

# MIMO Mobile Terminal Tracking Using Bayesian Probability Estimation

Konstantinos Papakonstantinou<sup>†</sup>, Merouane Debbah<sup>‡</sup> and Dirk Slock<sup>†</sup>

<sup>†</sup>Institute Eurecom, Mobile Communications Department

<sup>‡</sup> Supelec, Telecommunications Department

Email: papakons@eurecom.fr, merouane.debbah@supelec.fr, slock@eurecom.fr

**Abstract**— We present a simple and efficient way of tracking a Mobile Terminal (MT) using multiple antennas at the transmitter and the receiver. We consider the double directional model and use the Bayesian framework in order to estimate the speed and direction of the MT. We show in particular how to exploit the space dimension, inherently present in a MIMO system, to increase the estimation accuracy. In contrast to the majority of solutions that have been proposed so far for MT localization and tracking, this new method is suitable for a Non-Line-of-Sight (NLoS) propagation scenario and employs only one Base Station (BS)<sup>1</sup>.

## I. INTRODUCTION

Multiple-Input Multiple-Output (MIMO) systems have attracted much interest from an information-theoretic perspective, as it was proved that they increase the capacity linearly with the number of antennas [1]. In this paper we investigate how MIMO systems can be used in terms of localization. Antenna arrays have been extensively used in the past (at one or both sides of a communication system) for estimating the directions of arrival (DoAs) and/or the directions of departure (DoDs) of a signal (see for example the very appealing algorithms in [2],[3],[4]). This utility of MIMO systems will play a key role in the proposed tracking method.

One of the major problems that arise when trying to estimate location or movement dependent parameters from measured channel impulse responses is the errors due to the presence of NLoS components. To improve the accuracy under these conditions, researchers have adopted two entirely different approaches: Mitigate the errors, by trying to reduce or completely eliminate the impact of these components [5], or take these components into account and use the information contained in them to increase accuracy [6]. We will adopt the second approach, since this is the only applicable when there is no LoS component, a situation which is met often in real propagation environments. However, in contrast to most of the work that has been presented so far, in which a single bounce of the signal to any scatterer is the basic principle behind channel modeling [7],[8], in our tracking method, a more general model is used.

The rest of the paper is organized as follows: In section II we present the channel model. In section III we briefly describe the principles of Bayesian inference and estimation and we derive a formula which enables us to estimate parameters

<sup>1</sup>The work presented herein was supported by the project Semaphore.

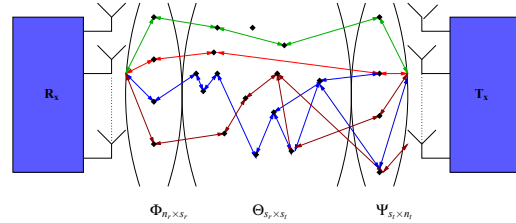


Fig. 1. Double directional model for a MIMO system.

in separate steps, easily. Simulation results for systems with different number of antennas and at different Signal-to-Noise (SNR) ratios are presented in section IV. Finally conclusions and suggestions for future work are given in section V.

**Notation:** Throughout the paper, upper case and lower case boldface symbols will represent matrices and column vectors respectively.  $(\cdot)^t$  will denote the transpose,  $(\cdot)^*$  the conjugate and  $(\cdot)^\dagger$  the conjugate transpose of any vector or matrix. For a  $M \times N$  matrix  $\mathbf{A} = [\mathbf{a}_1, \dots, \mathbf{a}_N]$ ,  $\text{vec}(\mathbf{A}) = [\mathbf{a}_1^t, \dots, \mathbf{a}_N^t]^t$  is a vector of length  $MN$ , while for a  $M \times 1$  vector  $\mathbf{a} = [a_1, \dots, a_M]$ ,  $\text{diag}(\mathbf{a})$  is an  $M \times M$  diagonal matrix with  $\mathbf{a}$ 's entries along its main diagonal. Finally by  $f(\mathbf{A})$  we denote a matrix whose  $\{i, j\}$  entry is  $f(a_{i,j})$ .

