

Information and multiaccess interference in a complexity-constrained vector channel

Rodrigo de Miguel¹, Ori Shental^{2,4}, Ralf R Müller¹ and Ido Kanter³

¹ Department of Electronics and Telecommunications, Norwegian University of Science and Technology, 7491 Trondheim, Norway

² Department of Electrical Engineering-Systems, Tel-Aviv University, Tel Aviv 69978, Israel

³ Minerva Center and Department of Physics, Bar-Ilan University, Ramat-Gan 52900, Israel

E-mail: demiguel@iet.ntnu.no, shentalo@eng.tau.ac.il, oshental@ucsd.edu, ralf@iet.ntnu.no and kanter@mail.biu.ac.il

Received 16 January 2007, in final form 30 March 2007

Published 30 April 2007

Online at stacks.iop.org/JPhysA/40/5241

Abstract

A noisy vector channel operating under a strict complexity constraint at the receiver is introduced. According to this constraint, detected bits, obtained by performing hard decisions directly on the channel's matched filter output, must be the same as the transmitted binary inputs. An asymptotic analysis is carried out using mathematical tools imported from the study of neural networks, and it is shown that, under a bounded noise assumption, such complexity-constrained channel exhibits a non-trivial Shannon-theoretic capacity. It is found that performance relies on rigorous interference-based multiuser cooperation at the transmitter and that this cooperation is best served when all transmitters use the same amplitude.

PACS numbers: 89.70.+c, 05.20.-y, 84.35.+i

(Some figures in this article are in colour only in the electronic version)

1. Introduction

Multiple-access systems described by vector channels (MIMO, CDMA) are prominent in modern wireless communications. As such, revealing their information-theoretic properties has been a fruitful source of ongoing research (see, e.g., [1, 2]). In these systems, information is conveyed simultaneously from a group of sources to another group of sinks over the same physical medium and bandwidth. These transmissions are not orthogonal and interfere with each other. Detrimental effects such as multiaccess interference and noise can be completely

⁴ Current address: Center for Magnetic Recording Research (CMRR), University of California, San Diego, 9500 Gilman Drive, La Jolla, CA 92093, USA.

eliminated in theory by adopting optimal detection schemes and sophisticated error-correcting codes [3].

As in many problems of reliable (i.e., error-free) communication via a detrimental channel, typical information-theoretic investigations of multiaccess systems are often carried out under an upper-bounded transmission power or limited bandwidth, but usually no restrictions on complexity are imposed. In the era of pervasive and ubiquitous wireless communication systems, involving only limited computing power, there is an emerging interest in the workings of a stricter complexity-constrained scenario [4].

In this contribution, such a reduced complexity setting is introduced and analysed. This reduced complexity scenario requires that detected bits, sliced at the output of a bank of filters matched to the vector channel, be the same as the transmitted binary inputs. This constraint is analogous to the constraint on single-neuron (or spin) flip metastable states of the Hopfield model for neural networks [5, 6]. We term this scheme *vector input preselection (VIP)*, as transmitted bit combinations are preselected so as to comply with this strict constraint. Such a transmission scheme would result in the appealing use of low-cost receivers, in the context of reduced signal processing and computing. In a sense, this scheme is equivalent to outsourcing part of the detection complexity back to the transmitters. However, although the receivers need only do minimal signal processing, communication does not conclude until they properly interpret the received bits. This can be done using a decoding table identical to the transmitters' joint coding table. Although the *complexity* of this task cannot be outsourced back to the transmitters, it is solved by proper memory allocation [3].

The Hopfield model has been used before for developing suboptimal multiuser detectors [7–9]. In this contribution, the Hopfield model is used to compute the Shannon capacity of the noisy VIP channel in the many-users limit, revealing the cost in ‘information rate’ caused by limiting the channel complexity. Interestingly, we find that contrary to intuition, the VIP channel yields a non-trivial capacity.

We analyse the effect that the power distribution among the transmitters has on the channel capacity. Although equal power interference is found to be the worst case interference scenario for linear multiuser receivers [2] we find that, for the VIP channel, optimal cooperative interference occurs when all transmitters use equal amplitudes. Indeed we find that cooperation through interference is a key feature of the VIP channel, which is in contrast to systems without input preselection, where higher channel capacity is obtained by interference mitigation at the receiver [2, 10, 11]. It is important to note that the VIP rule employed in this paper is conceptually different to the known technique of multiuser precoding (or pre-equalization) [12–14] in which there is no restriction on the input signalling.

The organization of the present paper is as follows. The basic vector channel model is outlined in section 2. The asymptotic system analysis is briefly introduced in section 3. In section 4, the expression for the asymptotic information capacity of the channel is presented. The results are shown in section 5 and are further discussed in section 6. Appendices A and B provide full mathematical details.

2. Channel model

Consider a synchronous noisy vector channel with T transmitters and N receivers. The channel is characterized by an $N \times T$ random matrix \mathbf{S} with independent identically distributed entries with unit variance and zero mean. The matrix \mathbf{S} is assumed to be perfectly known at both ends of the channel. The input vector consists of T binary entries which are subject to transformation by the channel matrix \mathbf{S} . The N -length received column vector is projected

back into the T -dimensional information space by passing through a bank of filters matched to the channel matrix \mathbf{S} .

Thus overall, an input column vector $\mathbf{i} \triangleq (i_1, \dots, i_T)$ is transformed into an output $\boldsymbol{\theta} \triangleq (\theta_1, \dots, \theta_T)$ as follows:

$$\boldsymbol{\theta} = \frac{1}{N} \mathbf{S}^\dagger \mathbf{S} \mathbf{A} \mathbf{i} + \mathbf{n}, \quad (1)$$

where the diagonal matrix $\mathbf{A} \triangleq \text{diag}\{a_1, \dots, a_T\}$ controls the transmission power. The vector \mathbf{n} is a T -dimensional random vector with arbitrary additive noise components bounded to the interval $(-\kappa, \kappa)$, where the threshold κ is a non-negative constant.

2.1. Vector input preselection

Consider a potential energy surface in \mathfrak{R}^T where each $\mathbf{i} \in \{-1, 1\}^T$ has its own well of attraction. When an input vector \mathbf{i} is perturbed by the channel operator as in (1), the resulting state may remain within the initial well or jump to another well (or to a local maximum between several wells). In the model under consideration, the vector $\boldsymbol{\theta} \in \mathfrak{R}^T$ is within the well of attraction of \mathbf{i} when $\text{sgn}(\boldsymbol{\theta}) = \mathbf{i}$, where the equality and sign function apply to each of the T vector components.

To ease the task of detection at the receiver, the transmitter, having perfect channel state information, carries out a vector input preselection (VIP) process. To meet the VIP rule an input vector \mathbf{i} must be basin invariant under channel perturbation (1) or, equivalently, satisfy the following equality:

$$\mathbf{i} \equiv \text{sgn} \left\{ \frac{1}{N} \mathbf{S}^\dagger \mathbf{S} \mathbf{A} \mathbf{i} + \mathbf{n} \right\}. \quad (2)$$

Although the VIP process restricts the vectors eligible for transmission to a subset of the 2^T possible binary vectors, the receiver can be certain that reception is error free by a trivial use of conventional demodulation

$$\mathbf{i} = \text{sgn}(\boldsymbol{\theta}), \quad (3)$$

and without the use of complicated signal processing. The input vectors which satisfy (2) are, in the noiseless limit, equivalent to the metastable states of the Hopfield model.

The VIP rule is effectively outsourcing the complexity of detection from the receiver to the transmitter, which must coordinate its multiple inputs, thus limiting its choice of symbols for transmission. Figure 1 shows a diagram of the multiaccess VIP channel, which may describe, for example, a cooperative MIMO channel. An important feature in this scheme of simple detection by input preselection is that throughput bounds are invariant to the exact noise probability distribution, depending only on the noise upper and lower bounds $(-\kappa, \kappa)$.

3. Channel states and the thermodynamic limit

Equation (2), which represents the condition that a vector \mathbf{i} must satisfy in order to remain within its basin of attraction, may be rewritten as follows:

$$\int_{((\kappa-1)\beta^{-1}, \infty)^T} d^T \lambda \delta\{T^{-1} \mathbf{S}^\dagger \mathbf{S} \mathbf{A} \mathbf{i} - (\beta^{-1} \mathbf{I} + \boldsymbol{\Lambda}) \mathbf{i}\} = 1, \quad (4)$$

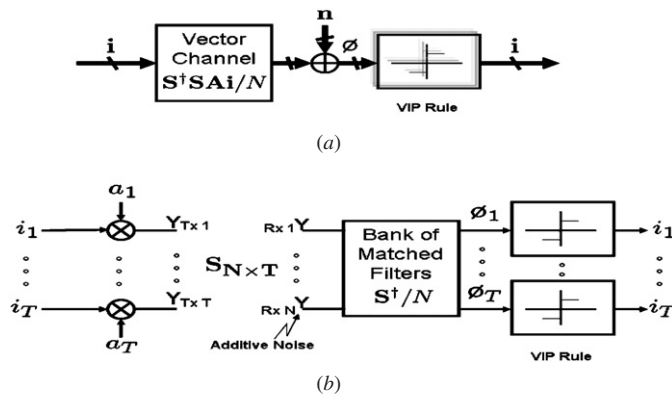


Figure 1. (a) The VIP-vector channel. (b) A VIP-MIMO channel example with T transmitters (Tx) and N receivers (Rx).

where $\beta \equiv T/N$, \mathbf{I} is the identity matrix and the T variables of integration are the eigenvalues of the diagonal matrix $\mathbf{\Lambda}$. The total number \mathbb{N} of input symbols that obey the VIP rule is thus given by

$$\mathbb{N}(\mathbf{S}; T, \beta, \mathbf{A}, \kappa) = \sum_{\mathbf{i}} \int_{((\kappa-1)\beta^{-1}, \infty)^T} d^T \lambda \delta\{T^{-1}\mathbf{S}^\dagger \mathbf{S} \mathbf{A} \mathbf{i} - (\beta^{-1}\mathbf{I} + \mathbf{\Lambda})\mathbf{i}\}, \quad (5)$$

where the sum is taken over all 2^T possible input vectors that may, in principle, be generated at the transmitter.

