i\} = 1 - \prod_{i=1}^{m}(1 - \Pr\{C_i < R\}); \quad (11)$$

where $C_i = \ln(1 + g_i\gamma_0)$ is the instantaneous capacity of i-th stream in [nat/s/Hz], $g_i = |\mathbf{h}_{i\perp}|^2$ is i-th stream power gain, $\mathbf{h}_{i\perp}$ is i-th column of the channel matrix projected onto the subspace orthogonal to $\mathtt{span}\{\mathbf{h}_{i+1}, \ldots, \mathbf{h}_m\}$; we also used the fact that in the i.i.d. Rayleigh fading channels different g_i are independent of each other [2][12].

When an optimum rate allocation is employed and capacity-achieving codes are used for each stream, the per-stream rates are set equal to the corresponding per-stream instantaneous capacities C_i and the total rate equals to the system capacity $C_{IRA} = \sum_i C_i$. The system outage probability is then given by

$$P^{IRA}_{out} = \Pr\left\{\sum_i C_i < mR\right\}, \quad (12)$$

i.e. the outage takes place only when the sum capacity is below the target rate mR (compare to (11) where an outage takes place when at least one $C_i < R$).

V. Instantaneous Power Allocation (IPA)

In this section, we consider optimum allocation of instantaneous power subject to the total power constraint and when all the stream rates are the same (e.g. the same modulation/coding to simplify the system design). In this case, an achievable per-stream rate R satisfies $R \le C_i = \ln(1 + g_i\alpha_i\gamma_0)\ \forall i$ so that the system capacity for given $\boldsymbol{\alpha}$ is

$$C(\boldsymbol{\alpha}) = m \max_R \{R : R \le C_i \ \forall i\} = m \min_i \ln(1 + \alpha_i g_i \gamma_0) \quad (13)$$

and the optimization problem can be formulated as follows:

$$C_{IPA} = m \max_{\boldsymbol{\alpha}} \min_i \ln(1 + \alpha_i g_i \gamma_0), \ \text{s.t.} \sum_i \alpha_i = m, \ \alpha_i \ge 0 \quad (14)$$

The following Theorem gives its solution.

Theorem 2 (IPA)**.** *The system capacity of the IPA in* (14) *under the condition* $g_i > 0$ *is*

$$C_{IPA} = m \ln(1 + \overline{g}\gamma_0) \quad (15)$$

and 0 otherwise, where $\overline{g} = (\frac{1}{m}\sum_i g_i^{-1})^{-1}$ *is the harmonic mean per-stream gain. It is achieved by the following power allocation:*

$$\alpha_i = \overline{g}/g_i, \quad (16)$$

i.e. the channel inversion is the optimum strategy in this case.

Proof: See Appendix. ■

Corollary 3. *The power allocation in* (16) *also achieves the minimum outage probability and the maximum outage capacity under the total power constraint and uniform rate allocation.*

Proof: Follows immediately from Theorem 1 and Corollary 1. ■

Note from (15) that all per-stream rates are the same and equal to $\ln(1 + \overline{g}\gamma_0)$. It follows from (16) that weak streams get more power than strong ones, which is similar to the power allocation of the uncoded V-BLAST in [6], and just the opposite of the optimum joint power/rate allocations in the next section. The condition $g_i > 0$ can be easily ensured in practice by deactivating all zero-gain streams.

The system capacity C_{IPA} of the IPA can be compared to that of the uniform power/rate allocation, whose capacity is, from (13), $C_u = m \ln(1 + g_{min}\gamma_0)$, where $g_{min} = \min_i g_i$ is the minimum stream gain.

Proposition 2. *The IPA offers an SNR gain* G_{IPA} *over the uniform power/rate allocation, defined from* $C_{IPA}(\gamma_0) = C_u(G_{IPA}\gamma_0)$*, as follows:*

$$1 \le G_{IPA} = \frac{\overline{g}}{g_{min}} \le m \quad (17)$$

The lower bound is tight when all stream gains are about the same, and the upper bound is tight when one stream gain is significantly smaller than the others.

Proof: The equality follows from the comparison of C_{IPA} and C_u and using the definition of G_{IPA}. The inequalities follow from the fact that $g_{min} \le \overline{g} \le m g_{min}$. ■

It follows from Proposition 2 that the IPA does not provide any diversity gain over the uniform allocation, just an SNR gain.

One may also consider a problem dual to that in (14),

$$\min_{\boldsymbol{\alpha}} \sum_i \alpha_i, \text{s.t. } C(\boldsymbol{\alpha}) = C_0, \ \alpha_i \ge 0 \quad (18)$$

i.e. minimizing the total power subject to the system capacity $C(\boldsymbol{\alpha})$ being equal to a target value C_0. This problem has the same solution as that in Theorem 2.

Theorem 3. *The problem in* (18) *has the same solution as that in Theorem 2 under the condition* $C_0 = C_{IPA}$*, where* C_{IPA} *is the optimum capacity in Theorem 2, i.e. the power allocation in* (16) *also minimizes the total power needed to achieve the system capacity* C_{IPA}*.*

Proof: A standard proof is to solve (18) via the Lagrange multiplier technique similar to that in Appendix. We give below a simpler contradiction-type proof that provides an insight as to why this duality holds. Let $\boldsymbol{\alpha}^*$ and $\boldsymbol{\alpha}'$ be the optimal power allocations in (14) and (18) when $C_0 = C_{IPA}$. It follows that $\sum_i \alpha_i' \leq \sum_i \alpha_i^*$ (since $\boldsymbol{\alpha}'$ minimizes the total power). Assume that $\sum_i \alpha_i' < \sum_i \alpha_i^* = m$ and define $\alpha_i'' = m\alpha_i'/\sum_i \alpha_i' > \alpha_i'$, so that $\sum_i \alpha_i'' = m$. Since $C(\boldsymbol{\alpha})$ is an increasing function, $C(\boldsymbol{\alpha}'') > C(\boldsymbol{\alpha}') = C_{IPA}$, which is impossible since C_{IPA} is the largest system capacity. Therefore, $\sum_i \alpha_i' = \sum_i \alpha_i^* = m$. Since $C(\boldsymbol{\alpha})$ is strictly concave, the problem in (14) has a unique solution, so that $\alpha_i' = \alpha_i^*$. ∎

Finally, we remark that Theorems 2 and 3, Corollary 3 and Proposition 2 also apply to any multi-stream transmission system with independent streams and not only to ZF V-BLAST (e.g. spatial multiplexing over the channel eigenmodes, OFDM etc.), since we have not used any V-BLAST specifics except for stream independence.

