Joint Symbol Detection and Multi-Channel Acquisition in Fast-Fading Narrowband Wireless Environments

João Xavier, Victor Barroso, Tiago Patrão
Instituto de Sistemas e Robótica – Instituto Superior Técnico, Portugal

Abstract—**We address the problem of joint source symbol detection and multi-channel estimation in time-selective digital communication scenarios. Our approach is based on a statistical model which decouples the time dynamics of the multi-channel vector in amplitude and direction. We compute the most probable emitted symbol sequence and channel realization for this statistical model, given the set of array observations. Our maximum a posterior (MAP) receiver consists of a bank of parallel processors. Each processor finds the most probable channel realization for a given symbol sequence via a relaxed semidefinite programming (SDP) re-formulation of the original estimation problem. Computer simulations are included to assess the capability of our technique in acquiring fast-changing flat-fading channels.**

I. Problem Formulation

CONSIDER a wireless communication scenario in which a multiple antenna receiver observes a mobile digital source, as depicted in figure 1. The source transmits the bandpass com-

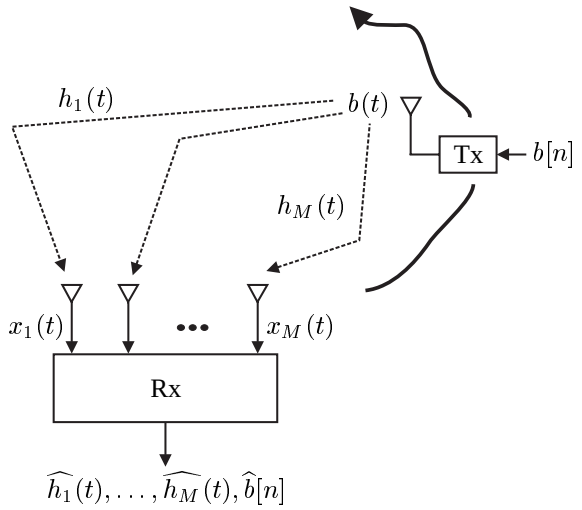

Fig. 1. Flat-fading multi-link channel with a mobile source (baseband model)

plex signal $s(t) = b(t)e^{j2\pi f_c t}$, where f_c denotes the frequency of the complex sinusoidal carrier. The baseband information-bearing signal is given by $b(t) = \sum_{n=-\infty}^{+\infty} b[n]\, p\,(t - nT)$ where $b[n]$ denotes the nth emitted information symbol, T is the symbol period and $p(t)$ is a unit amplitude rectangular shaping pulse of duration T seconds. We assume a flat-fading channel between the source and each one of the M receiving antennas, *i.e.*, the maximum delay spread of the multipath channel linking the source to each spatial sensor is a small fraction of the symbol period T (narrowband source assumption). Assuming that the mth receiver is time-synchronized to the transmitted signal, the complex bandpass signal picked by the mth antenna is $r_m(t) = x_m(t)e^{j2\pi f_c t}$ where $x_m(t) = h_m(t)\, b(t) + w_m(t)$ stands for its baseband equivalent. Here,

$$h_m(t) = c_m(t)e^{j\theta_m(t)} \tag{1}$$

denotes the net baseband gain, $c_m(t)$ is the complex-valued fading channel, $\theta_m(t)$ models carrier phase drifts between the source and the mth spatial sensor, and $w_m(t)$ stands for zero-mean complex additive white Gaussian noise (AWGN) with power spectral density (PSD) $2N_0$ Watts/Hz, that is, $\mathrm{E}\,\{w_m(t)w_m(t-\tau)^*\} = 2N_0\delta(\tau)$. Let the lowpass signal $x_m(t)$, which is available at the receiver after $r_m(t)$ is demodulated to baseband, be oversampled by an integrate-and-dump (I&D) circuit yielding the discrete-time sequence

$$x_m[nP+p] = \frac{1}{\Delta}\int_{nT+p\Delta}^{nT+(p+1)\Delta} x_m(t)\, dt, \quad p = 0, 1, \ldots, P-1,$$

where the integer $P = T/\Delta \geq 1$ denotes the number of data samples taken in each symbol period. Assuming that $h_m(t)$ is constant during the interval of integration, we have

$$x_m[nP+p] = h_m[nP+p]\, b[n] + w_m[nP+p], \tag{2}$$

where $h_m[nP+p] = h_m\,(nT+p\Delta)$ and

$$w_m[nP+p] = \frac{1}{\Delta}\int_{nT+p\Delta}^{nT+(p+1)\Delta} w_m(t)\, dt.$$

Let N be the number of successive symbol intervals thus recorded by the receiver, starting at $t = T$ for convenience of notation. Letting $k = nP + p$ denote the time-index in (2), stacking the data sequences $x_m[k]$ into the complex vector $\boldsymbol{x}[k] = (x_1[k], x_2[k], \ldots, x_M[k])^T$, and collecting the vectors $\boldsymbol{x}[k], k = 1, 2, \ldots, K = NP$ into the data matrix

$$\boldsymbol{X} = \left[\begin{array}{cccc} \boldsymbol{x}[1] & \boldsymbol{x}[2] & \cdots & \boldsymbol{x}[K] \end{array}\right], \tag{3}$$

yields the matricial data model

$$\boldsymbol{X} = \boldsymbol{H}\,\mathrm{diag}\,(\boldsymbol{b} \otimes \boldsymbol{1}_P) + \boldsymbol{W}. \tag{4}$$

Here,

$$\boldsymbol{H} = \left[\begin{array}{cccc} \boldsymbol{h}[1] & \boldsymbol{h}[2] & \cdots & \boldsymbol{h}[K] \end{array}\right] \tag{5}$$

denotes the sequence of vector channel realizations, $\boldsymbol{h}[k] = (h_1[k], h_2[k], \ldots, h_M[k])^T$, $\boldsymbol{b} = (b[1], b[2], \ldots, b[N])^T$ denotes the transmitted sequence of N information symbols, and $\boldsymbol{W} = [\,\boldsymbol{w}[1]\,\boldsymbol{w}[2]\,\cdots\,\boldsymbol{w}[K]\,]$ represents the noise matrix, $\boldsymbol{w}[k] =$