The channel matrix \mathbf{S} can exist in many different quenched states. We may consider, for simplicity and illustrative purposes, a channel matrix consisting of independent identically distributed binary entries. While *microscopically* we specify the state of the channel as one of 2^{NT} equiprobable states \mathbf{S} (henceforth microstates), *macroscopically* the behaviour, or macrostate, of the channel may be expressed as the number of inputs it allows to fulfil the VIP rule. Although the specific set of such inputs depends on the channel microstate, it is only the cardinality of the set that determines a macrostate. Therefore, more than one microstate may correspond to the same macrostate. Let $\Omega(\mathbb{N}; T, \beta, \mathbf{A}, \kappa)$ be the number of channel microstates that allow exactly and no more than \mathbb{N} inputs to satisfy the VIP rule, that is the number of microstates that correspond to macrostate- \mathbb{N} for T transmitters, a channel load β , an amplitude distribution \mathbf{A} and a noise bound κ . There are a total of 2^{NT} possible microstates, and 2^T possible macrostates, which entails

$$\sum_{\mathbb{N}=1}^{2^T} \Omega(\mathbb{N}; T, \beta, \mathbf{A}, \kappa) = 2^{NT}. \quad (6)$$

Figure 2 shows a macrostates–microstates diagram for the case of a noiseless channel with $N = T = 4$ and equal amplitudes. Because all of the microstates are equiprobable we may directly convert the number of microstates into probability dividing by 2^{NT} . For the case shown in figure 2, the most likely macrostate (that is, the one realized by more microstates than any other) is that which allows six of the sixteen possible inputs to fulfil the VIP rule.

The so-called *basic postulate of statistical mechanics* is the equiprobability of microstates in a closed system, and central to ensemble theory is the assumption that all thermodynamic quantities can be written as an ensemble average of a suitable microscopic observable [15]. We anticipate that \mathbb{N} grows exponentially with T , and therefore the typical number \mathbb{N} of VIP

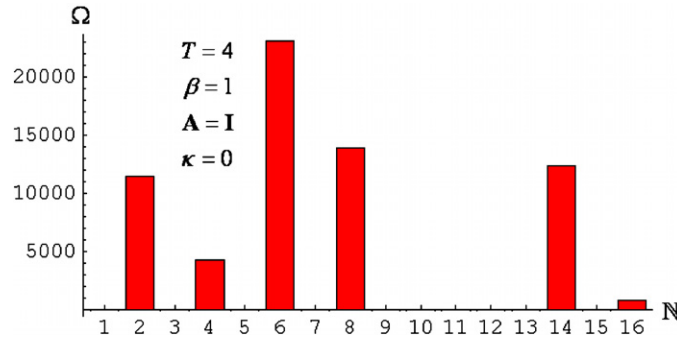


Figure 2. Macrostates–microstates diagram for a noiseless channel represented by a binary matrix, with four transmitters and four receivers. The transmitters all employ equal amplitudes. The most likely macrostate is that which allows six of the sixteen possible inputs to fulfil the VIP constraint.

vectors in the large system is given by the extensive average $\exp \overline{\ln \mathbb{N}}$, where the overbar denotes the configurational average with respect to all channel microstates \mathbf{S} . In the following, the assumption is made that as the dimensions of the system go to infinity at constant ratio ($T, N \rightarrow \infty; T/N = \beta$), the microstate (probability) distribution converges to its own average and fluctuations away from the most likely macrostate vanish⁵. This entails that the distribution of \mathbb{N} is largely peaked around $\overline{\mathbb{N}}$, which allows us to simplify the problem by taking the so-called *annealed approximation* and directly performing the average of \mathbb{N} over the quenched channel states \mathbf{S} . This assumption is not uncommon in the study of various spin and neural systems exhibiting quenched interactions (see, e.g., [5, 6, 16–22]). The annealed approximation is for now accepted, and its validity will be further addressed in section 6, where it will be discussed and tested against finite-size simulations. The self-averaging assumption allows us to write

$$\overline{\mathbb{N}(\mathbf{S}; T, \beta, \mathbf{A}, \kappa)} \xrightarrow{T \rightarrow \infty} \mathcal{N}(T, \beta, \mathbf{A}, \kappa) \equiv \overline{\mathbb{N}(\mathbf{S}; T, \beta, \mathbf{A}, \kappa)}, \quad (7)$$

where \mathcal{N} is (an upper bound on) the typical number of valid inputs. \mathcal{N} may be rewritten as follows:

$$\mathcal{N}(T, \beta, \mathbf{A}, \kappa) = \sum_{\mathbf{i}} \int_{((\kappa-1)\beta^{-1}, \infty)^T} d^T \lambda \overline{\delta\{T^{-1} \mathbf{S}^\dagger \mathbf{S} \mathbf{A} \mathbf{i} - (\beta^{-1} \mathbf{I} + \mathbf{\Lambda}) \mathbf{i}\}}. \quad (8)$$

4. Information capacity

As discussed in section 3, it is assumed that all channel states will exhibit similar behaviour as the channel dimensions become larger. This behaviour is characterized by the number \mathcal{N} of inputs that meet the VIP criterion (2). The number of VIP symbols in the large-system limit may be found using the asymptotic mathematical trickery shown in appendix A (subsection A.1); the resulting expression is

$$\mathcal{N}(T, \beta, \mathbf{A}, \kappa) = \frac{2^T T^2}{4\pi^2 \beta^2} \exp \left\{ \frac{T}{\beta} g(\beta, \mathbf{A}, \kappa) \right\}, \quad (9)$$

where

$$g(\beta, \mathbf{A}, \kappa) = \chi + \frac{1}{2} + \frac{\chi^2 \overline{U}}{2\phi} + \frac{1}{2} \ln \frac{\phi}{\overline{U}} + \beta \int_0^\infty dU \psi(U) \ln Q(\sigma), \quad (10)$$

⁵ This dominant macrostate is equivalent to what in information theory is known as the set of typical sequences.

and the function $\psi(U)$ is the probability density function for the power of the T transmitters (squared eigenvalues of \mathbf{A}). The quantities ϕ and χ are the simultaneous solutions to the following coupled equations:

$$\left. \begin{aligned} 1 + \frac{\chi \bar{U}}{\phi} + \sqrt{\frac{\beta}{\phi}} \int_0^\infty dU \psi(U) \sqrt{U} \frac{Q'(\sigma)}{Q(\sigma)} &= 0 \\ 1 - \frac{\chi^2 \bar{U}}{\phi} - \beta \int_0^\infty dU \psi(U) \frac{Q'(\sigma)}{Q(\sigma)} \sigma &= 0 \end{aligned} \right\}, \quad (11)$$

where $Q(\sigma) = \frac{1}{\sqrt{2\pi}} \int_\sigma^\infty dt \exp(-\frac{t^2}{2})$ and $\sigma = \frac{\kappa + \chi \sqrt{\bar{U}}}{\sqrt{\beta \phi}}$.

As all inputs eligible for reliable transmission are *a priori* equiprobable, the uncertainty function for the set of valid inputs at the transmitter is equivalent to the Boltzmann entropy. The uncertainty function is known in information theory as Shannon entropy, and in the VIP channel the Shannon entropy at the transmitter directly yields the Shannon capacity, which is the theoretical upper bound on the information transfer rate [3]. The Shannon capacity of an asymptotic VIP channel is given in bits/symbol/transmitter as

$$C_\infty(\beta, \mathbf{A}, \kappa) = \lim_{T \rightarrow \infty} \frac{\log_2 \mathcal{N}(T, \beta, \mathbf{A}, \kappa)}{T}. \quad (12)$$

5. Results

5.1. Effect of the amplitude distribution on the asymptotic capacity of the VIP channel

In order to evaluate the performance of the VIP channel when the transmitters send their information with unequal amplitudes, power is assigned to them according to a chi-square distribution, which is often used to model multipath diversity in multiple-access channels [23]. The power distribution function $\psi(U)$ has variance $2/r$ and average power \bar{U} :

$$\psi(U; r, \bar{U}) = \frac{r}{\bar{U}} \frac{(rU/\bar{U})^{r/2-1}}{2^{r/2} \Gamma(r/2)} \exp\left(-\frac{rU}{2\bar{U}}\right), \quad (13)$$

where $\Gamma(\cdot)$ is the Euler gamma function. We find that, regardless of the value of κ , the number of VIP inputs \mathcal{N} is maximized by a power distribution which assigns equal power to each of the T transmitters. This result is in contrast to the case of linear multiuser receivers, where equal power interference is the worst interference case scenario [2]. Figure 3 shows how the Shannon capacity of the channel is affected by the power distribution among the T transmitters.

5.2. Noise effects on the asymptotic capacity of the VIP channel

Figure 4 displays the asymptotic capacity C_∞ versus the channel load β for different noise bounds κ . We observe that, for noise bounds not greater than the transmitter's amplitude (henceforth taken to be 1 for all T transmitters), the channel capacity decreases with the channel load and the noise bound. An interesting phenomenon is that when the noise bound is sufficiently large, the multiuser interference becomes beneficial, as it may neutralize part of the noise. Figure 5 shows a plot of the asymptotic capacity C_∞ versus the noise threshold κ for different channel loads. Two regimes may be identified: for noise thresholds smaller than the transmitting amplitude, an increase in the channel load results in a lower capacity; on the other hand, as the noise threshold becomes slightly larger than 1, multiuser interference plays a constructive role and, as we can see in figure 6, the optimal channel load may adopt a non-trivial value. This indicates that, in the right conditions, the channel capacity per transmitter

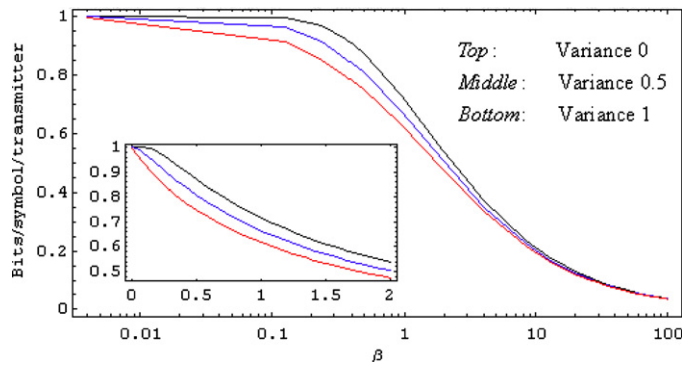


Figure 3. Asymptotic capacity C_∞ as a function of the load β for different distributions of power among the transmitters. The curves shown are for a noiseless ($\kappa = 0$) system and they correspond to different chi-square power distributions with average 1. The inset represents a linear zoom into the $\beta = 0$ to $\beta = 2$ region.