VI. Joint Instantaneous Power/Rate Allocation (IPRA)

In this section, we study the joint optimum allocation of instantaneous power and rate. The key observation here is that, contrary to what one would expect [11], the conventional waterfilling algorithm (WF) does not provide an optimal solution to (5). Indeed, an implicit assumption behind the conventional WF is that the channel gains do not depend on the allocated power. This is not so for the V-BLAST because the SIC forces the equivalent channel gains $g_i = |\mathbf{h}_{i\perp}|^2$ to depend on the allocated powers, albeit in a binary way: if some transmitters are not active ($\alpha_i = 0$), there is no need to project out interference from those streams. Thus, turning off $i-$th stream affects the gains of lower-level streams $g_1 ... g_{i-1}$. This results in (5) being a non-convex problem for the V-BLAST, for which the conventional WF is in general not a solution. However, the problem can be split into 2^{m-1} convex sub-problems, one per each set of inactive transmitters, and each of the sub-problems can be solved via the conventional WF algorithm. The following theorem makes this idea precise.

Theorem 4 (FWF). *The joint optimum allocation of instantaneous power/rate for the coded V-BLAST (i.e. (5)) is given by the Fractional Waterfilling Algorithm (FWF) below:*

• *A. Split the problem:* `for` $l = 1, ..., 2^{m-1}$

Select a set of participating transmitters: if i-th bit in m-digit binary representation of $2l-1$ (l is an index of the set) is $l^{(i)} = 1$, then transmitter i participates in l-th set (1st transmitter always participates).

Calculate the per-stream gains with interference from yet-to-be-detected symbols projected out, $g_i^l = |\mathbf{h}_{i\perp}^l|^2$, $\mathbf{h}_{i\perp}^l \perp \left\{\mathbf{h}_{i+1} l^{(i+1)}, \ldots, \mathbf{h}_m l^{(m)}\right\}$, for $i = 1, \ldots, m$.

• *B. Do the WF on the set of participating transmitters:*

Calculate the power allocation:

$$\alpha_i^l = l^{(i)}(\nu_l - 1/(\gamma_0 g_i^l))_+, \tag{19}$$

where $x_+ = x$ if $x > 0$ and 0 otherwise, and the water level ν_l is found from the total power constraint $\sum_{i=1}^m \alpha_i^l = m$. The per-stream and total capacities are:

$$C_i^l = \ln\left(1 + \gamma_0 \alpha_i^l g_i^l\right), \ C^l = \sum_{i=1}^m C_i^l.$$

• *C. Finalize:* `End for` *(l)*

The optimum power and rate allocations are given by $\alpha_i^{l^\star}$ and $C_i^{l^\star}$, where $l^\star = \arg\max_l C^l$.

Proof: The optimization problem of simultaneous power and rate allocation is formally stated as $\boldsymbol{\alpha}^\star = \arg\max_{\boldsymbol{\alpha}\in S} C(\boldsymbol{\alpha})$, where $S = \{\boldsymbol{\alpha} : \alpha_i \geq 0, \sum \alpha_i \leq m\}$, the system capacity $C(\boldsymbol{\alpha})$ is

$$C(\boldsymbol{\alpha}) = \sum_{i=1}^m C_i(\boldsymbol{\alpha}) = \sum_{i=1}^m \ln(1 + g_i \alpha_i \gamma_0), \tag{20}$$

where $g_i = |\mathbf{h}_{i\perp}|^2$, and the per-stream rates are set equal to C_i. Step A of the algorithm divides the entire feasible set S into disjoint subsets corresponding to different patterns of participating transmitters: $S = \bigcup_{l=1}^{2^{m-1}} S_l$, where l determines which transmitters $2, \ldots, m$ are participating [5]. It follows that

$$\begin{aligned} C^\star &= \max_{\boldsymbol{\alpha}\in S} C(\boldsymbol{\alpha}) = \max_{\boldsymbol{\alpha}\in \bigcup_{l=1}^{2^{m-1}} S_l} C(\boldsymbol{\alpha}) \\ &= \max_l \max_{\boldsymbol{\alpha}\in S_l} C(\boldsymbol{\alpha}) = \max_l C^l, \end{aligned}$$

where $C^l = \max_{\boldsymbol{\alpha}\in S_l} C(\boldsymbol{\alpha})$ has a unique solution via the conventional WF algorithm[6] applied to the participating set S_l. This is done in Step B. Step C is obvious. ∎

Similarly to the IPA (see Theorem 3), the IPRA via the FWF also minimizes the total power required to achieve the target system capacity C^*. While the FWF is more complex than the conventional WF, its incremental complexity is low for small m. The following corollary shows that the FWF is close to the conventional WF at high SNR for a given channel realization.