All authors are with the Instituto de Sistemas e Robótica, Instituto Superior Técnico, Torre Norte, Piso 7, Av. Rovisco Pais, 1049–001, Fax: +351 21 841 8291. E-mails: {jxavier,vab}@isr.ist.utl.pt, tiagofgp@yahoo.com.br. This work was supported by the FCT Programa Operacional Sociedade de Informação (POSI) in the frame of QCA III, under contract POSI/2001/CPS/38775

$(w_1[k], w_2[k], \ldots, w_M[k])^T$. We assume that the additive noise processes are spatially white,

$$\mathrm{E}\left\{\boldsymbol{w}[k]\boldsymbol{w}[k-l]^H\right\} = \frac{\sigma^2}{2}\boldsymbol{I}_M\delta[l], \tag{6}$$

where $\sigma^2 = 4N_0/\Delta$, $(\cdot)^H$ denotes the Hermitean operator (transpose conjugate) and $\delta[l]$ stands for the discrete-time Kronecker signal ($\delta[0] = 1$ and $\delta[l] = 0$ for $l \neq 0$). For a generic vector $\boldsymbol{v} = (v_1, v_2, \ldots, v_n)^T$, $\mathrm{diag}(\boldsymbol{v})$ denotes the diagonal matrix with $v_1, v_2, \ldots, v_n$ as its main diagonal entries. The symbol $\otimes$ stands for the Kronecker product and, for an integer n, $\boldsymbol{I}_n$ is the $n \times n$ identity matrix and $\boldsymbol{1}_n = (1, 1, \ldots, 1)^T$ represents the n-dimensional column vector with all entries equal to 1.

In this paper, we address the problem of jointly detecting the emitted information sequence $\boldsymbol{b}$ and estimating the channel matrix $\boldsymbol{H}$ from the available data matrix $\boldsymbol{X}$, see (4). We work under a Bayesian framework. Moreover, we assume that only small data bursts are available for processing, say, with length $N \simeq 4$ symbols. This precludes the usage of 2nd order statistics methods [1], [2]. We assign probabilistic priors to both random objects $\boldsymbol{H}$ (channel) and $\boldsymbol{b}$ (source), and present a sub-optimum implementation of their corresponding maximum-a-posteriori (MAP) estimators. Our paper is organized as follows. In section II, we describe and motivate the priors on the channel and source. We decouple the time dynamics of the channel vector $\boldsymbol{h}[k]$ in amplitude $\rho[k] = \|\boldsymbol{h}[k]\|$ and direction $\boldsymbol{u}[k] = \boldsymbol{h}[k]/\|\boldsymbol{h}[k]\|$. Separate priors are then assigned to the stochastic sequences $\{\rho[k]\}$ and $\{\boldsymbol{u}[k]\}$. For simplicity, we restrict ourselves in this paper to a constant fading envelope, *i.e.*, we assume that $\rho[k] = \rho$ for $k = 1, 2, \ldots, K$, where ρ denotes a random variable uniformly distributed in an interval $[0, A]$. The stochastic sequence $\{\boldsymbol{u}[k]\}$ is modeled as a first-order Markov process on the unit-sphere. The conditional transition probability of $\boldsymbol{u}[k] \,|\, \boldsymbol{u}[k-1]$ is a von Mises-Fisher distribution with mean (or mode) $\boldsymbol{u}[k-1]$ and concentration parameter κ. This 1-parameter (κ) model permits to capture the characteristics of several fast flat-fading digital communication channels. In section III, we discuss a sub-optimum implementation of the MAP estimators of $\boldsymbol{H}$ and $\boldsymbol{b}$ for the given probabilistic prior. We show how the optimization problem underlying the MAP estimation of $\boldsymbol{H}$, for a fixed data sequence $\boldsymbol{b}$, can be approximated by a semidefinite program (SDP). This class of convex programs, which extend linear and quadratic programming, has attracted much attention from the optimization community in the past recent years, leading to the development of powerful primal-dual interior-point solvers. These algorithms can find the global minimum of SDPs with polynomial worst-case complexity and exhibit very good performance in practice. In section IV, we present some computer simulations to assess the performance of our MAP estimator in acquiring fast flat-fading multichannels. Section V concludes our paper.