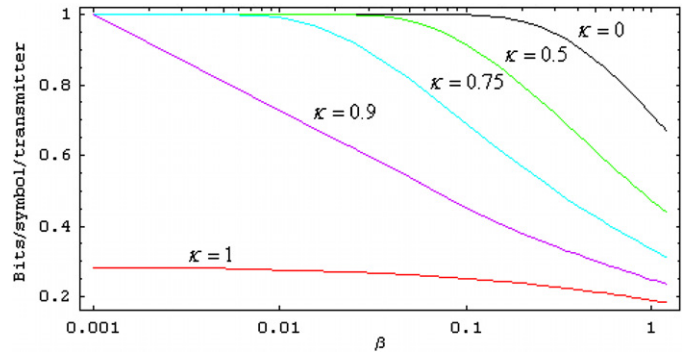


Figure 4. Asymptotic capacity C_∞ versus channel load for different noise bounds. The T transmitters transmit with unit amplitude.

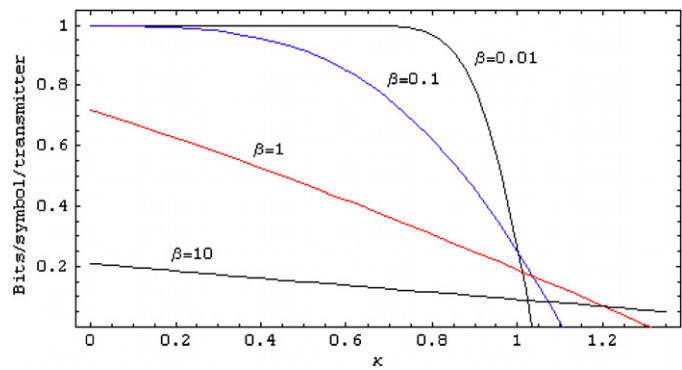


Figure 5. Asymptotic capacity C_∞ versus the noise threshold for different channel loads. The T transmitters transmit with unit amplitude.

may be increased by either adding more transmitters and/or by turning off some receivers. This effect may also appear for noise bounds smaller than 1 when the power distribution is not

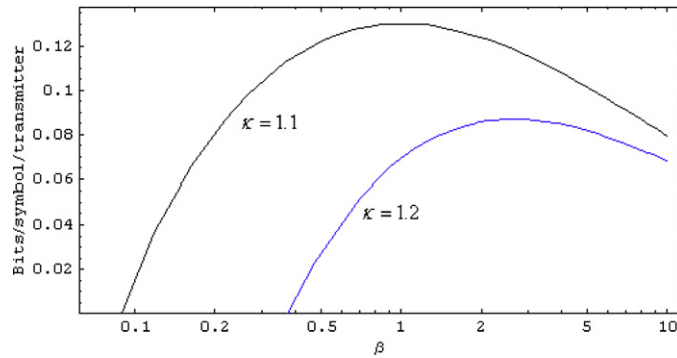


Figure 6. Asymptotic capacity C_∞ versus channel load for noise bounds greater than the transmitter's amplitude (taken to be 1 for all transmitters).

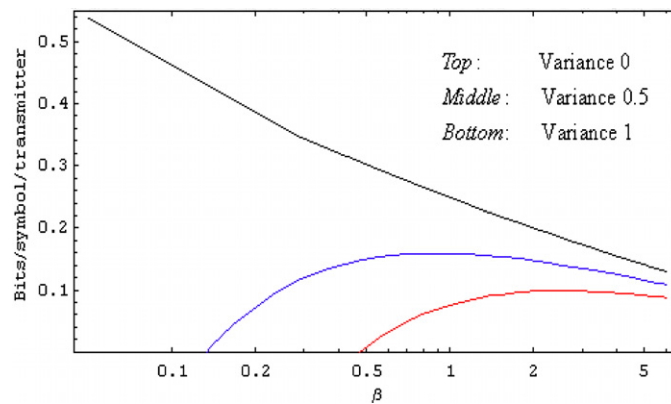


Figure 7. Asymptotic capacity C_∞ as a function of the load β for different distributions of power among the transmitters. The curves shown are for a noisy ($\kappa = 0.9$) system and they correspond to different chi-square power distributions with average 1.

optimal, as shown in figure 7. In figure 5 we can also observe how, depending on the channel load, when the noise threshold is sufficiently large, reliable communication in the simple VIP channel becomes infeasible at any rate.

6. Discussion

6.1. Multiaccess interference and cooperation in the VIP channel

Although the proposed channel model entails low complexity at the receiver end, it requires strict cooperation among transmitters to carry out the vector input preselection. As shown in figure 3 and contrary to linear multiuser receivers [2], this cooperation is best served when all transmitters send with identical amplitude. This indicates that, in this joint coding scheme, all transmitters are equally important when it comes to information and cooperative interference. While at moderate loads the effect of multiuser cooperation in the channel capacity is most significant, the influence of their power distribution disappears for vanishing channel loads, i.e. when multiaccess interference disappears. Conversely, as the channel load becomes very

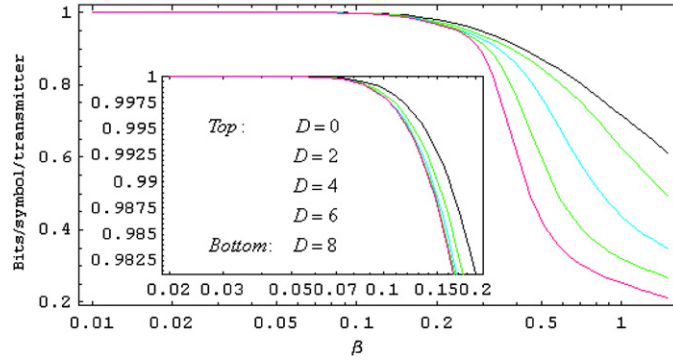


Figure 8. Asymptotic capacity C_∞ versus channel load for different amounts of linear parallel interference cancellation stages. The inset represents a zoom into the $\beta = 0.02$ to $\beta = 0.2$ region.

large, the adverse effects of multiaccess interference are overwhelming and any multiuser cooperation is hopeless.

Although increasing the load β of the channel is generally detrimental to the performance bounds, we observe that, for sufficiently large noise bounds, it may be beneficial to have a greater channel load as shown in figure 6. For instance, when the noise bound κ equals 1.1, bringing the channel load β from 0.1 to 1 increases the capacity per transmitter by a factor of 10. This means that turning off 90% of the receivers would increase the total channel capacity by a factor of 10 or, conversely, that if the number of transmitters is multiplied by 10, then the total channel capacity (capacity per transmitter \times number of transmitters) would increase by a factor of 100.

Consider the case of equal amplitudes and no channel noise; then the only factor hindering VIP performance appears to be the multiaccess interference as contained in the channel load β . This suggests, in the spirit of [2], introducing some stages of interference cancellation at the receiver as a strategy to seek improved VIP performance. If we introduce D stages of linear parallel interference cancellation (LPIC), the resulting channel model is

$$\mathbf{o} = \sum_{q=0}^D \left(\mathbf{I} - \frac{1}{N} \mathbf{S}^\dagger \mathbf{S} \right)^q \frac{1}{N} \mathbf{S}^\dagger \mathbf{S} \mathbf{i}, \quad (14)$$

and its associated VIP rule such that (3) yields error-free detection is

$$\mathbf{i} \equiv \text{sgn} \left\{ \sum_{q=0}^D \left(\mathbf{I} - \frac{1}{N} \mathbf{S}^\dagger \mathbf{S} \right)^q \frac{1}{N} \mathbf{S}^\dagger \mathbf{S} \mathbf{i} \right\}. \quad (15)$$

As shown in appendix A (subsection A.2), the number of VIP inputs, and hence the asymptotic capacity of this channel, may be found as a function of the channel load β and the number D of LPIC stages. The surprising result is that performance declines as the number of interference cancellation stages increases. This remarkable behaviour, shown in figure 8, reinforces the indication that, although an increased channel load results in a lower capacity, multiaccess interference itself is a key factor in the performance of the VIP channel. If the transmitters are expected to carry out a vector input preselection, they must be allowed to jointly code their input through interference-cooperation. LPIC neutralizes the constructive role interference takes in the vector input preselection process.

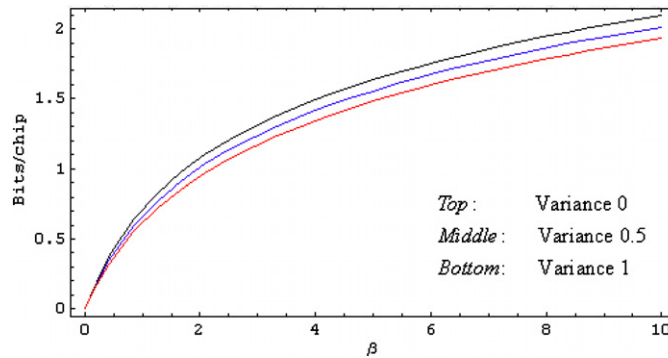


Figure 9. Asymptotic spectral efficiency ϵ_∞ as a function of the load β for different distributions of power among the transmitters. The curves shown are for a noiseless ($\kappa = 0$) system and they correspond to different chi-square power distributions with average 1.