Corollary 4. *For a given realization of a full rank channel, the fractional waterfilling algorithm converges to the conventional WF at high SNR, when both produce the uniform power allocation, $\alpha_i^\star \to 1$ when $\gamma_0 \to \infty$. For rank-deficient channels, both algorithms allocate no power to zero-gain dimensions and the same power to all active streams.*

Proof: From (19), $\lim_{\gamma_0\to\infty} \alpha_i^l = \nu_l$ and $\lim_{\gamma_0\to\infty} \nu_l = \frac{m}{\sum_i l^{(i)}}$. Let $C^{(p)} = \max_{l:\sum_i l^{(i)}=p} C^l$, i.e. $C^{(p)}$ is the maximum capacity attained with p participating transmitters. Let $g_i^{(p)}$ be the channel gains corresponding to $C^{(p)}$ (set $g_i^{(p)} = 0$ if the corresponding transmitter does not participates; all other $g_i^{(p)}$'s are strictly positive as follows from the full rank condition), and let $\nu_{(p)}$ be the corresponding water level.

[5] "participating" in the sense that they are not turned off at this step and thus affect the projection operation. They may be assigned zero power later on by the WF algorithm at step B, which however will not affect the projection operation. 1st transmitter always participates since it does not affect the gains of any other stream.

[6] "Conventional" in a sense that $h_{i\perp}$ are calculated only once taking into account all transmitters in the set S_l and the waterfilling is also done over this set. If some transmitters are assigned zero power by the WF algorithm, this does not lead to recalculation of $h_{i\perp}$.

Compare $C^{(p)}$ and $C^{(p+1)}$ as $\gamma_0 \to \infty$:

$$C^{(p+1)} - C^{(p)} = \ln\left(\frac{\prod_{i=1}^m \left(1+\nu_{(p+1)}\gamma_0 g_i^{(p+1)}\right)}{\prod_{i=1}^m \left(1+\nu_{(p)}\gamma_0 g_i^{(p)}\right)}\right) \to \ln\left(\gamma_0 \frac{\nu_{(p+1)}^{p+1}\prod_{i:g_i^{(p+1)}>0} g_i^{(p+1)}}{\nu_{(p)}^{p}\prod_{i:g_i^{(p)}>0} g_i^{(p)}}\right) > 0,$$

so that $C^m > C^{m-1} > \ldots > C^1$, and therefore the fully dimensional system (C^m) is optimal when $\gamma_0 \to \infty$. ■

Note that Corollary 4 does not imply that the respective outage probabilities converge to each other as $\gamma_0 \to \infty$. Indeed, as Corollary 9 demonstrates, they are very much different[7]. Unlike the average power and rate allocation (APRA) in [1], where the uniform power allocation is optimum when combined with the average rate allocation, the IPRA allocates power in a non-uniform way (except in some special cases). The following corollaries further characterize the FWF in Theorem 4.

Corollary 5. *If the allocation given by the FWF has all streams active (no $\alpha_i = 0$), it is the same as the allocation given by the conventional WF.*

Corollary 6. *If the conventional WF allocation has some inactive streams, the FWF one also has inactive stream(s) (not necessarily the same ones), but the converse is not necessarily true.*

Following Theorem 1, the FWF algorithm not only maximizes the instantaneous capacity C but also minimizes the system outage probability P_{out}.

VII. Performance Analysis

In this section, we present a comparative performance analysis of the unoptimized and optimized systems in different operating regimes.

A. Any SNR, any rate

Proposition 3. *For any channel realization, the system capacities of the coded V-BLAST with the FWF, the WF, the IRA, the IPA and the uniform power/rate allocation are bounded as follows:*

$$\ln(1+m\gamma_0 g_{max}) \le C_{FWF} \le m\ln(1+\gamma_0 g_{max}) \quad (21)$$

$$\ln(1+m\gamma_0 g_{max\perp}) \le C_{WF} \le m\ln(1+\gamma_0 g_{max\perp}) \quad (22)$$

$$\ln(1+\gamma_0 g_{max\perp}) \le C_{IRA} \le m\ln(1+\gamma_0 g_{max\perp}) \quad (23)$$

$$C_u \le C_{IPA} \le C_u(m\gamma_0) \quad (24)$$

$$C_u = m\ln\left(1+\gamma_0 \min_i |\mathbf{h}_{i\perp}|^2\right) \quad (25)$$

where $g_{max} = \max_i |\mathbf{h}_i|^2$, $g_{max\perp} = \max_i |\mathbf{h}_{i\perp}|^2$ are the maximum unprojected and projected stream gains. The bounds are tight (i.e. there are channel realizations that achieve them). The relationship $C_u \le C_{IRA} \le C_{WF} \le C_{FWF}$ always holds. Moreover, $C_u = C_{IRA} = C_{WF} = C_{FWF} = m\ln\left(1+\gamma_0|\mathbf{h}_1|^2\right)$ if and only if $\mathbf{H}^+\mathbf{H} = |\mathbf{h}_1|^2\mathbf{I}$, where $\mathbf{I}$ is the identity matrix, i.e the channel is orthogonal. No optimization is required in this case.

Proof: The left expression in (21) is the capacity in the regime with only one active transmitter, $\alpha_{i_{max}} = m$, where $i_{max} = \arg\max_i |\mathbf{h}_i|^2$ ($g_{i_{max}} = |\mathbf{h}_{i_{max}}|^2$ as there is no interference to project out when only one stream is active). Since the optimal capacity cannot be smaller, the lower bound in (21) holds. The upper bound in (21) is obtained as follows :

$$\begin{aligned} C_{FWF} &= \sum_i \ln\left(1+\alpha_i^\star\gamma_0|\mathbf{h}_{i\perp}|^2\right) \\ &\le \sum_i \ln\left(1+\alpha_i^\star\gamma_0 \max_i |\mathbf{h}_i|^2\right) \\ &\le m\ln\left(1+\gamma_0 \max_i |\mathbf{h}_i|^2\right) \end{aligned}$$