II. Channel and Source Probabilistic Priors

For certain idealized radio propagation environments, it is possible to deduce analytically some simplified statistical models for the fading channel $c_m(t)$ in (1). As an example, if the mobile is surrounded by many scatterers, several propagation paths (each with its own amplitude and phase) do exist between the narrowband source and the mth antenna. For a large number of these discrete propagation paths, it is plausible to invoke the central limit theorem and model the complex gain $c_m(t)$ as a wide-sense stationary (WSS) complex circular Gaussian process [3], [4], [5]. Moreover, $c_m(t)$ can be taken as zero-mean (Rayleigh fading model) if there is no direct line-of-sight (LOS) component, whereas a nonzero mean must be included if a direct specular component is present (Rice fading model). Besides the Rayleigh/Rice distribution, other first-order statistics modeling the envelope of the complex-valued fading channel, $|c_m(t)|$, have been considered in the literature, e.g., the Nakagami-q (Hoyt), the Nakagami-n, and the Nakagami-m models, see [4] and the references therein. The autocorrelation function of the WSS Gaussian process $c_m(t)$, $r_m(\tau) = \mathrm{E}\left\{c_m(t)c_m(t-\tau)^*\right\}$, dictating the second-order statistics of the fading channel, can be obtained for some specific scattering propagation configurations from Doppler-shift motion-induced physical considerations. For example, the assumption of an idealized isotropic scattering scenario (the mobile is surrounded by a cluster of scatterers uniformly distributed in angle), an uniform azimuthal power gain for the mth receiving antenna (omnidirectional sensor), and a constant vehicle speed, leads to the Clarke's model [5] with the fading autocorrelation given by $r_m(\tau) = \sigma_m^2 J_0(2\pi f_m \tau)$. Here, $\sigma_m^2 = \mathrm{E}\left\{|c_m(t)|^2\right\}$ denotes the power of the fading process, $J_0(\cdot)$ is the zero-order Bessel function of the first kind, $f_m = v_m f_c/c$ is the maximum Doppler frequency in Hz, v_m is the speed of the mobile source relative to the mth antenna in m/s, f_c stands for the central frequency of the transmitted signal in Hz, and $c = 3 \times 10^8$ m/s denotes the speed of light. Clarke's model is employed in land mobile scenarios [4]. Other fading autocorrelation models are available for distinct propagation scenarios, see [4]. The phase drift $\theta_m(t)$ in (1) models non-channel induced phase shifts between the transmitter and the mth receiver, e.g., transmitter and local oscillator asynchronism. As an illustrative example, consider an $M = 2$ antenna array receiver which observes a digital source with symbol period $T = 0.1$ ms and carrier frequency $f_c = 1$ GHz. We assume that the vehicle moves with a speed of $v = 120$ Km/h. As the fading channel model, we take $c_1(t)$ and $c_2(t)$ as zero-mean unit-power statistically independent complex Gaussian processes, each with autocorrelation function predicted by Clarke's model. Moreover, we assume that the crystal oscillator at the receiver has a stability of $\eta = 0.5$ ppm (part per million), *i.e.*, we let $\theta_m(t) = 2\pi f_\theta t$, where the oscillator frequency error $f_\theta = \eta f_c = 500$ Hz, for $m = 1, 2$. Let $\boldsymbol{h}(t) = \rho(t)\,\boldsymbol{u}(t)$ denote the decomposition of the channel vector $\boldsymbol{h}(t)$ in amplitude $\rho(t) = \|\boldsymbol{h}(t)\| \geq 0$ and direction $\boldsymbol{u}(t) = \boldsymbol{h}(t)/\|\boldsymbol{h}(t)\| = (u_1(t), u_2(t))^T$. In figures 2 and 3, we see a realization of $\boldsymbol{h}(t)$, through its components $\rho(t)$ and $\boldsymbol{u}(t)$, over a time span of $N = 4$ symbol periods. We have taken 40 equi-spaced time samples of $\rho(t)$ and $\boldsymbol{u}(t)$ in the observation period $[0, NT]$, and computed some statistics. For this realization, the mean value of the channel amplitude is $\overline{\rho} = 1.9234$ with a standard deviation of $\sigma_\rho = 0.0907$. Thus, the envelope exhibits a fluctuation of $\sigma_\rho/\overline{\rho} = 4.7$ % about its nominal value $\overline{\rho}$. The mean values of $\mathrm{Re}\,u_1(t), \mathrm{Re}\,u_2(t), \mathrm{Im}\,u_1(t), \mathrm{Im}\,u_2(t)$ are $-0.5699, 0.4923, -0.4793$ and 0.2794, respectively. The corresponding standard deviations are $0.1659, 0.1334, 0.2210$ and 0.1845, leading to fluctuations of 29.11 %, 27.10 %, 46.11 %

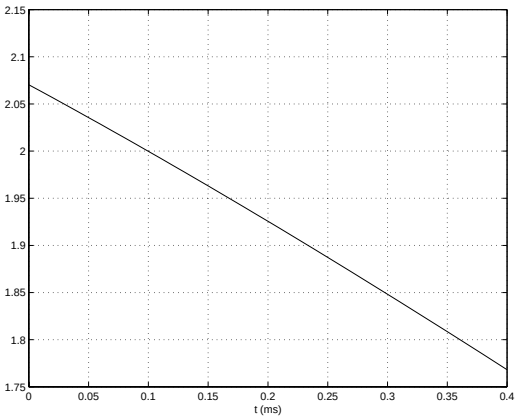

Fig. 2. Realization of $\rho(t)$ (example)

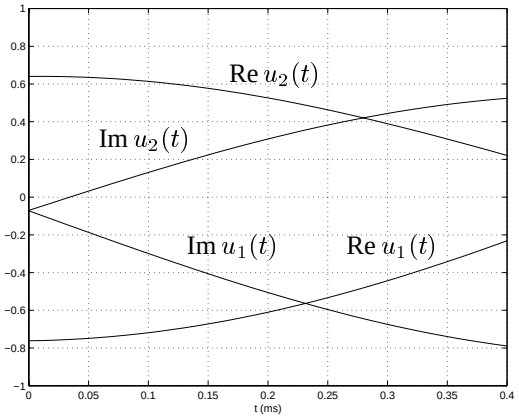

Fig. 3. Realization of $\boldsymbol{u}(t)$ (example)

and 66.02 %, respectively. From these data, we can conclude that the time variation of the channel vector $\boldsymbol{h}(t)$ over the restricted interval $[0, NT]$ is mainly due to the time variation of the phase in each mth entry of $\boldsymbol{h}(t)$, *i.e.*, the net effect of the time variation of the phase of the fading channel $c_m(t)$ and the phase drift $\theta_m(t)$, which $\boldsymbol{u}(t)$ preserves up to a multiplicative factor. The channel amplitude $\rho(t)$ is nearly constant over the time interval considered. This asymmetric behavior of $\rho(t)$ and $\boldsymbol{u}(t)$ becomes more noticeable if more statistically independent antennas are employed at the receiver (spatial diversity), or if a Rice channel model is considered, as both of these scenarios tend to stabilize the amplitude of the channel vector. For example, as it is well-known, the Ricean fading channel approaches the classical non-fading (constant amplitude) AWGN channel as the Rician factor tends to infinity. The fact that the amplitude $\rho(t) = \|\boldsymbol{h}(t)\|$ of the source spatial signature varies slower, for small time intervals, than the "phase" vector $\boldsymbol{u}(t) = \boldsymbol{h}(t)/\|\boldsymbol{h}(t)\|$ is in agreement with experimental measurements (e.g., see [6]), and generalizes the typical behavior of single-channel systems: notice that, for $M = 1$ channel given by $h(t) = A(t)e^{j\phi(t)}$, we have $\rho(t) = A(t) \geq 0$, and the vector $\boldsymbol{u}(t)$ specializes to the pure (unit-amplitude) phasor $\boldsymbol{u}(t) = e^{j\phi(t)}$, where $\phi(t)$ accounts for the joint time variation of the phase of the fading channel and carrier phase drift. Motivated by this behavior of the time dynamics of the channel vector $\boldsymbol{h}(t)$ over short time intervals, characteristic of many fading scenarios, we work in the sequel with the following statistical model for the length K vector channel sequence defined in (5): we let $\boldsymbol{h}[k] = \rho\,\boldsymbol{u}[k]$, where