6.2. Efficiency versus information rate in the VIP channel

Both multiple-input/multiple-output (MIMO) and code-division multiple-access (CDMA) channels may be modelled by equation (1). While in the MIMO case the T transmitters represent transmitting antennas and the N receivers correspond to receiving antennas, in the case of CDMA the T transmitters represent synchronous users sending their bits to a base station, and the number N of receivers represents the spreading factor at the base station, i.e. the number of chips per received symbol. Note, in passing, that for heavily overloaded systems (i.e. $\beta = T/N \rightarrow \infty$) the capacity curve decay of the noiseless case coincides with the capacity of the Hopfield model (see equation (12) in [5] for an analytical approximation of the capacity curve at large loads).

In the case of CDMA communication, a useful measure of efficiency is the *spectral efficiency*, which is the total capacity per chip, or the total number of bits per chip that can be transmitted reliably. The asymptotic spectral efficiency ϵ_∞ of the channel is given by the product of the load β and the asymptotic capacity C_∞ :

$$\epsilon_\infty(\beta, \mathbf{A}, \kappa) \equiv \beta C_\infty(\beta, \mathbf{A}, \kappa). \quad (16)$$

Spectral efficiency of matched filter demodulation without user cooperation is known [24] to monotonically increase with the load, but to be upper bounded by 1.44 bits/chip. As shown in appendix B it can be easily verified that the spectral efficiency ϵ_∞ of the cooperative VIP channel is an unbounded monotonically increasing function of β as shown in figure 9. This means that in order to achieve maximum spectral efficiency the load needs to be infinitely large, which in turn means zero capacity or data transmission rate. Conversely, if the channel were to achieve maximum capacity (which is 1 bit per channel use and user) the bandwidth burnt by spreading would be extremely high, and the spectral efficiency zero.

Drawing an analogy between spectral efficiency and *exergy efficiency* allows us to make a connection with irreversible thermodynamics. Just like maximum exergy efficiency (second law efficiency) can only be attained at extremely slow energy conversion rates (cyclical processes), it is only at vanishing data rates that this channel approaches maximum spectral efficiency. This is in synchrony with the inherently irreversible nature of communication processes. The second law of thermodynamics states that in any real (irreversible) process some power will inevitably be dissipated and that it is only in ideal (quasistatic) processes that dissipation vanishes and maximum exergy efficiency can be reached.

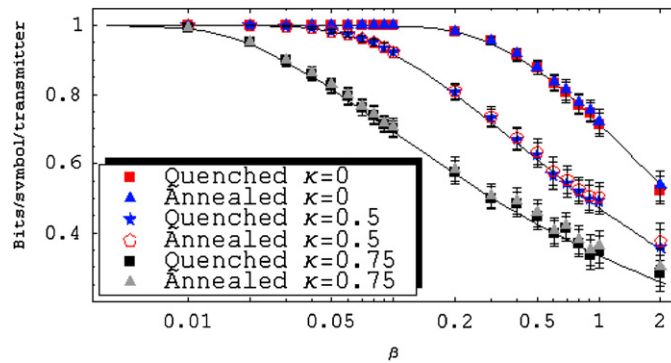


Figure 10. Asymptotic curves obtained by the asymptotic annealed approximation (solid curves) are tested against both annealed ($T^{-1} \log \bar{N}$) and quenched ($T^{-1} \log \bar{N}$) finite-size simulations for $T = 20$ transmitters. The solid curves are for noise bounds $\kappa = 0$ (top), $\kappa = 0.5$ (middle) and $\kappa = 0.75$ (bottom). An ensemble of 100 random binary channel matrices was used in the finite-size simulations.

6.3. The annealed approximation

As mentioned in section 3, as the total number of VIP inputs scales exponentially with T , the relevant quantity to be computed is the quenched average ($\log \bar{N}$). We made use of the annealed approximation and computed instead the annealed average ($\log \bar{N}$) which, as Jensen's inequality tells us, provides an upper bound to the former. We carried out both annealed and quenched finite-size simulations and found that the results obtained with the two methods are essentially identical within error bars for as few as $T = 20$ transmitters. As shown in figure 10, they are also in good agreement with the asymptotic annealed approximation; the agreement appears to weaken for increasing β as decreasing N takes the finite-size simulations further away from the asymptotic limit.

An asymptotic calculation of $T^{-1} \overline{\log \bar{N}}$ would require use of the replica method [25]. Bray and Moore [26, 27] found that for infinite range spin glasses both $T^{-1} \log \bar{N}$ and $T^{-1} \log \bar{N}$ (with \bar{N} being the number of metastable states within an infinitesimal energy band) become identically equal for vanishing off-diagonal order parameters in replica space. They also conducted a stability analysis and found that these diagonal solutions are indeed locally stable, which suggests that the number of metastable states is itself self-averaging. The agreement among both finite-size averages in the VIP channel, as well as between them and the asymptotic approximation, suggests that a similar conclusion applies to the number of VIP inputs given by equation (2) or at least that the annealed approximation tightly upper bounds the information capacity of the channel.

7. Conclusion

We have examined the asymptotic capacity of a vector channel under a strict input preselection routine. With this routine, the complexity of detection is partially outsourced from the receivers back to the transmitters, which must observe a rigorous interference-based multiaccess cooperation scheme. In order to validate the analytically derived results, we evaluated the capacity of a noisy VIP channel with a large yet finite number T of transmitters. Using exhaustive search simulations we verified that convergence is fast and the asymptotic approximation reasonable for as few as 20 transmitters. Determining the explicit VIP

transmissions in a diagrammatic manner (rather than via brute-force enumeration, which becomes infeasible for large T) remains an interesting open research question.

Acknowledgment

The authors wish to thank the anonymous reviewers for their helpful comments and suggestions. This work was partially supported by the Research Council of Norway (Grant 171 133) and the Israel Science Foundation (Grant 296/03). OS gratefully acknowledges the support of the Advanced Communication Center (ACC) and the Yitzhak and Chaya Weinstein Research Institute for Signal Processing, Tel Aviv University.

Appendix A. Counting MIMO-VIP inputs in the self-averaging limit

Consider the VIP channel (1):

$$\boldsymbol{\phi} = \frac{1}{N} \mathbf{S}^\dagger \mathbf{S} \mathbf{A} \mathbf{i} + \mathbf{n}. \quad (\text{A.1})$$

If we introduce D stages of linear parallel interference cancellation (LPIC) at the receiver, the channel becomes

$$\boldsymbol{\phi} = \sum_{q=0}^D \left(\mathbf{I} - \frac{1}{N} \mathbf{S}^\dagger \mathbf{S} \right)^q \frac{1}{N} \mathbf{S}^\dagger \mathbf{S} \mathbf{i} + \mathbf{n}, \quad (\text{A.2})$$

and its associated VIP rule is

$$\mathbf{i} \equiv \text{sgn} \left\{ \sum_{q=0}^D \left(\mathbf{I} - \frac{1}{N} \mathbf{S}^\dagger \mathbf{S} \right)^q \frac{1}{N} \mathbf{S}^\dagger \mathbf{S} \mathbf{i} + \mathbf{n} \right\}. \quad (\text{A.3})$$

Note that if we let $D = 0$ (no interference mitigation) we recover the original VIP rule (2) from section 2. The VIP rule (A.3) may be restated as follows:

$n(\mathbf{S}; \mathbf{i}; T, \beta, \mathbf{A}, \kappa, D)$

$$\equiv \int_{((\kappa-1)\beta^{-1}, \infty)^T} d^T \lambda \delta \left\{ T^{-1} \sum_{q=0}^D \left(\mathbf{I} - \frac{1}{N} \mathbf{S}^\dagger \mathbf{S} \mathbf{A} \right)^q \mathbf{S}^\dagger \mathbf{S} \mathbf{A} \mathbf{i} - (\beta^{-1} \mathbf{I} + \boldsymbol{\Lambda}) \mathbf{i} \right\} = 1, \quad (\text{A.4})$$

where the T variables of integration are the eigenvalues of the diagonal matrix $\boldsymbol{\Lambda}$. The total number \mathbb{N} of VIP inputs is obtained by summing over all possibilities:

$$\mathbb{N}(\mathbf{S}; T, \beta, \mathbf{A}, \kappa, D) = \sum_{\mathbf{i}} n(\mathbf{S}; \mathbf{i}; T, \beta, \mathbf{A}, \kappa, D). \quad (\text{A.5})$$

As discussed in section 3, as the dimensions of the system go to infinity at constant ratio ($T, N \rightarrow \infty; T/N = \beta$) then the self-averaging property holds, and we may write

$$\mathbb{N}(\mathbf{S}; T, \beta, \mathbf{A}, \kappa, D) \xrightarrow{T \rightarrow \infty} \mathcal{N}(T, \beta, \mathbf{A}, \kappa, D) \equiv \overline{\mathbb{N}(\mathbf{S}; T, \beta, \mathbf{A}, \kappa, D)}, \quad (\text{A.6})$$

where the overbar denotes the configurational average with respect to the channel matrix \mathbf{S} . Then, in the large-system limit equation (A.5) may be rewritten as follows:

$$\mathcal{N}(T, \beta, \mathbf{A}, \kappa, D) = \sum_{\mathbf{i}} \overline{n(\mathbf{S}; \mathbf{i}; T, \beta, \mathbf{A}, \kappa, D)}. \quad (\text{A.7})$$

We may rewrite the δ -function in (A.4) as a Fourier sum, resulting in

$$\begin{aligned} \mathcal{N}(T, \beta, \mathbf{A}, \kappa, D) &= \frac{1}{(2\pi)^T} \int_{((\kappa-1)\beta^{-1}, \infty)^T} d^T \lambda \int_{\mathbb{R}^T} d^T \omega \\ &\quad \times \sum_{\mathbf{i}} \overline{\exp\{j\beta^{-1}\boldsymbol{\omega}^\dagger \mathbf{i}\} \exp\{j\boldsymbol{\omega}^\dagger \Lambda \mathbf{i}\} \exp\left\{-jT^{-1} \sum_{q=0}^D \boldsymbol{\omega}^\dagger \left(\mathbf{I} - \frac{1}{N} \mathbf{S}^\dagger \mathbf{S} \mathbf{A}\right)^q \mathbf{S}^\dagger \mathbf{S} \mathbf{A} \mathbf{i}\right\}}. \end{aligned} \quad (\text{A.8})$$