where $\boldsymbol{\alpha}^\star$ is the optimum power allocation for $g_i = |\mathbf{h}_{i\perp}|^2$. The first inequality is due to the fact that $|\mathbf{h}_{i\perp}|^2 \le \max_i |\mathbf{h}_i|^2$. To obtain the second inequality, consider a fictitious MIMO channel with orthogonal subchannels of equal gains such that $|\mathbf{h}_{i\perp}|^2 = \max_i |\mathbf{h}_i|^2$. As seen from (19), the optimal FWF solution for such channel is the uniform power allocation and $\boldsymbol{\alpha}^\star$ is sub-optimal in general in this channel, which yields the upper bound in (21). The bounds for conventional waterfilling (22) follow from the same reasoning. The difference between the two stems from the fact that the WF is oblivious to the possibility that some transmitters may be inactive and hence do not require projecting out their subspace, so that the maximum possible gain for the WF is $\max_i |\mathbf{h}_{i\perp}|^2$. The lower bound for IRA capacity in (23) is the largest term in the sum $C_{IRA} = \sum_i \ln\left(1+\gamma_0|\mathbf{h}_{i\perp}|^2\right)$. The upper bound is obtained by upper bounding each term of the sum. For the regular V-BLAST (uniform power/rate allocation), the weakest stream dominates the performance so that (25) follows from (13). (24) follows from (17). ■

One can also envision an iterative water-filling (IWF) algorithm, where the per-stream gains are re-computed every time one stream is assigned 0 power (to eliminate orthogonal projection over respective $\mathbf{h}_i$; this may require several iterations, at most $m-1$, to converge). It is straightforward to see that its system capacity is between those of the WF and FWF, $C_{WF} \le C_{IWF} \le C_{FWF}$.

Comparing the system capacities of the FWF applied to the unordered ZF V-BLAST and of the WF applied to the optimally-ordered ZF V-BLAST, we observe that the former is in general sub-optimal. However, its computational complexity is also much smaller: while the optimally-ordered WF V-BLAST requires $m!$ WF runs, the FWF requires only 2^{m-1} ($< m!$ for $m > 2$) WF runs, most of which are sparse (i.e. many streams are inactive). Furthermore, these two capacities coincide in the low SNR regime, as follows from (39) and (40). The same can be demonstrated, with some effort, for the high SNR regime.

Let us consider the cases when some optimization strategies offer no advantage.

Corollary 7. *The following holds:*

• $C_u = C_{IPA} = C_{IRA} = C_{WF}$ *if and only if* $|\mathbf{h}_{1\perp}|^2 = \ldots =$

[7]This happens because, for any finite SNR, there are always bad channel realizations for which the WF and the FWF power/rate allocations are different, and it is these bad realizations that dominate the outage performance.

$|\mathbf{h}_{m\perp}|^2$.

• $C_u = C_{WF}$ *if* $C_u = C_{IRA}$ *or* $C_u = C_{IPA}$, *i.e. if there is no advantage in the IRA or IPA (compared to the unoptimized system), there is no advantage in the WF either, only the FWF may bring an improvement.*

Proposition 4. *The outage probabilities of the FWF, the WF, the IRA, the IPA and the uniform power/rate allocation are ordered in an arbitrary-fading channel as follows:*

$$P_{out}^{FWF} \leq P_{out}^{WF} \leq P_{out}^{IRA} \leq P_{out}^{u} \quad (26)$$

$$P_{out}^{u}(m\gamma_0) \leq P_{out}^{IPA} \leq P_{out}^{u} \quad (27)$$

with the equalities if $m = 1$. *In the i.i.d. Rayleigh-fading channel, the equalities are achieved only if* $m = 1$.

Proof: Each inequality in (26) follows from the fact that its left-hand side corresponds to optimization over a feasible set that is larger compared to that of its right-hand side. (27) follows from (17). If $m = 1$, there is nothing to optimize - hence, the equality. The "only if" part follows from the fact that a Gaussian density is strictly positive for any finite value of its argument so that the difference between outage sets corresponding to two different optimization strategies has non-zero measure. ■

Based on Proposition 3, the outage probabilities can now be characterized in a more specific way.

Corollary 8. *For any* R *and any* γ_0, *the outage probabilities of the coded V-BLAST with the uniform power/rate allocation, the IRA, the IPA, the WF and the FWF in the i.i.d. Rayleigh fading channel are bounded as follows:*

$$F_n^m(z) \leq P_{out}^{FWF} \leq F_n^m(z_m/m) \quad (28)$$

$$\prod_{i=1}^{m} F_{n-m+i}(z) \leq P_{out}^{WF} \leq \prod_{i=1}^{m} F_{n-m+i}(z_m/m) \quad (29)$$

$$\prod_{i=1}^{m} F_{n-m+i}(z) \leq P_{out}^{IRA} \leq \prod_{i=1}^{m} F_{n-m+i}(z_m) \quad (30)$$

$$P_{out}^{u} = 1 - \prod_{i=1}^{m}\left(1 - F_{n-m+i}(z)\right) \quad (31)$$

where $z = (e^R - 1)/\gamma_0$, $z_m = (e^{mR} - 1)/\gamma_0$, $F_k(x) = 1 - e^{-x}\sum_{l=0}^{k-1} x^l/l!$ *is the outage probability of* k*-th order MRC.*

Proof: Observe that $|\mathbf{h}_i|^2 \sim \chi^2_{2n}$, all independent of each other, and $|\mathbf{h}_{i\perp}|^2 \sim \chi^2_{2(n-m+i)}$, and also independent of each other [12]. For $X \sim \chi^2_{2k}$, we have $\Pr\{X < x\} = F_k(x)$. Using these facts, the bounds of Corollary 8 follow from the bounds of Proposition 3. ■

B. Low outage probability regime

The diversity gains can now be characterized in the low-outage regime based on Corollary 8. The diversity gain can be found from [8]

$$d = -\lim_{\gamma_0 \to \infty} \frac{\ln P_{out}}{\ln \gamma_0}. \quad (32)$$

or by inspection when a closed-form low-outage approximation of P_{out} is available.