## II. SYSTEM MODEL

In the following, we consider the double directional model, which describes a time-variant, frequency-selective channel, taking into account the DoAs, the DoDs, the Delays, the Bandwidth and the transmitted and received powers. Based on the Maximum Entropy principle, Debbah et al. validated this model in [9]. The authors also proved in [10], that the double directional model encompasses the “Kronecker”, “Müller”, “Virtual Representation” and “keyhole” models as special cases. According to the model, the  $n_r \times n_t$  MIMO matrix  $\mathbf{H}$  is given by:

$$\mathbf{H}_{n_r \times n_t}(f, t) = \frac{1}{\sqrt{s_r s_t}} \hat{\Phi}_{n_r \times s_r}(t) \mathbf{P}_r(\Theta_{s_r \times s_t} \odot \mathbf{D}_{s_r \times s_t}(f)) \mathbf{P}_t \hat{\Psi}_{s_t \times n_t}(t) \quad (1)$$

where  $n_r, n_t, s_r, s_t$  are the number of receiving and transmitting antennas and the number of scatterers in the area of the receiver and the transmitter respectively (See figure 1). The

entries of  $\Theta$  are i.i.d. complex Gaussian with zero mean and unit variance.  $\mathbf{P}_r$  and  $\mathbf{P}_t$  are diagonal matrices containing the powers of the steering directions. The rest of the matrices on the right hand side (r.h.s.) of (1) are defined as :

$$\hat{\Phi} = e^{(\mathbf{B}^r + 2\pi \frac{f_c}{c} v_r \cos(\Phi - \alpha_r) \mathbf{I})t} \quad (2)$$

$$\hat{\Psi} = e^{(\mathbf{B}^t + 2\pi \frac{f_c}{c} v_t \cos(\Psi - \alpha_t) \mathbf{I})t} \quad (3)$$

$$\mathbf{D} = e^{-j2\pi f \mathbf{T}} \quad (4)$$

so that their  $(k, l)$  entries are:

$$\hat{\phi}_{k,l} = e^{j(\beta_{kl}^r + 2\pi \frac{f_c}{c} v_r \cos(\phi_{k,l} - \alpha_r)t)} \quad (5)$$

$$\hat{\psi}_{k,l} = e^{j(\beta_{kl}^t + 2\pi \frac{f_c}{c} v_t \cos(\psi_{k,l} - \alpha_t)t)} \quad (6)$$

$$d_{k,l} = e^{-j2\pi f \tau_{k,l}} \quad (7)$$

respectively.  $\beta_{kl}^r$  is the initial phase of the signal from scatterer  $l$  to receiving antenna  $k$  and  $\phi_{k,l}$  is the angle between a line perpendicular to the antenna array and the wavefront's direction.  $v_r$  and  $\alpha_r$  are the receiver's speed and direction. The parameters  $\beta_{kl}^t$  and  $\psi_{k,l}$  along with  $v_t$  and  $\alpha_t$  can be defined in a similar way for the transmitter.  $\tau_{k,l}$  is the unknown delay required for the signal to propagate from scatterer  $l$  at the transmitter's side to scatterer  $k$  at the receiver's side.

The factors  $2\pi \frac{f_c}{c} v_r \cos(\phi_{kl} - \alpha_r)t$  and  $2\pi \frac{f_c}{c} v_t \cos(\psi_{kl} - \alpha_t)t$  in the exponents of the entries of the matrices  $\hat{\Phi}$  and  $\hat{\Psi}$  respectively, represent the shift in frequency due to the movement of the receiver and the transmitter (doppler effect). We are particularly interested in the entries of  $\hat{\Phi}$ , since these depend on the speed  $v_r$  and the direction  $\alpha_r$  of the MT. Although the analysis will be carried out for the general channel model, in our simulations we focus on Uniform Linear Arrays (ULA) at both the receiver and the transmitter. These are of practical interest and simplify our channel model (by reducing the number of unknown random variables), so that the  $(k, l)$  entries of  $\hat{\Phi}$  and  $\hat{\Psi}$  become<sup>2</sup>:

$$\hat{\phi}_{k,l} = e^{j2\pi \left( \frac{d(k-1) \sin(\phi_l)}{\lambda} + \frac{f_c}{c} v_r \cos(\phi_l - \alpha_r)t \right)} \quad (8)$$

$$\hat{\psi}_{k,l} = e^{j2\pi \left( \frac{d(l-1) \sin(\psi_k)}{\lambda} + \frac{f_c}{c} v_t \cos(\psi_k - \alpha_t)t \right)} \quad (9)$$

$$(10)$$

where we have used the fact that under the far field approximation, the initial phase  $\beta_{kl}^r$  ( $\beta_{kl}^t$ ) varies linearly with the sine of  $\phi_{k,l}$  ( $\psi_{k,l}$ ).

### III. BAYESIAN ESTIMATION OF LOCATION-DEPENDENT PARAMETERS

We consider the estimation of the speed and the direction of the MT via the transmission of a training sequence. The design of an optimal<sup>3</sup> training sequence for frequency selective MIMO channels is beyond the scope of this work. The training sequence considered in this paper consists of a set of  $N$  orthogonal sub-vectors of size  $n_t$  each and is assumed to be

<sup>2</sup>Any two-dimensional antenna array can be considered instead of ULAs with a simple modification of the exponent of the entries of these matrices.

<sup>3</sup>A training sequence that minimizes the estimation error.

transmitted within the channel's coherence time  $T^4$ . Each sub-vector is given by  $\mathbf{x}_l = [\mathbf{0}_{(l-1)}^t, x, \mathbf{0}_{(nt-l)}^t]^t$ ,  $l = 1, \dots, N$ , where  $x$  is the training symbol which is known at the receiver. This training sequence was derived in [11] and was proven to be optimal<sup>5</sup> for MIMO OFDM systems with cyclic prefix.

The discrete-time model describing the input-output relationship of a time-variant frequency-selective MIMO channel is :

$$\mathbf{y}_l(f_i, t_j) = \mathbf{H}(f_i, t_j) \mathbf{x}_l + \mathbf{n}_l(f_i, t_j) \quad (11)$$

where  $i = 1, \dots, N_f$ ,  $j = 1, \dots, N_t$ .  $N_f$  and  $N_t$  are the numbers of different frequency and time samples respectively.  $\mathbf{n}$  is the  $n_r \times 1$  noise vector whose entries are i.i.d. complex Gaussian with mean 0 and variance  $\sigma^2$ .

Let  $\mathbf{X} = [\mathbf{x}_1, \dots, \mathbf{x}_N]$  denote the  $n_t \times N$  channel input training matrix,  $\mathbf{Y}(f_i, t_j) = [\mathbf{y}_1(f_i, t_j), \dots, \mathbf{y}_N(f_i, t_j)]$  denote the  $n_r \times N$  channel output matrix and  $\mathbf{N}(f_i, t_j) = [\mathbf{n}_1(f_i, t_j), \dots, \mathbf{n}_N(f_i, t_j)]$  denote the  $n_t \times N$  noise matrix. The input-output relationship becomes:

$$\mathbf{Y}(f_i, t_j) = \mathbf{H}(f_i, t_j) \mathbf{X} + \mathbf{N}(f_i, t_j) \quad (12)$$

Without loss of generality we will choose  $N = n_t$  and the training symbol  $x = 1$  so that the input training matrix is  $\mathbf{X} = \mathbf{I}_{n_t}$ . This reduces (12) to the simpler form:

$$\mathbf{Y}(f_i, t_j) = \mathbf{H}(f_i, t_j) + \mathbf{N}(f_i, t_j) \quad (13)$$