$$\rho \sim \mathcal{U}([0, A]) \tag{7}$$

denotes a random variable uniformly distributed over the interval $[0, A]$, where $A > 0$ denotes some fixed constant; the sequence $\{\boldsymbol{u}[1], \boldsymbol{u}[2], \ldots, \boldsymbol{u}[K]\}$ is statistically independent of ρ and is taken as a first-order Markov process on the unit-sphere. More specifically, let $\boldsymbol{v}[k] = \left(\mathrm{Re}\,\boldsymbol{u}[k]^T\ \mathrm{Im}\,\boldsymbol{u}[k]^T\right)^T$. Notice that each $2M$-dimensional vector $\boldsymbol{v}[k]$ lies in $\mathbb{S}^{2M-1}$. Here, and for further reference, $\mathbb{S}^{n-1} = \{\boldsymbol{v} \in \mathbb{R}^n : \|\boldsymbol{v}\| = 1\}$ denotes the $(n-1)$-dimensional unit-sphere of the Euclidean space $\mathbb{R}^n$. We let $\boldsymbol{v}[1]$ be uniformly distributed over the sphere $\mathbb{S}^{2M-1}$, written

$$\boldsymbol{v}[1] \sim \mathcal{U}\left(\mathbb{S}^{2M-1}\right), \tag{8}$$

and let the one-step transition probability be given by

$$\boldsymbol{v}[k]\,|\,\boldsymbol{v}[k-1] \sim \mathcal{M}_{2M}\left(\boldsymbol{v}[k-1], \kappa\right). \tag{9}$$

Here, $\mathcal{M}_p(\boldsymbol{\mu}, \kappa)$ denotes the von Mises-Fisher distribution on $\mathbb{S}^{p-1}$ with the unit-norm vector $\boldsymbol{\mu}$ as mean direction and the non-negative scalar κ as the concentration parameter, see [7]. The density of the von Mises-Fisher distribution $\mathcal{M}_p(\boldsymbol{\mu}, \kappa)$ with respect to the uniform distribution on the unit-sphere is

$$f(\boldsymbol{v}) = \alpha_p(\kappa)\,\exp\left(\kappa\,\boldsymbol{\mu}^T\boldsymbol{v}\right), \tag{10}$$

where $\boldsymbol{v} \in \mathbb{S}^{p-1}$ and $\alpha_p(\kappa)$ denotes the normalizing constant. This distribution reduces to the uniform distribution on the unit-sphere for $\kappa = 0$, and exhibits a mode at $\boldsymbol{\mu}$ for $\kappa > 0$. As κ increases, the probability mass becomes more concentrated around the mean direction $\boldsymbol{\mu}$. As an example, we plot in figures 4 and 5, the density of the von Mises-Fisher distribution $\mathcal{M}_p(\boldsymbol{\mu}, \kappa)$ on the circle ($p = 2$) with mean direction $\boldsymbol{\mu} = (\cos(\pi/4), \sin(\pi/4))^T$ and concentration parameters $\kappa = 0.5$ and $\kappa = 5$, respectively. Some remarks are in order

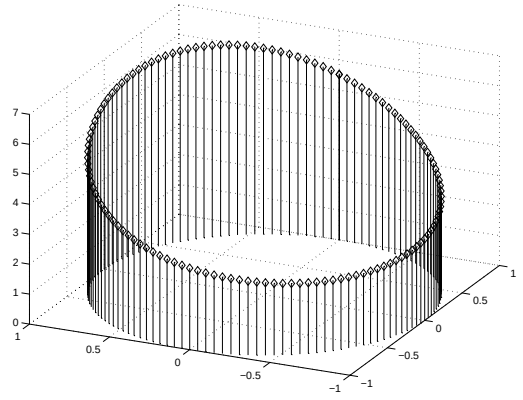

Fig. 4. Density of $\mathcal{M}_2(\boldsymbol{\mu}, \kappa) : \boldsymbol{\mu} = (\cos(\pi/4), \sin(\pi/4))^T$ and $\kappa = 0.5$

regarding our channel statistical model. **i)** It is a 1-parameter model (with parameter κ) that does not rely on any special assumption about the scattering environment, antenna directivity pattern, etc. Its main purpose is to be able of reproducing the typical time variation of the channel vector $\boldsymbol{h}(t)$ over small observation intervals, which occurs in many flat-fading propagation scenarios. Of course, if an alternative, more sophisticated

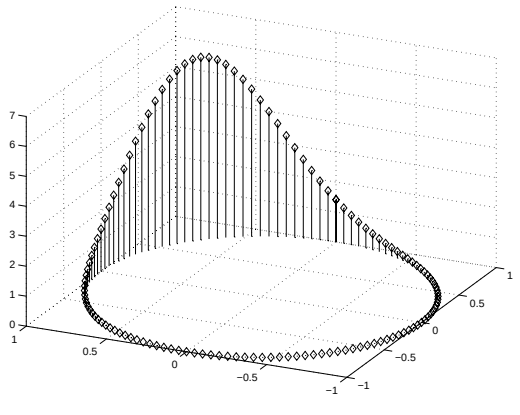

Fig. 5. Density of $\mathcal{M}_2(\boldsymbol{\mu}, \kappa) : \boldsymbol{\mu} = (\cos(\pi/4), \sin(\pi/4))^T$ and $\kappa = 5$