Corollary 9. *For fixed* R, *the diversity gains of the unoptimized V-BLAST, the IRA, the IPA, the WF and the FWF are related as follows:*

$$d_u = n - m + 1 = d_{IPA}$$
$$\leq d_{WF} = d_{IRA} = \sum_{i=1}^{m}(n - m + i) \leq d_{FWF} = nm$$

The equality is achieved for $m = 1$ *only, i.e. only the FWF achieves the full MIMO channel diversity* nm *for* $m > 1$. [8]

Proof: Using the well-known approximation $F_k(x) = x^k/k! + o(x^k)$, $x \to 0$, in the upper and lower bounds to P_{out} in each equation of Corollary 8 and substituting it into (32), one observes that the lower and upper bounds give the same diversity gain, which is therefore the diversity gain. ■

The instantaneous rate allocation is the most efficient of all the techniques in terms of incremental improvement as it brings significant diversity gain and keeps the rate close to the capacity. When $m > 1$, the full MIMO channel diversity is achieved by the FWF only.

C. Any SNR, low rate (wideband) regime

The $R \ll 1$ regime here is also known as the wideband regime [13][14] (since R is in [nat/s/Hz], i.e. the rate per unit bandwidth), which is a popular solution for many systems (e.g. CDMA).

Theorem 5. *For any* γ_0 *and low rate* $R \ll 1$, *the outage probabilities of the coded V-BLAST with the uniform power/rate allocation, the IRA, the WF and the FWF in the i.i.d. Rayleigh fading channel are given by*[9]

$$P_{out}^{u} = 1 - \prod_{i=1}^{m}\left(1 - F_{n-m+i}(x)\right) \approx \frac{x^{n-m+1}}{(n-m+1)!}, \quad (33)$$

$$P_{out}^{IRA} \approx F_{d_{IRA}}(mx) \approx \frac{1}{d_{IRA}!}(mx)^{d_{IRA}}, \quad (34)$$

$$P_{out}^{WF} \approx \prod_{i=1}^{m} F_{n-m+i}(x) \approx \frac{x^{d_{IRA}}}{\prod_{i=1}^{m}(n-m+i)!} \quad (35)$$

$$P_{out}^{FWF} \approx F_n^m(x) \approx \frac{x^{nm}}{(n!)^m} \quad (36)$$

where the second approximation in each case holds at the low outage regime, $x = R/\gamma_0 \ll 1$.

Proof: 1st equality in (36) and (35) follows immediately from Corollary 8 by applying the low rate approximation $e^{mR} - 1 \approx mR$ and observing that the lower and upper bounds coincide. To show 1st equality in (34), notice that

$$P_{out}^{IRA} = \Pr\left\{\sum_i \ln\left(1 + \gamma_0|\mathbf{h}_{i\perp}|^2\right) < mR\right\}$$
$$\overset{(a)}{\approx} \Pr\left\{\gamma_0 \sum_i |\mathbf{h}_{i\perp}|^2 < mR\right\} \overset{(b)}{=} F_{d_{IRA}}\left(\frac{mR}{\gamma_0}\right)$$

[8] This conclusion is not in contradiction to Corollary 9 of [18] since the latter (as well as Theorem 7 in [18]) requires all streams to be active, which is not the case for fixed R and $\gamma_0 \to \infty$.

[9] to the best of our knowledge, it is the first time when the outage probability of the waterfilling algorithm is found in an explicit, closed form.

where (a) follows from $\ln(1+x) \approx x$ when $x \ll 1$, so that $\ln(1+\gamma_0|\mathbf{h}_{i\perp}|^2) \approx \gamma_0|\mathbf{h}_{i\perp}|^2$ under an outage event $C_{IRA} < mR \ll 1$; (b) follows from the fact that $\sum_i |\mathbf{h}_{i\perp}|^2 \sim \chi^2_{2d_{IRA}}$. 2nd equality in all cases follows from the standard approximation of $F_k(x)$ for small x. ■

Note that the FWF not only has a higher diversity gain, but also an SNR gain of $\prod_{i=1}^m n!/(n-m+i)!$ over the WF.

D. Low SNR regime

The next result characterizes the optimization strategies at the low SNR regime. We emphasize that low SNR does not imply high error rate in coded systems, unlike uncoded ones. For example, the outage probability of the coded system is small at low SNR as long as $R/\gamma_0 \ll 1$. Many practical systems (e.g. CDMA) operate in this regime [8].

Theorem 6. *In the low SNR regime, $m\gamma_0 \max_i |\mathbf{h}_i|^2 \ll 1$, the instantaneous capacities of the regular (unoptimized) V-BLAST and of the IRA, the WF, and the FWF are given by*

$$C_u \approx m\gamma_0 \min_i |\mathbf{h}_{i\perp}|^2 \quad (37)$$

$$C_{IRA} \approx \gamma_0 \sum_{i=1}^{m} |\mathbf{h}_{i\perp}|^2 \quad (38)$$

$$C_{WF} \approx m\gamma_0 \max_i |\mathbf{h}_{i\perp}|^2 \quad (39)$$

$$C_{FWF} \approx m\gamma_0 \max_i |\mathbf{h}_i|^2 \quad (40)$$

and these capacities are attained by the following power/rate allocations:

- (38) *is attained by* $R_i^{IRA} = \gamma_0|\mathbf{h}_{i\perp}|^2$,
- (39) *is attained by* $\alpha_{i_{max}}^{WF} = m$, $R_{i_{max}}^{WF} = m\gamma_0|\mathbf{h}_{i_{max}\perp}|^2$, *and 0 otherwise, where* $i_{max} = \arg\max_i |\mathbf{h}_{i\perp}|^2$ *is the strongest projected channel,*
- (40) *is attained by* $\alpha_{i_{max}}^{FWF} = m$, $R_{i_{max}}^{FWF} = m\gamma_0|\mathbf{h}_{i_{max}}|^2$ *and 0 otherwise, where* $i_{max} = \arg\max_i |\mathbf{h}_i|^2$ *is the strongest unprojected channel.*