The joint conditional density of all the received matrices, denoted hereafter as  $f(\mathbf{Y} | \{\mathbf{H}_{11} \dots \mathbf{H}_{N_f N_t}\}) = f(\mathbf{Y} | \mathbb{S}_{\mathbf{H}})$ , is given by:

$$f(\mathbf{Y} | \mathbb{S}_{\mathbf{H}}) = \prod_{i=1}^{N_f} \prod_{j=1}^{N_t} \frac{1}{(\pi \sigma^2)^{n_r n_t}} e^{-\frac{|\mathbf{Y}(f_i, t_j) - \mathbf{H}(f_i, t_j)|^2}{\sigma^2}} \quad (14)$$

$$= \prod_{i=1}^{N_f} \prod_{j=1}^{N_t} \prod_{p=1}^{n_t} \prod_{q=1}^{n_r} \frac{1}{(\pi \sigma^2)} e^{-\frac{|y_{p,q}(f_i, t_j) - h_{p,q}(f_i, t_j)|^2}{\sigma^2}}$$

Based on (14) a Maximum Likelihood (ML) estimator can be implemented to compute the speed and the direction of a MT. The ML estimator is actually equivalent to the Maximum A-Posteriori (MAP) estimator, since both the speed and the direction have uniform a-priori distributions. Let  $\mathbf{p}_{int}$  be the vector containing the parameters to be estimated ("parameters of interest") and  $\mathbf{p}_{nuis}$  be the vector of all the rest parameters in  $\mathbf{H}$  ("nuisance parameters"). These two vectors will be explicitly defined later. For the time being it suffices to state that  $\mathbf{p}_{nuis} = [\boldsymbol{\theta}, \boldsymbol{\tau}, \dots]$ , where  $\boldsymbol{\theta} = \text{vec}(\Theta)$  and  $\boldsymbol{\tau} = \text{vec}(\mathbf{T})$ . Under the Bayesian Framework [12], we can infer on the a-priori distributions of all the random variables composing the entries of  $\mathbf{H}$ , therefore also composing  $\mathbf{p}_{int}$  and  $\mathbf{p}_{nuis}$ . Specifically the p.m.f. of the number of scatterers at both sides is  $f(s) = \frac{1}{N_f N_t}$ ,  $s = 1, \dots, N_f N_t$ , where  $s$  represents either

<sup>4</sup>We assume that the random matrices composing  $\mathbf{H}$ , do not change within  $T$  and choose to send a set of orthogonal training sub-vectors  $\mathbf{x}$  that span the column space of the MIMO matrix during this interval.

<sup>5</sup>In the sense that it maximizes capacity when used for channel estimation.

$s_r$  or  $s_t$ . The priors of the continuous random variables are given below:

$$\begin{aligned} \alpha_r, \alpha_t, \phi_{q,k}, \beta_{q,k}^r, \psi_{l,p}, \beta_{l,p}^t &\sim U[0, 2\pi] \\ v_r &\sim U[0, v_{max}^r], v_t \sim U[0, v_{max}^t] \\ \tau_{k,l} &\sim U[0, \tau_{max}], \theta_{k,l} \sim CN(0, 1) \\ p_{k,k}^r &\sim U[0, p_{max}^r], p_{l,l}^t \sim U[0, p_{max}^t] \end{aligned}$$

$\forall q = 1, \dots, n_r, k = 1, \dots, s_r, p = 1, \dots, n_t, l = 1, \dots, s_t$

According to ML estimation,

$$\hat{\mathbf{p}}_{int} = \underset{\mathbf{p}_{int}}{\operatorname{argmax}} f(\mathbf{Y}|\mathbf{p}_{int}) \quad (15)$$

where  $f(\mathbf{Y}|\mathbf{p}_{int})$  is the joint density of the received data matrices  $\mathbf{Y}(f_i, t_j)$  conditioned only on the parameters of interest. We can obtain this density by marginalizing over all the ‘‘nuisance parameters’’, according to:

$$f(\mathbf{Y}|\mathbf{p}_{int}) = \int_{\mathcal{A}} f(\mathbf{Y}|\mathcal{S}_{\mathbf{H}}) \cdot f(\mathbf{p}_{nuis}) d\mathbf{p}_{nuis} \quad (16)$$