The integrals over the T components of $\boldsymbol{\omega}$ run from $-\infty$ to ∞ . We may then multiply all the components in $\boldsymbol{\omega}$ by the corresponding component in \mathbf{i} without altering the result (recall that the components of \mathbf{i} are unit-amplitude binary). Hence we may rewrite (A.8) as follows:

$$\mathcal{N}(T, \beta, \mathbf{A}, \kappa, D) = \frac{1}{(2\pi)^T} \int_{((\kappa-1)\beta^{-1}, \infty)^T} d^T \lambda \int_{\mathbb{R}^T} d^T \omega \exp\{j\beta^{-1}\boldsymbol{\omega}^\dagger \mathbf{1}\} \exp\{j\boldsymbol{\omega}^\dagger \Lambda \mathbf{1}\} \sum_{\mathbf{i}} E_{\mathbf{i}}, \quad (\text{A.9})$$

where $\mathbf{1}$ is the all-1 column vector, $\hat{\boldsymbol{\omega}}$ is a diagonal matrix such that $\boldsymbol{\omega} \equiv \hat{\boldsymbol{\omega}} \mathbf{1}$, and

$$E_{\mathbf{i}} \equiv \exp\left\{-jT^{-1} \sum_{q=0}^D \mathbf{i}^\dagger \hat{\boldsymbol{\omega}} \left(\mathbf{I} - \frac{1}{N} \mathbf{S}^\dagger \mathbf{S} \mathbf{A}\right)^q \mathbf{S}^\dagger \mathbf{S} \mathbf{A} \mathbf{i}\right\}. \quad (\text{A.10})$$

In order to deal with the expectation in (A.9), the following transform will be useful:

$$\begin{aligned} \exp\left\{-\frac{j}{T} \mathbf{C}^\dagger \mathbf{B}\right\} &= \frac{1}{(2\pi/T)^{\dim \mathbf{B}}} \int_{\mathbb{R}^{\dim \mathbf{B}}} d^{\dim \mathbf{B}} c \int_{\mathbb{R}^{\dim \mathbf{B}}} d^{\dim \mathbf{B}} b \\ &\quad \times \exp\left\{j\frac{T}{2} \mathbf{c}^\dagger \mathbf{c}\right\} \exp\left\{-j\frac{T}{2} \mathbf{b}^\dagger \mathbf{b}\right\} \cos\left\{\frac{\mathbf{c}^\dagger \mathbf{C} + \mathbf{b}^\dagger \mathbf{C} + \mathbf{c}^\dagger \mathbf{B} - \mathbf{b}^\dagger \mathbf{B}}{\sqrt{2}}\right\}. \end{aligned} \quad (\text{A.11})$$

If we let $\mathbf{C} \equiv \sum_{q=0}^D \mathbf{S}(\mathbf{I} - \mathbf{S}^\dagger \mathbf{S} \mathbf{A}/N)^q \hat{\boldsymbol{\omega}} \mathbf{i}$ and $\mathbf{B} \equiv \mathbf{S} \mathbf{A} \mathbf{i}$, then we may apply the transform (A.11) to the expectation in (A.10):

$$\begin{aligned} E_{\mathbf{i}} &\equiv \overline{\exp\left\{-jT^{-1} \sum_{q=0}^D \mathbf{i}^\dagger \hat{\boldsymbol{\omega}} \left(\mathbf{I} - \frac{1}{N} \mathbf{S}^\dagger \mathbf{S} \mathbf{A}\right)^q \mathbf{S}^\dagger \mathbf{S} \mathbf{A} \mathbf{i}\right\}} \\ &= \frac{1}{(2\pi/T)^{T\beta^{-1}}} \int_{\mathbb{R}^{T\beta^{-1}}} d^{T\beta^{-1}} c \int_{\mathbb{R}^{T\beta^{-1}}} d^{T\beta^{-1}} b \exp\left\{j\frac{T}{2} \mathbf{c}^\dagger \mathbf{c}\right\} \exp\left\{-j\frac{T}{2} \mathbf{b}^\dagger \mathbf{b}\right\} F_{\mathbf{i}}, \end{aligned} \quad (\text{A.12})$$

where

$$F_{\mathbf{i}} \equiv \overline{\cos\left\{\frac{\sum_{q=0}^D (\mathbf{c}^\dagger + \mathbf{b}^\dagger) \mathbf{S}(\mathbf{I} - \mathbf{S}^\dagger \mathbf{S} \mathbf{A}/N)^q \hat{\boldsymbol{\omega}} \mathbf{i} + (\mathbf{c}^\dagger - \mathbf{b}^\dagger) \mathbf{S} \mathbf{A} \mathbf{i}}{\sqrt{2}}\right\}}. \quad (\text{A.13})$$

Now two considerations may be made:

- (1) In order for the integrals in (A.12) to converge, the components of \mathbf{b} and \mathbf{c} must be of order $T^{-1/2}$. Hence, in the large T limit any powers of \mathbf{b} and \mathbf{c} larger than quadratic may be neglected.
- (2) Up to second order in x , the following equalities hold:

$$\overline{\cos x} = \overline{\exp(-x^2/2)} = \exp(-x^2/2). \quad (\text{A.14})$$

Taking these two considerations into account we may write

$$F_i = \exp \left\{ - \left(\frac{\sum_{q=0}^D (\mathbf{c}^\dagger + \mathbf{b}^\dagger) \mathbf{S} (\mathbf{I} - \mathbf{S}^\dagger \mathbf{S} \mathbf{A} / N)^q \hat{\boldsymbol{\omega}} \mathbf{i} + (\mathbf{c}^\dagger - \mathbf{b}^\dagger) \mathbf{S} \mathbf{A} \mathbf{i}}{2} \right)^2 \right\}. \tag{A.15}$$

The following results from random matrix theory [28] will be useful to evaluate the expectation (A.15):

$$(\mathbf{I} - \mathbf{S}^\dagger \mathbf{S} / N)^q = \sum_{\xi=0}^q \binom{q}{\xi} \left(\frac{-1}{N} \right)^\xi (\mathbf{S}^\dagger \mathbf{S})^\xi, \tag{A.16}$$

$$\overline{\mathbf{S} \mathbf{a} \mathbf{b}^\dagger \mathbf{S}^\dagger} = \mathbf{a}^\dagger \mathbf{b} \mathbf{I}, \tag{A.17}$$

$$\overline{\mathbf{S} (\mathbf{S}^\dagger \mathbf{S})^\xi \mathbf{a} \mathbf{b}^\dagger (\mathbf{S}^\dagger \mathbf{S})^\eta \mathbf{S}^\dagger} = f(T, \beta, \xi + \eta) \mathbf{a}^\dagger \mathbf{b} \mathbf{I}, \tag{A.18}$$

where f is a scalar function defined as follows:

$$f(T, \beta, \xi) \mathbf{I} \equiv \frac{1}{T} \overline{\mathbf{S} (\mathbf{S}^\dagger \mathbf{S})^\xi \mathbf{S}^\dagger} = T^\xi \left(\frac{\overline{\mathbf{S} \mathbf{S}^\dagger}}{T} \right)^{\xi+1}. \tag{A.19}$$

In the infinite matrix limit, the function f is given by [28]

$$f(T, \beta, \xi) = \frac{\beta^{-\xi-1} T^\xi}{\xi + 1} \sum_{i=1}^{\xi+1} \binom{\xi+1}{i} \binom{\xi+1}{i-1} \beta^i. \tag{A.20}$$

A.1. Power distribution in the VIP channel

In this section, we will analyse the case of no interference cancellation, that is $D = 0$. We do, however, allow for an arbitrary amplitude distribution.

Inserting (A.17) into (A.15) yields

$$\begin{aligned} F_i &= \exp \left\{ - \left(\frac{(\mathbf{c}^\dagger + \mathbf{b}^\dagger) \mathbf{S} \hat{\boldsymbol{\omega}} \mathbf{i} + (\mathbf{c}^\dagger - \mathbf{b}^\dagger) \mathbf{S} \mathbf{A} \mathbf{i}}{2} \right)^2 \right\} \\ &= \exp \left\{ - \frac{1}{4} \{ \boldsymbol{\omega}^\dagger \boldsymbol{\omega} (\mathbf{c}^\dagger + \mathbf{b}^\dagger) (\mathbf{c} + \mathbf{b}) + T \bar{U} (\mathbf{c}^\dagger - \mathbf{b}^\dagger) (\mathbf{c} - \mathbf{b}) + 2 \boldsymbol{\omega}^\dagger \mathbf{A} \mathbf{1} (\mathbf{c}^\dagger \mathbf{c} - \mathbf{b}^\dagger \mathbf{b}) \} \right\}, \end{aligned} \tag{A.21}$$

where \bar{U} is the average power (squared amplitude) per transmitter. Inserting the expression for F_i back into (A.12) we obtain

$$\begin{aligned} E_i &\equiv \exp \left\{ -j T^{-1} \sum_{q=0}^0 \mathbf{i}^\dagger \hat{\boldsymbol{\omega}} \left(\mathbf{I} - \frac{1}{N} \mathbf{S}^\dagger \mathbf{S} \mathbf{A} \right)^q \mathbf{S}^\dagger \mathbf{S} \mathbf{A} \mathbf{i} \right\} \\ &= \frac{1}{(2\pi/T)^{T\beta-1}} \int_{\mathbb{R}^{T\beta-1}} d^{T\beta-1} c \int_{\mathbb{R}^{T\beta-1}} d^{T\beta-1} b \exp \left\{ j \frac{T}{2} \mathbf{c}^\dagger \mathbf{c} \right\} \exp \left\{ -j \frac{T}{2} \mathbf{b}^\dagger \mathbf{b} \right\} \\ &\quad \times \exp \left\{ - \frac{1}{4} \{ \boldsymbol{\omega}^\dagger \boldsymbol{\omega} (\mathbf{c}^\dagger + \mathbf{b}^\dagger) (\mathbf{c} + \mathbf{b}) + T \bar{U} (\mathbf{c}^\dagger - \mathbf{b}^\dagger) (\mathbf{c} - \mathbf{b}) + 2 \boldsymbol{\omega}^\dagger \mathbf{A} \mathbf{1} (\mathbf{c}^\dagger \mathbf{c} - \mathbf{b}^\dagger \mathbf{b}) \} \right\}. \end{aligned} \tag{A.22}$$