Proof: In the capacity expressions of Proposition 3, apply the approximation $\ln(1+x) \approx x$, which is valid for $x < 1$, and observe that the lower and upper bounds coincide. ■

As a side remark, note that the outage probabilities in Theorem 5 are the same as those obtained based on the right hand sides of (37)-(40). Additionally, it follows from Theorem 6 that the FWF significantly outperforms the WF, $C_{WF} \ll C_{FWF}$, when $\max_i |\mathbf{h}_{i\perp}| \ll \max_i |\mathbf{h}_i|$ and their performance is close otherwise.

Based on Theorem 6, the following holds.

Corollary 10. *In the low SNR regime, the WF and FWF are related as follows: $C_{WF} = C_{FWF}$ if and only if $i_{max} = \arg\max_i |\mathbf{h}_{i\perp}|^2 = m$ or $\mathbf{h}_{i_{max}} \perp \{\mathbf{h}_{i_{max}+1}, \ldots, \mathbf{h}_m\}$.*

One may further consider the optimally-ordered WF and FWF (OWF and OFWF). Based on (40) and (39), the following relationship holds in the low-SNR regime: $C_{OFWF} = C_{FWF} = C_{OWF} \geq C_{WF}$, i.e. the ordering does not improve the FWF, but does improve the WF making it equal to the FWF.

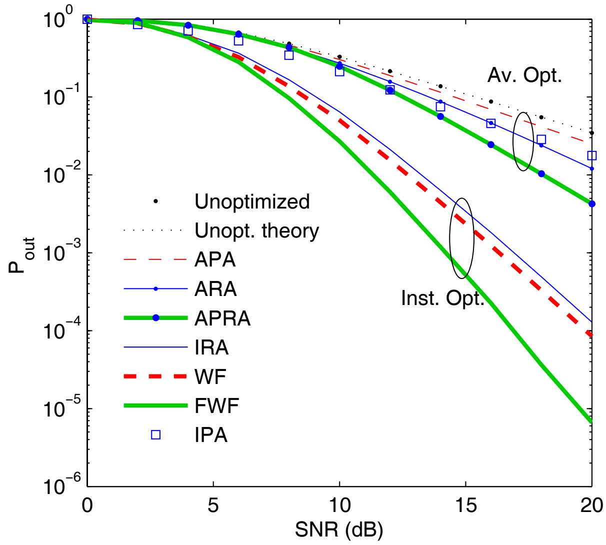
VIII. Examples

To obtain some additional insights and to validate the analytical results above, we consider here 2x2 V-BLAST in the i.i.d. Rayleigh fading channel under the optimization strategies discussed above. Outage probabilities of the V-BLAST with these optimization strategies obtained by Monte-Carlo simulations and the approximations above are shown in Fig. 2. As follows from the analysis, both the IRA and the FWF provide a significant improvement over the unoptimized system. As per Corollary 9, the conventional WF fails to achieve the minimum outage probability for a given total rate and to provide any diversity gain over the IRA (both have the diversity gain of 3), while the FWF achieves the full MIMO diversity of 4, outperforming both the IRA and the WF by a wide margin (about 3-5 dB at $P_{out} = 10^{-3}$). Note also a significant advantage of the instantaneous optimization over the average one: while the average power and rate allocation achieves the outage probability $\approx 10^{-2}$ at $\gamma_0 = 20$ dB in Fig. 2(a), the FWF achieves the outage probability $\approx 10^{-5}$ at the same SNR. The instantaneous optimization is also significantly superior to the average one in the low-rate low-SNR regime, as Fig. 2(b) shows, so that more feedback required for the former pays off well: the FWF brings the outage probability down to $\approx 10^{-6}$ from $\approx 10^{-1}$ for the unoptimized system at SNR = 0 dB. Note also that, for this system, the FWF requires only one extra run with only one stream active compared to the WF, i.e. very small incremental complexity. Apart from simulations, Fig. 2(b) also shows the low-rate approximations in Theorem 5, which are remarkably accurate in this regime.