channel model (perhaps based on field measurements) may be trusted, it should be employed. In sum, we are trading simplicity (only one parameter to tune) and robustness (no special propagation scenario is assumed) for accuracy. **ii)** The assumption that the vector channel amplitude is constant, $\|\boldsymbol{h}[k]\| = \rho$, is taken here because we work only with small data bursts. However, it should be said that the constant amplitude assumption is mainly taken for simplicity: for example, it can be seen that the inclusion of independent fading amplitudes per each data sample is easily accommodated in our SDP framework to be presented in the next section (see also [8]). The distribution $\rho \sim \mathcal{U}([0, A])$ acts as a non-informative prior, reflecting our ignorance about the initial channel state. **iii)** Under certain special circumstances, some models for the phase variation of the mth multipath channel $h_m(t)$ can be devised. For example, in the case of perfect transmitter and receiver oscillators and a propagation scenario with a dominant direct LOS component with weaker (negligible) indirect components, the phase variation can be approximated by a linear dynamic, the well-known Doppler effect due to vehicle motion. However, since we do not rely on any specific propagation scenario nor oscillator asynchronism model, this viewpoint is not taken here. Instead, we let the phase vary randomly from data sample to data sample, with the concentration parameter κ controlling the amount of "randomness" per transition: larger values of κ correspond to slower time-varying phases, whereas smaller values of κ can model fast-changing phase processes. Again, the uniform density assumption $\boldsymbol{v}[1] \sim \mathcal{U}\left(\mathbb{S}^{2M-1}\right)$ about the initial direction vector is adopted to reflect our ignorance about the initial vector channel state. As a final remark, one could complicate the model and allow for distinct concentration parameters per vector channel entry, or for a time-varying concentration parameter, or both. These and other model refinements are explored elsewhere [8].

Regarding the transmitter model, we consider that the information source emits a string $\{b[n]\}$ of independent and identically distributed symbols drawn from a finite modulation alphabet. For simplicity, we assume hereafter a binary-shift keying (BSK) digital source, *i.e.*, the symbols $b[n]$ are taken from the set $\{-1, 1\}$ and are equiprobable:

$$\Pr\{b[n] = -1\} = \Pr\{b[n] = 1\} = \frac{1}{2}. \tag{11}$$

III. MAP Channel and Symbol Estimators

The MAP estimates of the unknown channel random parameters ρ, $\boldsymbol{V} = [\,\boldsymbol{v}[1]\,\boldsymbol{v}[2]\,\cdots\,\boldsymbol{v}[K]\,]$ and the source sequence of bits $\boldsymbol{b} = (b[1], b[2], \ldots, b[N])^T$ correspond to the global solutions of the optimization problem

$$\left(\widehat{\rho}, \widehat{\boldsymbol{V}}, \widehat{\boldsymbol{b}}\right) = \underset{\rho,\, \boldsymbol{V},\, \boldsymbol{b}}{\arg\max}\; p(\rho, \boldsymbol{V}, \boldsymbol{b} \mid \boldsymbol{X}). \tag{12}$$

Recall from (3) that $\boldsymbol{X}$ denotes the matrix of observations. In (12), we have the explicit constraints $0 \leq \rho \leq A$ (recall the prior on the amplitude of the vector channel in (7)), $\|\boldsymbol{v}[k]\| = 1$ for $k = 1, 2, \ldots, K$ and $b[n] \in \{\pm 1\}$ for $n = 1, 2, \ldots, N$. We are assuming that both the noise variance σ^2 and the concentration parameter κ are known. In fact, only their product $\sigma^2\kappa$ needs to be assumed known (see (14) below). Using the Bayes rule and our statistical assumptions – see (6), (7), (8), (9), (10) and (11) – we have, after some trivial algebraic manipulations, the equivalent optimization problem

$$\left(\widehat{\rho}, \widehat{\boldsymbol{V}}, \widehat{\boldsymbol{b}}\right) = \underset{\rho,\, \boldsymbol{V},\, \boldsymbol{b}}{\arg\min}\; \psi(\rho, \boldsymbol{V}, \boldsymbol{b}) \tag{13}$$

where

$$\psi(\rho, \boldsymbol{V}, \boldsymbol{b}) = \rho^2 - \frac{2}{K}\sum_{k=1}^{K} \rho\, \boldsymbol{y}_{\boldsymbol{b}}[k]^T \boldsymbol{v}[k] - \frac{\sigma^2\kappa}{2K}\sum_{k=2}^{K} \boldsymbol{v}[k]^T \boldsymbol{v}[k-1], \tag{14}$$

and the sequence $\boldsymbol{y}_{\boldsymbol{b}}[k]$ is defined by

$$\boldsymbol{y}_{\boldsymbol{b}}[nP + p] = \begin{bmatrix} b[n]\,\mathrm{Re}\,\boldsymbol{x}[nP + p] \\ b[n]\,\mathrm{Im}\,\boldsymbol{x}[nP + p] \end{bmatrix} \tag{15}$$

for $p = 0, 1, \ldots, P-1$ and $n = 1, 2, \ldots, N$. Notice that the sequence $\boldsymbol{y}_{\boldsymbol{b}}[k]$ depends on the sequence of bits $\boldsymbol{b} = (b[1], b[2], \ldots, b[N])^T$ as equation (15) shows (hence, the subscript $\boldsymbol{b}$ in the notation $\boldsymbol{y}_{\boldsymbol{b}}[k]$).

The optimization problem in (13) is posed in terms of discrete ($\boldsymbol{b}$) and continuous ($\rho, \boldsymbol{V}$) variables. It may be solved by enumerating all bit sequences of length N and, for each one, say $\boldsymbol{b}$, optimize over ρ and $\boldsymbol{V}$ to yield the corresponding estimates $\rho_{\boldsymbol{b}}$ and $\boldsymbol{V}_{\boldsymbol{b}}$,

$$(\rho_{\boldsymbol{b}}, \boldsymbol{V}_{\boldsymbol{b}}) = \underset{\rho,\, \boldsymbol{V}}{\arg\min}\; \psi(\rho, \boldsymbol{V}, \boldsymbol{b}). \tag{16}$$