The following two unity expressions may be inserted into (A.22) without altering its result:

$$\begin{aligned} & \beta^{-1} \int_{-\infty}^{\infty} d\phi \delta \left\{ \phi \beta^{-1} - \frac{1}{2}(\mathbf{c}^\dagger + \mathbf{b}^\dagger)(\mathbf{c} + \mathbf{b}) \right\} \\ &= T\beta^{-1} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{d\phi d\Phi}{2\pi} \exp \left\{ jT\Phi \left(\phi \beta^{-1} - \frac{1}{2}(\mathbf{c}^\dagger + \mathbf{b}^\dagger)(\mathbf{c} + \mathbf{b}) \right) \right\} = 1, \quad (\text{A.23}) \end{aligned}$$

$$\begin{aligned} & \beta^{-1} \int_{-\infty}^{\infty} d\chi \delta \left\{ \beta^{-1}\chi - \frac{j}{2}(\mathbf{c}^\dagger \mathbf{c} - \mathbf{b}^\dagger \mathbf{b}) \right\} \\ &= T\beta^{-1} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{d\chi dP}{2\pi} \exp \left\{ jTP \left(\beta^{-1}\chi - \frac{j}{2}(\mathbf{c}^\dagger \mathbf{c} - \mathbf{b}^\dagger \mathbf{b}) \right) \right\} = 1. \quad (\text{A.24}) \end{aligned}$$

Inserting (A.23) and (A.24) into (A.22) and rewriting everything in terms of ϕ and χ as dictated by the delta functions yields

$$\begin{aligned} E_i &= \frac{T^2 \beta^{-2}}{(2\pi/T)^{T\beta^{-1}} (2\pi)^2} \int_{\mathbb{R}^{T\beta^{-1}}} d^{T\beta^{-1}} c \int_{\mathbb{R}^{T\beta^{-1}}} d^{T\beta^{-1}} b \int_{-\infty}^{\infty} d\phi \int_{-\infty}^{\infty} d\Phi \int_{-\infty}^{\infty} d\chi \int_{-\infty}^{\infty} dP \\ &\quad \times \exp\{T\beta^{-1}\chi\} \exp\{jT\Phi\phi\beta^{-1}\} \exp\{jTP\beta^{-1}\chi\} \\ &\quad \times \exp\{j\beta^{-1}\chi\boldsymbol{\omega}^\dagger \mathbf{A}\mathbf{1}\} \exp\left\{-\frac{1}{2}\phi\boldsymbol{\omega}^\dagger \boldsymbol{\omega}\right\} \exp\left\{\frac{TP}{2}(\mathbf{c}^\dagger \mathbf{c} - \mathbf{b}^\dagger \mathbf{b})\right\} \\ &\quad \times \exp\left\{-j\frac{T\Phi}{2}(\mathbf{c}^\dagger + \mathbf{b}^\dagger)(\mathbf{c} + \mathbf{b})\right\} \exp\left\{-\frac{T\bar{U}}{4}(\mathbf{c}^\dagger - \mathbf{b}^\dagger)(\mathbf{c} - \mathbf{b})\right\}. \quad (\text{A.25}) \end{aligned}$$

The \mathbf{c} and \mathbf{b} integrals may be performed analytically, and (A.25) becomes

$$\begin{aligned} E_i &= \frac{T^2 \beta^{-2}}{(2\pi/T)^{T\beta^{-1}} (2\pi)^2} \int_{-\infty}^{\infty} d\phi \int_{-\infty}^{\infty} d\Phi \int_{-\infty}^{\infty} d\chi \int_{-\infty}^{\infty} dP \\ &\quad \times \exp\{T\beta^{-1}\chi\} \exp\{j\beta^{-1}\chi\boldsymbol{\omega}^\dagger \mathbf{A}\mathbf{1}\} \exp\left\{-\frac{1}{2}\phi\beta^{-1}\boldsymbol{\omega}^\dagger \boldsymbol{\omega}\right\} \\ &\quad \times \exp\left\{T\beta^{-1} \left[j\Phi\phi + jP\chi + \ln \frac{2\pi/T}{\sqrt{2j\Phi - 2P + \bar{U}} \sqrt{\frac{jP^2 + 2\Phi\bar{U}}{2\Phi + 2jP - j\bar{U}}}} \right] \right\}. \quad (\text{A.26}) \end{aligned}$$

In the large T limit, the P and Φ integrals are dominated by the term for which the value of the exponent is largest. That term is given by $\Phi = -j(\phi - \chi^2\bar{U})/2\phi^2$ and $P = -j\chi\bar{U}/\phi$. Inserting these expressions into (A.26) we may deem the P and Φ variables as integrated, yielding

$$\begin{aligned} E_i &= 2^{-2} \pi^{-2} T^2 \beta^{-2} \bar{U}^{-T\beta^{-1}/2} \int_{-\infty}^{\infty} d\phi \int_{-\infty}^{\infty} d\chi \exp\{j\beta^{-1}\chi\boldsymbol{\omega}^\dagger \mathbf{A}\mathbf{1}\} \exp\left\{-\frac{1}{2}\phi\beta^{-1}\boldsymbol{\omega}^\dagger \boldsymbol{\omega}\right\} \\ &\quad \times \exp\left\{T\beta^{-1} \left[\chi + \frac{1}{2} + \frac{\chi^2\bar{U}}{2\phi} + \frac{1}{2} \ln \phi \right] \right\}. \quad (\text{A.27}) \end{aligned}$$

Inserting (A.27) back into (A.9) we obtain the following expression for the number of allowed inputs:

$$\begin{aligned} \mathcal{N}(T, \beta, \mathbf{A}, \kappa) &= 2^{-2} \pi^{-2-T} T^2 \beta^{-2} \int_{-\infty}^{\infty} d\phi \int_{-\infty}^{\infty} d\chi \int_{\mathbb{R}^T} d^T \omega \int_{((\kappa-1)\beta^{-1}, \infty)^T} d^T \lambda \\ &\quad \times \exp\{j\beta^{-1} \boldsymbol{\omega}^\dagger \mathbf{1}\} \exp\{j\boldsymbol{\omega}^\dagger \boldsymbol{\Lambda} \mathbf{1}\} \exp\{j\beta^{-1} \chi \boldsymbol{\omega}^\dagger \mathbf{A} \mathbf{1}\} \exp\left\{-\frac{1}{2} \phi \beta^{-1} \boldsymbol{\omega}^\dagger \boldsymbol{\omega}\right\} \\ &\quad \times \exp\left\{T\beta^{-1} \left[\chi + \frac{1}{2} + \frac{\chi^2 \bar{U}}{2\phi} + \frac{1}{2} \ln \frac{\phi}{U}\right]\right\}. \end{aligned} \quad (\text{A.28})$$

After performing the T $\boldsymbol{\omega}$ -integrals, expression (A.28) becomes

$$\begin{aligned} \mathcal{N}(T, \beta, \mathbf{A}, \kappa) &= 2^{T/2-2} \pi^{-2-T/2} T^2 \beta^{-2+T/2} \int_{-\infty}^{\infty} d\phi \int_{-\infty}^{\infty} d\chi \int_{((\kappa-1)\beta^{-1}, \infty)^T} d^T \lambda \\ &\quad \times \phi^{-T/2} \exp\left\{T\beta^{-1} \left[\chi + \frac{1}{2} + \frac{\chi^2 \bar{U}}{2\phi} + \frac{1}{2} \ln \frac{\phi}{U}\right]\right\} \\ &\quad \times \exp\left\{\frac{-\beta}{2\phi} (\beta^{-1} \mathbf{1}^\dagger + \beta^{-1} \chi \mathbf{1}^\dagger \mathbf{A} + \mathbf{1}^\dagger \boldsymbol{\Lambda}) (\beta^{-1} \mathbf{1} + \beta^{-1} \chi \mathbf{A} \mathbf{1} + \boldsymbol{\Lambda} \mathbf{1})\right\}. \end{aligned} \quad (\text{A.29})$$

The T λ -integrals may be performed analytically (recall λ are the components of the vector $\boldsymbol{\Lambda} \mathbf{1}$), yielding

$$\begin{aligned} \mathcal{N}(T, \beta, \mathbf{A}, \kappa) &= 2^{T-2} \pi^{-2} T^2 \beta^{-2} \int_{-\infty}^{\infty} d\phi \int_{-\infty}^{\infty} d\chi \\ &\quad \times \exp\left\{T\beta^{-1} \left[\chi + \frac{1}{2} + \frac{\chi^2 \bar{U}}{2\phi} + \frac{1}{2} \ln \frac{\phi}{U} + \frac{1}{\beta^{-1} T} \sum_{t=1}^T \ln Q(\xi_t)\right]\right\}, \end{aligned} \quad (\text{A.30})$$

where

$$Q(\xi_t) \equiv \frac{1}{\sqrt{2\pi}} \int_{\xi_t}^{\infty} dx e^{-\frac{x^2}{2}}, \quad (\text{A.31})$$

$$\xi_t \equiv \frac{\kappa + a_t \chi}{\sqrt{\beta \phi}}, \quad (\text{A.32})$$

and a_t is the amplitude of the t th transmitter, i.e. the t th eigenvalue of \mathbf{A} .