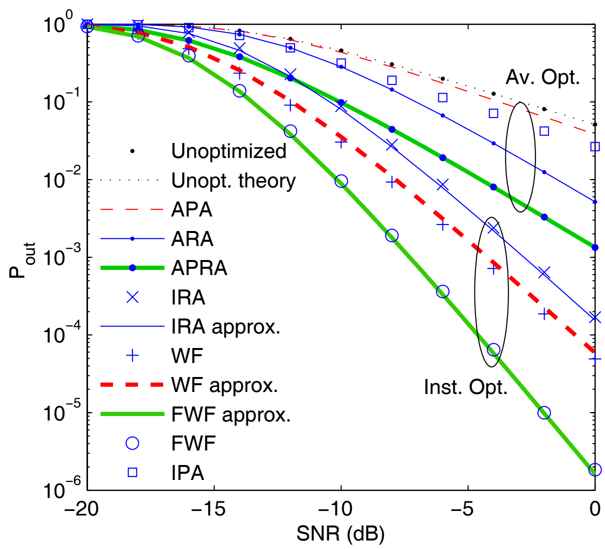
To illustrate Corollaries 4, 7 and 10, and Theorem 6, Fig. 3 shows the instantaneous capacity for two typical channel realizations: a 'good' channel (a) with almost orthogonal subchannels, $|\mathbf{h}_{1\perp}| = |\mathbf{h}_2| \approx |\mathbf{h}_1|$, and a 'bad' channel (b) with almost parallel subchannels, $|\mathbf{h}_{1\perp}| = |\mathbf{h}_2| \ll |\mathbf{h}_1|$. The per-stream gains in the ZF V-BLAST are the same in these two cases, $g_1 = g_2 = 1$, so that the WF allocates the same powers and rates to the two streams, $\alpha_1 = \alpha_2 = 1$, $R_1 = R_2 = R$, and is identical to the IRA, the IPA and the unoptimized system, $C_{WF} = C_{IRA} = C_{IPA} = C_u$, over the entire SNR range in both cases. This behavior agrees with the statement of Corollary 7, which says that if the IRA or IPA fails to provide any improvement over the unoptimized system then so does the WF. As predicted by Corollary 4, the FWF coincides with the WF at high SNR. In contrast, the FWF is the only strategy to yield a performance gain at low SNR: $C_{FWF} \approx 2C_{WF}$ for the 'good' case in Fig. 3(a), i.e. slightly better than the conventional WF, and $C_{FWF} \approx 100C_{WF}$ for the 'bad' one in Fig. 3(b), i.e. a significant advantage when $|\mathbf{h}_{1\perp}| = |\mathbf{h}_2| \ll |\mathbf{h}_1|$. The 100-fold gain of the FWF over the conventional WF (and, of course, over the unoptimized system) in the 'bad' channel is particularly advantageous for static or quasi-static channels (very slow fading), when a user might experience a 'bad' channel realization for a very long period of time.

IX. Conclusion

Instantaneous power, rate and joint power-rate allocations for the coded V-BLAST have been studied. Since the conventional waterfilling algorithm fails to maximize the capacity



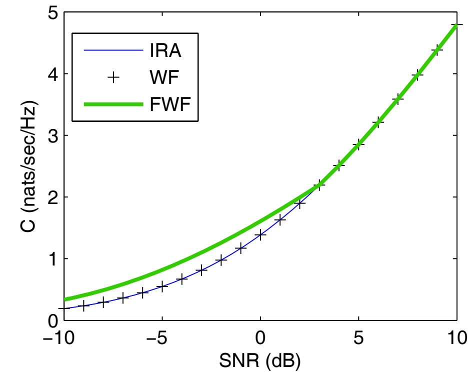
(a)



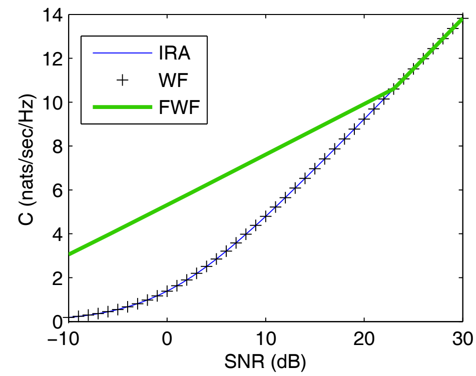
(b)

Fig. 2. Outage probability vs. average SNR for the average (Av.) and instantaneous (Inst.) optimization strategies for 2×2 V-BLAST in i.i.d. Rayleigh fading channel for (a) high rate $R = 3$ [nat/s/Hz] (simulations), and (b) low rate $R = 0.1$ [nat/s/Hz] (simulations (symbols) and the approximations in Theorem 5 (lines)). APA, ARA and APRA are the average power, rate and power/rate allocations from [1].

of the coded ZF V-BLAST, the fractional waterfilling algorithm has been proposed, which implements joint power and rate allocation that simultaneously maximizes instantaneous system capacity and minimizes the outage probability. Unlike the average optimization strategies and the IPA, the IRA and IPRA provide a diversity gain over the unoptimized system. The fractional waterfilling algorithm significantly outperforms the other strategies. As established by both the analysis and simulations, the FWF algorithm achieves the full MIMO diversity, while the IRA, IPA and the WF do not. A number of closed-form bounds and accurate approximations to the outage performance of these algorithms have been given. Many of these results also apply to generic multi-stream transmission systems (spatial multiplexing over channel eigenmodes, OFDM) or the systems relying on successive interference cancelation (multi-user detection, channel equalization).



(a)



(b)

Fig. 3. Instantaneous system capacity of 2×2 V-BLAST with the FWF and the conventional WF for two specific channel realizations; (a) - 'good", (b) - "bad". While the FWF exhibits only a small performance improvement for the "good" realization (a): $\mathbf{h}_1 = [1\ 1]^T$, $\mathbf{h}_2 = [0\ 1]^T$, its superiority for "bad" realization (b): $\mathbf{h}_1 = [1\ 10]^T$, $\mathbf{h}_2 = [0\ 1]^T$, is significant, especially at the low-SNR regime. Note that $C_{IRA} = C_{IPA} = C_{WF} = C_u$ in these examples, so that only the FWF provides an improvement.

X. Appendix

The problem in (14) is equivalent to the following:

$$\max_{\boldsymbol{\alpha}, R} R, \text{ s.t. } R \leq \ln(1 + \alpha_i g_i \gamma_0), \sum_i \alpha_i = m, \alpha_i \geq 0 \quad (41)$$

which is clearly a convex problem and for which the Lagrangian is

$$L = R - \sum_i \lambda_i (R - \ln(1 + \alpha_i g_i \gamma_0)) - \nu \left(\sum_i \alpha_i - m \right) + \sum_i \beta_i \alpha_i \quad (42)$$

where λ_i, β_i, ν are Lagrange multipliers. The corresponding KKT conditions (see e.g. [15]) are

$$\frac{\partial L}{\partial R} = 1 - \sum_i \lambda_i = 0 \quad (43)$$

$$\frac{\partial L}{\partial \alpha_i} = \frac{\lambda_i g_i \gamma_0}{1 + \alpha_i g_i \gamma_0} - \nu + \beta_i = 0 \quad (44)$$

$$\alpha_i, \beta_i, \lambda_i \geq 0, \ R \leq \ln(1 + \alpha_i g_i \gamma_0), \ \sum_i \alpha_i = m, \quad (45)$$

$$\beta_i \alpha_i = 0, \ \lambda_i (R - \ln(1 + \alpha_i g_i \gamma_0)) = 0 \quad (46)$$