In fact, since there is an unavoidable sign ambiguity in the variables $\boldsymbol{V}$ and $\boldsymbol{b}$, because $\psi(\rho, \boldsymbol{V}, \boldsymbol{b}) = \psi(\rho, -\boldsymbol{V}, -\boldsymbol{b})$, we may fix a bit, e.g., $b[1] = -1$, and enumerate over all 2^{N-1} bit sequences $\{b[2], \ldots, b[N]\}$ solving, for each one, the optimization problem in (16). The MAP estimates of the channel and source realizations in (12) are then given by $\widehat{\rho} = \rho_{\boldsymbol{b}}$, $\widehat{\boldsymbol{V}} = \boldsymbol{V}_{\boldsymbol{b}}$ and $\widehat{\boldsymbol{b}} = \boldsymbol{b}$, where $\boldsymbol{b}$ denotes the sequence minimizing $\psi(\rho_{\boldsymbol{b}}, \boldsymbol{V}_{\boldsymbol{b}}, \boldsymbol{b})$ over all bit sequences considered. This approach may be implemented through a bank of 2^{N-1} parallel processors, which is feasible for small sequence lengths N (as we are assuming throughout the paper). Each processor solves problem (16) for a fixed sequence of bits $\boldsymbol{b}$. Hereafter, we treat the vector $\boldsymbol{b}$ (hence, $\boldsymbol{y}_{\boldsymbol{b}}[k]$) as a constant, and focus on the prototype optimization problem in (16) with constraints $0 \leq \rho \leq A$ and $\|\boldsymbol{v}[k]\| = 1$ for $k = 1, 2, \ldots, K$. This is a highly nonlinear problem with

no apparent closed-form solution. It can be tackled by general-purpose iterative algorithms for constrained problems described in the standard references of optimization theory, e.g., see [9]. These iterative procedures are usually only locally-convergent. Thus, several time consuming re-initializations might be required to find the global solution. Here, we pursue a distinct approach. We exploit the special structure of the constraints to relax problem (16) into a nearby semidefinite program. This particular class of convex optimization problems has been under intense investigation over the past few years. It contains linear (and quadratic) programs as special cases, and admits globally convergent efficient algorithms based on iterating interior points which either follow central paths or reduce a potential function. Moreover, SDP finds applications in combinatorial optimization (providing polynomial-time bounds for NP-hard problems) [10], [11], systems and control theory [12], eigenvalue optimization problems [13], etc. In the sequel, we assume that the reader is familiar with optimization theory and, in particular, semidefinite programming theory. Survey papers can be found in [14], [10]. Several SDP resources (reviews, bibliography, software packages) are available from Christoph Helmberg's homepage `http://www.zib.de/helmberg/semidef.html`. We obtain the SDP relaxation of problem (16) by relaxing a rank 1 constraint that appears in one of its equivalent formulations. This relaxation technique is usually employed in the context of $(0,1)$-integer optimization [11]. In our case, we start by rewriting the primal problem (16) with $\boldsymbol{b}$ fixed and variables ρ and $\boldsymbol{v}[1],\ldots,\boldsymbol{v}[K]$ as

$$\begin{array}{c}\min \\ \rho^2 \leq A^2 \\ \{\boldsymbol{v}[k]^T\boldsymbol{v}[k]=1\}\end{array} \quad \left[\begin{array}{cc} \boldsymbol{v}^T & \rho \end{array}\right] \boldsymbol{\Gamma} \left[\begin{array}{c} \boldsymbol{v} \\ \rho \end{array}\right], \tag{17}$$

where $\boldsymbol{v}=\mathrm{vec}(\boldsymbol{V})$. Here, and for further reference, vec denotes the vectorization operator: for an arbitrary $n\times m$ rectangular matrix $\boldsymbol{A}=[\,\boldsymbol{a}_1\,\boldsymbol{a}_2\,\cdots\,\boldsymbol{a}_m\,]$, we have the nm-dimensional column vector $\mathrm{vec}(\boldsymbol{A})=\left(\boldsymbol{a}_1^T,\boldsymbol{a}_2^T,\ldots,\boldsymbol{a}_m^T\right)^T$. In (17),

$$\boldsymbol{\Gamma}=\left[\begin{array}{cc} \boldsymbol{T} & \boldsymbol{y} \\ \boldsymbol{y}^T & 1 \end{array}\right]$$

where $\boldsymbol{y} = \mathrm{vec}\left(-\frac{1}{K}\,[\,\boldsymbol{y_b}[1]\,\boldsymbol{y_b}[2]\,\cdots\,\boldsymbol{y_b}[K]\,]\right)$, $\boldsymbol{T} = -(\sigma^2\kappa)/(4K)\,\boldsymbol{R}\otimes\boldsymbol{I}_{2M}$, and $\boldsymbol{R}$ is a $K\times K$ matrix with 1's in the first upper and lower main diagonals and 0's elsewhere. We can reformulate (17) as

$$\begin{array}{c}\min \\ \boldsymbol{Z}=\left[\begin{array}{c} \boldsymbol{v} \\ \rho \end{array}\right]\left[\begin{array}{cc} \boldsymbol{v}^T & \rho \end{array}\right] \\ \boldsymbol{Z}_{2KM+1,2KM+1}\leq A^2 \\ \{\mathrm{tr}\,(\boldsymbol{Z}_k)=1\}\end{array} \quad \mathrm{tr}\,(\boldsymbol{\Gamma}\boldsymbol{Z})\,, \tag{18}$$

where tr denotes the trace operator, $\boldsymbol{Z}_{i,j}$ is the (i,j)th entry of the matrix $\boldsymbol{Z}$, and $\boldsymbol{Z}_k$ (for $k=1,2,\ldots,K$) stands for the $2M\times 2M$ submatrix of $\boldsymbol{Z}$ obtained by retaining the rows and columns ranging from $\alpha_k=(k-1)2M+1$ to $\beta_k=2kM$. The SDP relaxation of (18) consists in relaxing the rank 1 constraint

$$\boldsymbol{Z}=\left[\begin{array}{c} \boldsymbol{v} \\ \rho \end{array}\right]\left[\begin{array}{cc} \boldsymbol{v}^T & \rho \end{array}\right]=\left[\begin{array}{cc} \boldsymbol{v}\boldsymbol{v}^T & \rho\,\boldsymbol{v} \\ \rho\,\boldsymbol{v}^T & \rho^2 \end{array}\right] \tag{19}$$

into the positive semidefinite constraint $\boldsymbol{Z}\succeq\boldsymbol{0}$, yielding the semidefinite program

$$\begin{array}{c}\min \\ \boldsymbol{Z}\succeq\boldsymbol{0} \\ \boldsymbol{Z}_{2KM+1,2KM+1}\leq A^2 \\ \{\mathrm{tr}\,(\boldsymbol{Z}_k)=1\}\end{array} \quad \mathrm{tr}\,(\boldsymbol{\Gamma}\boldsymbol{Z})\,. \tag{20}$$