The sum inside the exponential in (A.30) may be viewed as an average over the amplitude (or power) distribution of the T transmitters. Let $\psi(U)$ be the normalized power distribution for the transmitters, then as T goes to infinity we may write

$$\frac{1}{T} \sum_{t=1}^T \ln Q(\xi_t) = \int_0^{\infty} dU \psi(U) \ln \left\{ \frac{1}{\sqrt{2\pi}} \int_{\frac{\kappa + \sqrt{U} \chi}{\sqrt{\beta \phi}}}^{\infty} dx e^{-\frac{x^2}{2}} \right\}. \quad (\text{A.33})$$

And we may rewrite (A.30) as

$$\begin{aligned} \mathcal{N}(T, \beta, \mathbf{A}, \kappa) &= 2^{T-2} \pi^{-2} T^2 \beta^{-2} \int_{-\infty}^{\infty} d\phi \int_{-\infty}^{\infty} d\chi \\ &\quad \times \exp\left\{T\beta^{-1} \left[\chi + \frac{1}{2} + \frac{\chi^2 \bar{U}}{2\phi} + \frac{1}{2} \ln \frac{\phi}{U} + \beta \int_0^{\infty} dU \psi(U) \ln Q(\sigma)\right]\right\}, \end{aligned} \quad (\text{A.34})$$

where

$$\sigma \equiv \frac{\kappa + \sqrt{U}\chi}{\sqrt{\beta\phi}}. \quad (\text{A.35})$$

As T goes to infinity, all that is required to integrate ϕ and χ is maximizing the exponent in (A.34) with respect to both variables. The maximum is given by the simultaneous solution to the equations

$$\left. \begin{aligned} 1 + \frac{\chi\bar{U}}{\phi} + \sqrt{\frac{\beta}{\phi}} \int_0^\infty dU \psi(U) \sqrt{U} \frac{Q'(\sigma)}{Q(\sigma)} &= 0 \\ 1 - \frac{\chi^2\bar{U}}{\phi} - \beta \int_0^\infty dU \psi(U) \frac{Q'(\sigma)}{Q(\sigma)} \sigma &= 0 \end{aligned} \right\}. \quad (\text{A.36})$$

A.2. Linear parallel interference cancellation in the VIP channel

In this section, we restrict the analysis to the case of a uniform amplitude distribution $\mathbf{A} = \mathbf{I}$, but allow for an arbitrary number D of interference cancellation stages. Once we expand the square inside the expectation (A.15), equations (A.16)–(A.18) are the expressions we need to evaluate F_i , which may be rewritten as follows:

$$F_i = \exp \left[-\frac{\Xi_D}{4} (\mathbf{c}^\dagger + \mathbf{b}^\dagger)(\mathbf{c} + \mathbf{b}) \boldsymbol{\omega}^\dagger \boldsymbol{\omega} - \frac{\Psi_D}{2} \boldsymbol{\omega}^\dagger \mathbf{1} (\mathbf{c}^\dagger \mathbf{c} - \mathbf{b}^\dagger \mathbf{b}) - \frac{T}{4} (\mathbf{c}^\dagger - \mathbf{b}^\dagger)(\mathbf{c} - \mathbf{b}) \right], \quad (\text{A.37})$$

where Ξ_D and Ψ_D are defined as

$$\Xi_D(T, \beta) \equiv \sum_{q=0}^D \sum_{\xi=0}^q \sum_{t=0}^D \sum_{\eta=0}^t \binom{q}{\xi} \binom{t}{\eta} \left(\frac{-1}{T\beta^{-1}} \right)^{\xi+\eta} f(T, \beta, \xi + \eta), \quad (\text{A.38})$$

$$\Psi_D(T, \beta) \equiv \sum_{q=0}^D \sum_{\xi=0}^q \binom{q}{\xi} \left(\frac{-1}{T\beta^{-1}} \right)^\xi f(T, \beta, \xi). \quad (\text{A.39})$$

Inserting the expression for F_i back into (A.12) we obtain

$$\begin{aligned} E_i &= \frac{1}{(2\pi/T)^{T\beta^{-1}}} \int_{\mathbb{R}^{T\beta^{-1}}} d^{T\beta^{-1}} c \int_{\mathbb{R}^{T\beta^{-1}}} d^{T\beta^{-1}} b \exp \left\{ j \frac{T}{2} \mathbf{c}^\dagger \mathbf{c} \right\} \exp \left\{ -j \frac{T}{2} \mathbf{b}^\dagger \mathbf{b} \right\} \\ &\quad \times \exp \left[-\frac{\Xi_D}{4} (\mathbf{c}^\dagger + \mathbf{b}^\dagger)(\mathbf{c} + \mathbf{b}) \boldsymbol{\omega}^\dagger \boldsymbol{\omega} - \frac{\Psi_D}{2} \boldsymbol{\omega}^\dagger \mathbf{1} (\mathbf{c}^\dagger \mathbf{c} - \mathbf{b}^\dagger \mathbf{b}) - \frac{T}{4} (\mathbf{c}^\dagger - \mathbf{b}^\dagger)(\mathbf{c} - \mathbf{b}) \right]. \end{aligned} \quad (\text{A.40})$$

Inserting expressions (A.23) and (A.24) into (A.40) yields

$$\begin{aligned} E_i &= \frac{T^2 \beta^{-2}}{(2\pi/T)^{T\beta^{-1}} (2\pi)^2} \int_{\mathbb{R}^{T\beta^{-1}}} d^{T\beta^{-1}} b \int_{\mathbb{R}^{T\beta^{-1}}} d^{T\beta^{-1}} c \int_{-\infty}^\infty d\phi \int_{-\infty}^\infty d\Phi \int_{-\infty}^\infty d\chi \int_{-\infty}^\infty dP \\ &\quad \times \exp\{T\beta^{-1}\chi\} \exp\{jT\Phi\phi\beta^{-1}\} \exp\{jTP\beta^{-1}\chi\} \\ &\quad \times \exp \left\{ -\frac{\Xi_D}{2} \phi\beta^{-1} \boldsymbol{\omega}^\dagger \boldsymbol{\omega} \right\} \exp\{j\Psi_D\beta^{-1}\chi \boldsymbol{\omega}^\dagger \mathbf{1}\} \exp \left\{ \frac{TP}{2} (\mathbf{c}^\dagger \mathbf{c} - \mathbf{b}^\dagger \mathbf{b}) \right\} \\ &\quad \times \exp \left\{ -\frac{T}{4} (\mathbf{c}^\dagger - \mathbf{b}^\dagger)(\mathbf{c} - \mathbf{b}) \right\} \exp \left\{ -j \frac{T\Phi}{2} (\mathbf{c}^\dagger + \mathbf{b}^\dagger)(\mathbf{c} + \mathbf{b}) \right\}. \end{aligned} \quad (\text{A.41})$$

The \mathbf{c} and \mathbf{b} integrals may be performed analytically, and expression (A.41) becomes

$$E_i = \frac{T^2 \beta^{-2}}{(2\pi/T)^{T\beta^{-1}} (2\pi)^2} \int_{-\infty}^{\infty} d\phi \int_{-\infty}^{\infty} d\Phi \int_{-\infty}^{\infty} d\chi \int_{-\infty}^{\infty} dP \\ \times \exp\{T\beta^{-1}\chi\} \exp\left\{-\frac{\Xi_D}{2}\phi\beta^{-1}\boldsymbol{\omega}^\dagger\boldsymbol{\omega}\right\} \exp\{j\Psi_D\beta^{-1}\chi\boldsymbol{\omega}^\dagger\mathbf{1}\} \\ \times \exp\left\{T\beta^{-1}\left(jP\chi + j\Phi\phi + \ln\frac{2\pi/T}{\sqrt{2j\Phi - 2P + 1}\sqrt{\frac{jP^2 + 2\Phi}{2\Phi + 2jP - j}}}\right)\right\}. \quad (\text{A.42})$$

In the large T limit the P and Φ integrals are dominated by the term for which the value of the exponent is largest. That term is given by $\Phi = -j(\phi - \chi^2)/2\phi^2$ and $P = -j\chi/\phi$. Inserting these expressions into (A.42) we may deem the P and Φ variables as integrated, yielding

$$E_i = 2^{-2}\pi^{-2}T^2\beta^{-2} \int_{-\infty}^{\infty} d\phi \int_{-\infty}^{\infty} d\chi \exp\left\{-\frac{\Xi_D}{2}\phi\beta^{-1}\boldsymbol{\omega}^\dagger\boldsymbol{\omega}\right\} \exp\{j\Psi_D\beta^{-1}\chi\boldsymbol{\omega}^\dagger\mathbf{1}\} \\ \times \exp\left\{T\beta^{-1}\left(\chi + \frac{1}{2} + \frac{\chi^2}{2\phi} + \frac{1}{2}\ln\phi\right)\right\}. \quad (\text{A.43})$$

Inserting (A.43) back into (A.9) we obtain the following expression for the number of VIP inputs:

$$\mathcal{N}(T, \beta, \kappa, D) = 2^{-2}\pi^{-2-T}T^2\beta^{-2} \int_{-\infty}^{\infty} d\phi \int_{-\infty}^{\infty} d\chi \int_{\mathbb{R}^T} d^T\boldsymbol{\omega} \int_{((\kappa-1)\beta^{-1}, \infty)^T} d^T\boldsymbol{\lambda} \\ \times \exp\{j\boldsymbol{\omega}^\dagger\boldsymbol{\Lambda}\mathbf{1}\} \exp\left\{-\frac{\Xi_D}{2}\phi\beta^{-1}\boldsymbol{\omega}^\dagger\boldsymbol{\omega}\right\} \exp\{j\Psi_D\beta^{-1}\chi\boldsymbol{\omega}^\dagger\mathbf{1}\} \exp\{j\beta^{-1}\boldsymbol{\omega}^\dagger\mathbf{1}\} \\ \times \exp\left\{T\beta^{-1}\left(\chi + \frac{1}{2} + \frac{\chi^2}{2\phi} + \frac{1}{2}\ln\phi\right)\right\}. \quad (\text{A.44})$$

After performing the T $\boldsymbol{\omega}$ -integrals, expression (A.45) becomes

$$\mathcal{N}(T, \beta, \kappa, D) = 2^{T/2-2}\pi^{-2-T/2}T^2\beta^{-2+T/2}\Xi_D^{-T/2} \int_{-\infty}^{\infty} d\phi \int_{-\infty}^{\infty} d\chi \int_{((\kappa-1)\beta^{-1}, \infty)^T} d^T\boldsymbol{\lambda} \\ \times \exp\left\{\frac{-\beta}{2\phi\Xi_D}(\beta^{-1}(1 + \chi\Psi_D)\mathbf{1}^\dagger + \mathbf{1}^\dagger\boldsymbol{\Lambda})(\beta^{-1}(1 + \chi\Psi_D)\mathbf{1} + \boldsymbol{\Lambda}\mathbf{1})\right\} \\ \times \phi^{-T/2} \exp\left\{T\beta^{-1}\left(\chi + \frac{1}{2} + \frac{\chi^2}{2\phi} + \frac{1}{2}\ln\phi\right)\right\}. \quad (\text{A.45})$$