Using these conditions and observing that $\alpha_i > 0 \ \forall i$ (otherwise $R = 0$), one obtains, after some manipulations,

$$\alpha_i = \frac{e^R - 1}{g_i} = \frac{m}{g_i \sum_i g_i^{-1}} = \frac{\overline{g}}{g_i} \tag{47}$$

from which $R = \ln(1 + \overline{g}\gamma_0)$ follows. □

REFERENCES

[1] V. Kostina and S. Loyka, "Optimum power and rate allocation for coded V-BLAST: average optimization," *IEEE Trans. Commun.*, vol. 59, no. 3, pp. 877-887, Mar. 2011.

[2] N. Prasad and M. Varanasi, "Analysis of decision feedback detection for MIMO Rayleigh-fading channels and the optimization of power and rate allocations," *IEEE Trans. Inf. Theory*, vol. 50, no. 6, pp. 1009-1025, June 2004.

[3] S. Nam, O. Shin, and K. Lee, "Transmit power allocation for a modified V-BLAST system," *IEEE Trans. Commun.*, vol. 52, no. 7, pp. 1074-1079, July 2004.

[4] R. Kalbasi, D. Falconer, and A. Banihashemi, "Optimum power allocation for a V-BLAST system with two antennas at the transmitter," *IEEE Commun. Lett.*, vol. 9, no. 9, pp. 826-828, Sep. 2005.

[5] N. Wang and S. Blostein, "Approximate minimum BER power allocation for MIMO spatial multiplexing systems," *IEEE Trans. Commun.*, vol. 55, no. 1, pp. 180-187, Jan. 2007.

[6] V. Kostina and S. Loyka, "On optimum power allocation for the V-BLAST," *IEEE Trans. Commun.*, vol. 56, no. 6, pp. 999-1012, June 2008.

[7] I. Sason and S. Shamai, "Performance analysis of linear codes under maximum-likelihood decoding: a tutorial," *Foundations Trends Commun. Inf. Theory*, vol. 3, no. 1/2, pp. 1-222, 2006.

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

[9] G. Caire and K. Kumar, "Information theoretic foundations of adaptive coded modulation," *Proc. IEEE*, vol. 95, no. 12, pp. 2274-2298, Dec. 2007.

[10] J. Cioffi, *Digital Communications (course notes)*. Stanford University, 2007.

[11] R. Zhang and J. Cioffi, "Approaching MIMO-OFDM capacity with zero-forcing V-BLAST decoding and optimized power, rate, and antenna-mapping feedback," *IEEE Trans. Signal Process.*, vol. 56, no. 10, pp. 5191-5203, Oct. 2008.

[12] S. Loyka and F. Gagnon, "V-BLAST without optimal ordering: analytical performance evaluation for Rayleigh fading channels," *IEEE Trans. Commun.*, vol. 54, no. 6, p. 1109-1120, June 2006.

[13] S. Verdu, "Recent results on the capacity of wideband channels in the low-power regime," *IEEE Wireless Commun.*, pp. 40-45, Aug. 2002.

[14] S. Verdu, "Spectral efficiency in the wideband regime," *IEEE Trans. Inf. Theory*, vol. 48, no. 6, pp. 1319-1343, June 2002.

[15] L. Boyd, S. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.

[16] S. Verdu and T. S. Han, "A general formula for channel capacity," *IEEE Trans. Inf. Theory*, vol. 40, no. 4, pp. 1147-1157, July 1994.

[17] M. Effros, A. Goldsmith, and Y. Liang, "Generalizing capacity: new definitions and capacity theorems for composite channels," *IEEE Trans. Inf. Theory*, vol. 56, no. 7, pp. 3069-3087, July 2010.

[18] V. Kostina and S. Loyka, "Optimum power and rate allocation for coded V-BLAST," *IEEE ICC*, June 2009.

Victoria Kostina received a Bachelor's degree with honors in applied mathematics and physics from the Moscow Institute of Physics and Technology, Russia, in 2004, where she was with the Institute for Information Transmission Problems, and a Master's degree in electrical engineering from the University of Ottawa, Canada, in 2006. She is currently working towards her Ph.D. at Princeton University, USA. Her research interests lie in information theory, probability theory, coding and wireless communications.

Sergey Loyka (M'96-SM'04) was born in Minsk, Belarus. He received the Ph.D. degree in Radio Engineering from the Belorussian State University of Informatics and Radioelectronics (BSUIR), Minsk, Belarus in 1995 and the M.S. degree with honors from Minsk Radioengineering Institute, Minsk, Belarus in 1992. Since 2001, he has been a faculty member at the School of Electrical Engineering and Computer Science, University of Ottawa, Canada. Prior to that, he was a research fellow in the Laboratory of Communications and Integrated Microelectronics (LACIME) of Ecole de Technologie Superieure, Montreal, Canada; a senior scientist at the Electromagnetic Compatibility Laboratory of BSUIR, Belarus; an invited scientist at the Laboratory of Electromagnetism and Acoustic (LEMA), Swiss Federal Institute of Technology, Lausanne, Switzerland. His research areas include wireless communications and networks, MIMO systems and smart antennas, RF system modeling and simulation, and electromagnetic compatibility, in which he has published extensively. Dr. Loyka is a technical program committee member/chair of several IEEE conferences and a reviewer for numerous IEEE periodicals and conferences. He received a number of awards from the URSI, the IEEE, the Swiss, Belarus and former USSR governments, and the Soros Foundation.