Let $\boldsymbol{Z}$ denote the solution of (20). We obtain an approximation for $\rho_{\boldsymbol{b}}$ and $\boldsymbol{V_b}=[\,\boldsymbol{v_b}[1]\,\cdots\,\boldsymbol{v_b}[K]\,]$ in (16) by letting $\rho_{\boldsymbol{b}}=\sqrt{\boldsymbol{Z}_{KM+1,KM+1}}$ and $\boldsymbol{v_b}[k]=\boldsymbol{z}_k/\|\boldsymbol{z}_k\|$, $k=1,\ldots,K$, where $\boldsymbol{z}_k$ denotes the $2M$-dimensional subvector of the last column of $\boldsymbol{Z}$ ranging from rows α_k to β_k. These choices for $\rho_{\boldsymbol{b}}$ and $\boldsymbol{V_b}$ are exact if $\boldsymbol{Z}$ has rank 1, as seen from the last column of the identity in (19). The number of variables in each SDP (20), hence its computational complexity, can be reduced by approximating the time variation of the channel direction vector, *i.e.*, $\boldsymbol{V}=[\,\boldsymbol{v}[1]\,\boldsymbol{v}[2]\,\cdots\,\boldsymbol{v}[K]\,]$ with a piecewise constant vector sequence. For example, assuming K is even and taking constant segments of length $L=2$ would lead to the parameterization $\boldsymbol{V}=[\,\boldsymbol{\omega}[1]\,\boldsymbol{\omega}[1]\,\boldsymbol{\omega}[2]\,\boldsymbol{\omega}[2]\,\cdots\,\boldsymbol{\omega}[K/2]\,\boldsymbol{\omega}[K/2]\,]$, in which we only let the channel direction $\boldsymbol{v}[k]$ change every $L=2$ data samples. With this approach, the number of variables is halved. The general case of constant segments of length L (K a multiple of L) leads to the reduced-size SDP

$$\begin{array}{c}\min \\ \widetilde{\boldsymbol{Z}}\succeq\boldsymbol{0} \\ \widetilde{\boldsymbol{Z}}_{2KM/L+1,2KM/L+1}\leq A^2 \\ \{\mathrm{tr}\left(\widetilde{\boldsymbol{Z}}_k\right)=1\}\end{array} \quad \mathrm{tr}\left(\widetilde{\boldsymbol{\Gamma}}\widetilde{\boldsymbol{Z}}\right), \tag{21}$$

where $\widetilde{\boldsymbol{\Gamma}}=\boldsymbol{\Omega}^T\boldsymbol{\Gamma}\boldsymbol{\Omega}$, $\boldsymbol{\Omega}$ being the diagonal concatenation of $\boldsymbol{I}_{K/L}\otimes\boldsymbol{1}_L\otimes\boldsymbol{I}_{2M}$ with 1. Notice that the variable $\boldsymbol{Z}$ to be optimized in (20) has size $(2KM+1)\times(2KM+1)$, whereas in (21) the variable $\widetilde{\boldsymbol{Z}}$ has size $(2KM/L+1)\times(2KM/L+1)$. Let $\widetilde{\boldsymbol{Z}}$ denote the solution of (21). We obtain $\rho_{\boldsymbol{b}}$ and $\boldsymbol{\omega}[k]$ as before: $\rho_{\boldsymbol{b}}=\sqrt{\widetilde{\boldsymbol{Z}}_{2KM/L+1,2KM/L+1}}$ and $\boldsymbol{\omega}[k]=\widetilde{\boldsymbol{z}}_k/\|\widetilde{\boldsymbol{z}}_k\|$ for $k=1,\ldots,K/L$, where $\widetilde{\boldsymbol{z}}_k$ denotes the $2M$-dimensional subvector of the last column of $\widetilde{\boldsymbol{Z}}$ ranging from rows α_k to β_k. In [8], we study in more detail the performance of this piecewise constant model for the time variation of the channel vector direction. Moreover, we derive a low-cost computational iterative scheme, based on differential-geometric concepts, to refine these sub-optimum estimates.

IV. Computer simulations

We conducted computer simulations to analyse the performance of our proposed MAP estimator. We considered a $M=1$ antenna receiver. We assumed an oversampling factor of $P=3$ and process $N=4$ consecutive bits. Thus, the data packet length is $K=PN=12$. Each data packet is generated according to our channel and source priors. We have fixed the vector channel amplitude throughout the simulations, $\rho=1$ (ignored at the receiver) and considered as von Mises-Fisher concentration parameters $\kappa=5,10,15,20,25$. For each κ, we varied the signal-to-noise ratio (SNR) from $\mathrm{SNR}_{\min}=-5$ dB to $\mathrm{SNR}_{\max}=20$ dB in steps of $\Delta=2.5$ dB. The SNR is defined as $\mathrm{SNR}=\mathrm{E}\left\{\|\boldsymbol{h}[\cdot]b[\cdot]\|^2\right\}/\mathrm{E}\left\{\|\boldsymbol{w}[\cdot]\|^2\right\}=2/\sigma^2$. For each

SNR, 2000 statistically independent Monte-Carlos runs were performed. Each Monte-Carlo run involves detecting the transmitted bits $\boldsymbol{b}$ and estimating the channel realization $\rho, \boldsymbol{V}$ with our MAP receiver. Figure 6 shows the bit error rates (BER) obtained for each κ over the considered range of SNRs. Results beyond SNR $= 7.5$ dB are not shown because they are not statistically significant (more Monte-Carlo runs are required). In

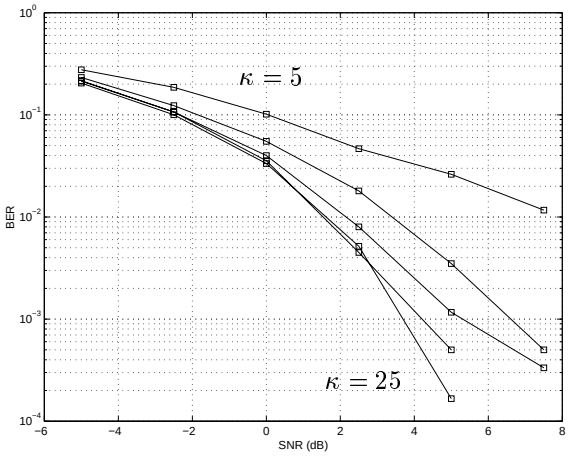

Fig. 6. Bit error rate (BER) versus signal-to-noise ratio (SNR)

figure 7, we plot the mean of the estimate $\widehat{\rho}$ versus SNR. As can be seen, the estimate converges to the true value $\rho = 1$, as the SNR tends to infinity. Suppose the sequence of bits is known,