The T $\boldsymbol{\lambda}$ -integrals may be performed analytically (recall that $\boldsymbol{\lambda}$ are the components of the vector $\boldsymbol{\Lambda}\mathbf{1}$), yielding

$$\mathcal{N}(T, \beta, \kappa, D) = 2^{T-2}\pi^{-2}T^2\beta^{-2} \int_{-\infty}^{\infty} d\phi \int_{-\infty}^{\infty} d\chi \\ \times \exp\left\{\frac{T}{2\beta}\left(2\chi + 1 + \frac{\chi^2}{\phi} + \ln\phi + 2\beta\ln Q(\xi_D)\right)\right\}, \quad (\text{A.46})$$

where

$$\xi_D \equiv (\chi + \kappa)\sqrt{\frac{1}{\phi\beta}\frac{\Psi_D}{\sqrt{\Xi_D}}}. \quad (\text{A.47})$$

As T goes to infinity, all that is required to integrate ϕ and χ is maximizing the exponent in (A.46) with respect to both variables. The maximum is given by the simultaneous solution to the equations

$$\left. \begin{aligned} 1 - \frac{\chi^2}{\phi} - \beta \frac{Q'(\xi_D)}{Q(\xi_D)} \xi_D &= 0 \\ 1 + \frac{\chi}{\phi} + \beta \frac{\xi_D}{\chi + \kappa} \frac{\sqrt{\Xi_D}}{\Psi_D} \frac{Q'(\xi_D)}{Q(\xi_D)} &= 0 \end{aligned} \right\}. \quad (\text{A.48})$$

Appendix B. Spectral efficiency of the CDMA VIP channel

In this section, the unbounded monotonic behaviour of ϵ_∞ as a function of β is shown. Inserting equations (9) and (12) into (16) we obtain

$$\epsilon_\infty(\beta, \mathbf{A}, \kappa) \equiv \beta C_\infty(\beta, \mathbf{A}, \kappa) = \beta + \frac{1}{\ln 2} g(\beta, \mathbf{A}, \kappa). \quad (\text{B.1})$$

For simplicity and without loss of generality, the channel is assumed noiseless ($\kappa = 0$) and the average power per user is assumed to be unity ($\bar{U} = 1$). In this case, the value of χ which solves equation (11) is $\chi = -1$, and equations (10), (11) and (13) simplify to

$$g(\beta, \mathbf{A}) = -\frac{1}{2} + \frac{1}{2\phi} + \frac{1}{2} \ln \phi + \beta \int_0^\infty dU \psi(U) \ln Q(\sigma), \quad (\text{B.2})$$

$$1 - \frac{1}{\phi} - \beta \int_0^\infty dU \psi(U) \frac{Q'(\sigma)}{Q(\sigma)} \sigma = 0, \quad (\text{B.3})$$

$$\psi(U; r) = r \frac{(rU)^{r/2-1}}{2^{r/2} \Gamma(r/2)} \exp\left\{-\frac{rU}{2}\right\}, \quad (\text{B.4})$$

where $Q(\sigma) = \frac{1}{\sqrt{2\pi}} \int_\sigma^\infty dt \exp\{-\frac{t^2}{2}\}$ and $\sigma = \frac{-\sqrt{U}}{\sqrt{\beta\phi}}$.

The asymptotic capacity C_∞ is upper bounded by 1 so it is clear that equation (B.1) approaches zero from above as the channel load β vanishes. Indeed as $\beta \rightarrow 0$ it is verified that ϕ becomes 1 for any value of r , which in turn means that the asymptotic capacity C_∞ approaches 1 and the spectral efficiency ϵ_∞ approaches zero. However, as β gets arbitrarily large the parameter ϕ becomes linearly proportional to β ; it is verified that the constant of proportionality is in the interval $(\frac{4}{\pi^2}, \frac{2}{\pi})$ and it is determined by the power distribution parameter r , which represents the number of degrees of freedom in the chi-square power distribution $\psi(U)$. The constant of proportionality is smallest for $r = 1$ and largest for $r \rightarrow \infty$. As β gets large, equation (B.2) becomes

$$g(\beta \rightarrow \infty, \mathbf{A}) = -\frac{1}{2} + \frac{1}{2\xi\beta} + \frac{1}{2} \ln \xi\beta + \beta \int_0^\infty dU \psi(U) \ln Q(\sigma), \quad (\text{B.5})$$

where $\sigma = \frac{-\sqrt{U}}{\beta\sqrt{\xi}}$ and ξ is a number in the interval $(\frac{4}{\pi^2}, \frac{2}{\pi})$. If (B.5) is inserted into (B.1), the following expansion can be made about $\beta \rightarrow \infty$:

$$\epsilon_\infty(\beta \rightarrow \infty, \mathbf{A}) = \frac{1}{\ln 2} \sqrt{\frac{2}{\pi\xi}} \int_0^\infty dU \psi(U) \sqrt{U} + \frac{\ln\{\beta\xi\} - 1}{2 \ln 2} + \vartheta(\beta^{-1}). \quad (\text{B.6})$$

Expression (B.6) is a monotonous increasing function of β , which means that ϵ_∞ never stops growing with the channel load.

References

- [1] Hochwald B M, Marzetta T L and Hassibi B 2001 Space-time autocoding *IEEE Trans. Inform. Theory* **47** 2761–81
- [2] Müller R R and Verdú S 2001 Design and analysis of low-complexity interference mitigation on vector channels *IEEE J. Select. Areas Commun.* **19** 1429–41
- [3] Cover T M and Thomas J A 1991 *Elements of Information Theory* (New York: Wiley)
- [4] Shental O, Kanter I and Weiss A J 2006 Capacity of complexity-constrained noise-free CDMA *IEEE Commun. Lett.* **10** 10–2
- [5] Gardner E J 1986 Structure of metastable states in the Hopfield model *J. Phys. A: Math. Gen.* **19** L1047–52
- [6] Singh M P 2001 Hopfield model with self-coupling *Phys. Rev. A* **64** 051 912–1–9
- [7] Kechriotis G I and Manolakos E S 1996 Hopfield neural network implementation of the optimal CDMA multiuser detector *IEEE Trans. Neural Netw.* **7** 131–41
- [8] Tanaka T 2001 Analysis of bit error probability of direct-sequence CDMA multiuser demodulators *Advances in Neural Information Processing Systems* vol 13 ed T Leen, T Dietterich and V Tresp (Cambridge, MA: MIT Press) pp 315–21
- [9] Tanaka T 2001 Statistical mechanics of CDMA multiuser demodulation *Europhys. Lett.* **54** 540–6
- [10] Verdú S 1998 *Multiuser Detection* (Cambridge: Cambridge University Press)
- [11] Verdú S and Shamai S (Shitz) 1999 Spectral efficiency of CDMA with random spreading *IEEE Trans. Inform. Theory* **45** 622–40
- [12] Vojčić B R and Jang W M 1998 Transmitter precoding in synchronous multiuser communications *IEEE Trans. Commun.* **46** 1346–55
- [13] Peel C B, Hochwald B M and Swindlehurst A L 2005 A vector-perturbation technique for near-capacity multi-antenna multiuser communication: Part I. Channel inversion and regularization *IEEE Trans. Commun.* **53** 195–202
- [14] Peel C B, Hochwald B M and Swindlehurst A L 2005 A vector-perturbation technique for near-capacity multi-antenna multiuser communication: Part II. Perturbation *IEEE Trans. Commun.* **53** 537–44
- [15] Greiner W, Stocker H and Neise L 2004 *Thermodynamics and Statistical Mechanics (Classical Theoretical Physics)* (Berlin: Springer)
- [16] Tanaka F and Edwards S F 1980 Analytic theory of the ground state properties of a spin glass: I. Ising spin glass *J. Phys. F: Met. Phys.* **10** 2769–78
- [17] Waugh F R, Marcus C M and Westervelt R M 1990 Fixed-point attractors in analog neural computation *Phys. Rev. Lett.* **64**
- [18] Fukai T and Shiino M 1990 Large suppression of spurious states in neural networks of nonlinear analog neurons *Phys. Rev. A* **42** 7459–66
- [19] Korutcheva E 1993 The number of metastable states of a simple perceptron with gradient descent learning algorithm *J. Phys. A: Math. Gen.* **26** L1021–7
- [20] Coolen A C C and Sherrington D 1994 Order-parameter flow in the fully connected Hopfield model near saturation *Phys. Rev. E* **49** 1921–34
- [21] Singh M P, Chengxiang Z and Dasgupta C 1995 Fixed points in a Hopfield model with random asymmetric interactions *Phys. Rev. E* **52** 5261–72
- [22] Bruce A D, Gardner E J and Wallace D J 1987 Dynamics and statistical mechanics of the Hopfield model *J. Phys. A: Math. Gen.* **20** 2909–34
- [23] Proakis J G 2000 *Digital Communications* 4th edn (New York: McGraw-Hill)
- [24] Viterbi A 1990 Very low rate convolutional codes for maximum theoretical performance of spread-spectrum multiple-access channels *IEEE J. Select. Areas Commun.* **8** 641–9
- [25] Edwards S F and Anderson P W 1975 Theory of spin glasses *J. Phys. F: Met. Phys.* **5** 965–74
- [26] Bray A J and Moore M A 1980 Metastable states in spin glasses *J. Phys. C: Solid State Phys.* **13** L469–76
- [27] Bray A J and Moore M A 1981 Metastable states in spin glasses with short-ranged interactions *J. Phys. C: Solid State Phys.* **14** 1313–27
- [28] Müller R R 2003 Applications of large random matrices in communications engineering *Proc. Int. Conf. on Advances in the Internet, Processing, Systems, and Interdisciplinary Results (IPSI) (Sveti Stefan, Montenegro, October 2003)*