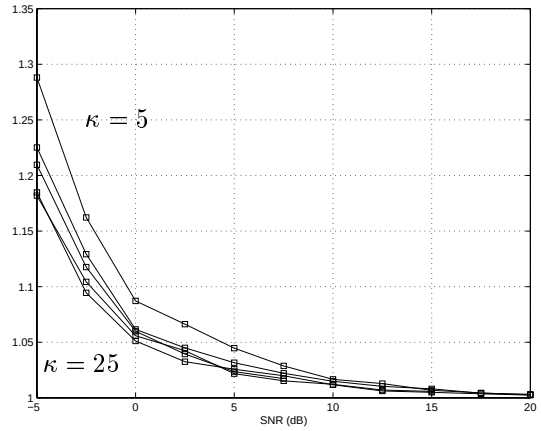

Fig. 7. Estimated fading amplitude ($\widehat{\rho}$) versus SNR

and we only face the task of estimating the channel. That is, we must solve (16) for a given (training) sequence of bits $\boldsymbol{b}$. Since $M = 1$, each channel direction $\boldsymbol{u}[k]$ is a point in the unit-radius circle of the complex plane. In figure 8, we show the average absolute phase error obtained by our estimator. Notice that with our von Mises-Fisher channel model, the phase of the channel has a mean jump of $22.2, 15.8, 13.2, 11.8$ and 10.8 degrees from $\boldsymbol{u}[k]$ to $\boldsymbol{u}[k+1]$ (separated in time by one third of the symbol period), for $\kappa = 5, 10, 15, 20$ and 25, respectively. Thus, we are tackling a scenario with rapidly varying channel phase.

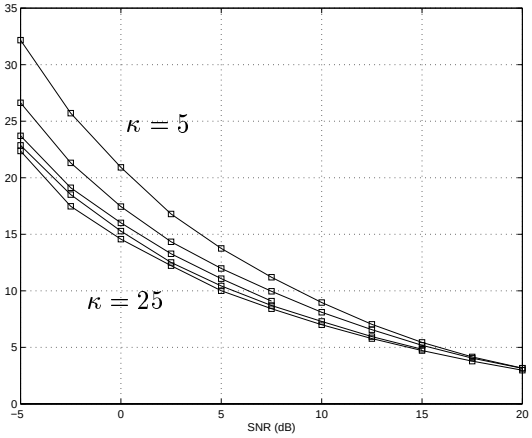

Fig. 8. Mean phase error (degrees) versus SNR

V. Conclusions

We addressed the problem of joint source symbol detection and multi-channel estimation in the context of flat-fading wireless communications. We rely on a simple vector channel model which captures its typical behavior in many idealized flat-fading propagation scenarios. We decouple the time dynamics of the amplitude and direction of the multi-channel vector over short time intervals. We let the amplitude remain constant and model the time variation of the channel direction as a Markov process on the unit-sphere. We implemented (sub-optimally) the MAP estimators of the emitted symbol sequence and channel realization. We exploited the special structure of the MAP optimization problem, and found a nearby SDP reformulation which can be efficiently solved by recently developed interior-point algorithms. Preliminary results assessed the ability of our method in acquiring fast-changing channels.

References

[1] V. Barroso, J. Xavier, and J. M. F. Moura, "Blind Array Channel Division Multiple Access (AChDMA) for Mobile Communications", *IEEE Trans. on Signal Proc.*, vol. 46, pp. 737–752, March 1998.

[2] J. Xavier, V. Barroso, and J. M. F. Moura, "Closed-form Correlative Coding (CFC_2) Blind Identification of MIMO Channels: Isometry Fitting to Second-Order Statistics," *IEEE Trans. on Signal Proc.*, vol. 49, no. 5, pp. 1073 – 1086, May 2001.

[3] J. Cavers, *Mobile Channel Characteristics*, Kluwer Academic Pub., 2000.

[4] M. Simon and M. Alouini, *Digital Communication over Fading Channels* , John Wiley & Sons, 2000.

[5] R. H. Clarke, "A statistical theory of mobile radio reception," *Bell System Technical Journal*, vol. 47, pp. 957–1000, 1968.

[6] S. Jeng, H. Lin, G. Okamoto, G.Xu, and W. Vogel, "Multipath direction finding with subspace smoothing," *in Proceedings of the IEEE Int. Conf. on Acoust., Speech and Sig. Proc. (ICASSP'97)*, pp. 3485–3488, Munich 1997.

[7] K. Mardia and P. Jupp, *Directional Statistics*, John Wiley & Sons, 2000.

[8] J. Xavier, V. Barroso, and T. Patrão, "Joint MAP Symbol Sequence Detection and Channel Estimation in Fast-Fading Narrowband Wireless Systems with Spatial Diversity: a Bayesian Approach," in preparation.

[9] D. Bertsekas, *Nonlinear Programming*, 2nd ed., Athena Scientific, 1999.

[10] F. Alizadeh, "Interior point methods in semidefinite programming with applications to combinatorial optimization," *SIAM Journal on Optimization*, vol. 5, pp. 13–51, 1995.

[11] N. Shor, "Dual quadratic estimates in polynomial and Boolean programming," *Annals of Operations Research*, vol. 25, pp. 163–168, 1990.

[12] S. Boyd, L. Ghaoui, E. Feron, and V. Balakrishnan. *Linear Matrix Inequalities in System and Control Theory*, vol. 15 of *Studies in Applied mathematics.* SIAM, Philadelphia, PA, June 1994.

[13] A. Lewis and M. Overton, "Eigenvalue optimization," *Acta Numerica*, vol. 5, pp. 149–190, 1996.

[14] L. Vandenberghe and S. Boyd, "Semidefinite programming," *SIAM Review*, vol. 38, no. 1, pp. 49–95, March 1996.