since  $f(\mathbf{Y}|\mathcal{S}_{\mathbf{H}}) = f(\mathbf{Y}|\mathbf{p}_{int}, \mathbf{p}_{nuis})$ . By  $\mathcal{A}$  we denote the vector space where the vector  $\mathbf{p}_{nuis}$  lies.  $f(\mathbf{p}_{nuis})$  is the product of the a-priori densities of all the ‘‘nuisance parameters’’, some of which are conditioned on  $s_r, s_t$  or both. Since all the parameters except  $\boldsymbol{\theta}$  are uniformly distributed, we can write:

$$f(\mathbf{p}_{nuis}) = O(s_r, s_t) e^{-\boldsymbol{\theta}^\dagger \boldsymbol{\theta}} \quad (17)$$

We can proceed first by marginalizing over the Gaussian vector  $\boldsymbol{\theta}$  as follows:

$$\begin{aligned} f(\mathbf{Y}|\mathbf{p}_{int}) &= \int_{\tilde{\mathcal{A}}} \int_{\mathcal{C}^{s_r s_t}} f(\mathbf{Y}|\mathbf{p}_{int}, \mathbf{p}_{nuis}) f(\mathbf{p}_{nuis}) d\boldsymbol{\theta} d\tilde{\mathbf{p}}_{nuis} \\ &= \int_{\tilde{\mathcal{A}}} \int_{\mathcal{C}^{s_r s_t}} \prod_{i=1}^{N_f} \prod_{j=1}^{N_t} \prod_{p=1}^{n_t} \prod_{q=1}^{n_r} \left( \frac{1}{(\pi\sigma^2)} \right) \\ &\quad e^{-\frac{|y_{p,q}(f_i, t_j) - h_{p,q}(f_i, t_j)|^2}{\sigma^2}} O(s_r, s_t) e^{-\boldsymbol{\theta}^\dagger \boldsymbol{\theta}} d\boldsymbol{\theta} d\tilde{\mathbf{p}}_{nuis} \\ &= \int_{\tilde{\mathcal{A}}} \left[ \int_{\mathcal{C}^{s_r s_t}} \prod_{m=1}^M \left( \frac{1}{(\pi\sigma^2)} e^{-\frac{|y_m - \sum_{s=1}^S c_{m,s} \theta_s|^2}{\sigma^2}} \right) \right. \\ &\quad \left. e^{-\sum_{s=1}^S |\theta_s|^2} d\boldsymbol{\theta} \right] O(s_r, s_t) d\tilde{\mathbf{p}}_{nuis} \\ &= \int_{\tilde{\mathcal{A}}} \mathcal{J}_{\boldsymbol{\theta}} O(s_r, s_t) d\tilde{\mathbf{p}}_{nuis} \end{aligned} \quad (18)$$

where  $\tilde{\mathbf{p}}_{nuis}$  is vector comprised of all the elements of  $\mathbf{p}_{nuis}$  except the  $\boldsymbol{\theta}$  and lies on the space  $\tilde{\mathcal{A}}$ , while the vector  $\boldsymbol{\theta}$  lies on  $\mathcal{C}^{s_r s_t}$ . In the second equality above we used the fact that according to our model,  $h_{p,q}(f_i, t_j)$  can be expressed as:

$$\begin{aligned} h_{p,q}(f_i, t_j) &= \\ &= \frac{1}{\sqrt{s_r s_t}} \sum_{k=1}^{s_r} \sum_{l=1}^{s_t} p_{k,k}^r \hat{\phi}_{p,k}(t_j) \hat{\psi}_{l,q}(t_j) p_{l,l}^t d_{k,l}(f_i) \theta_{k,l} \\ &= \sum_{k=1}^{s_r} \sum_{l=1}^{s_t} c_{p,q,k,l}(f_i, t_j) \theta_{k,l} = \sum_{s=1}^S c_{m,s} \theta_s \end{aligned} \quad (19)$$

where, for ease of notation, we replaced the subscripts  $p, q, i, j$  with a single subscript  $m = 1, \dots, M$  and the subscripts  $k, l$  with  $s = 1, \dots, S$ .  $M = N_f N_t n_t n_r$  and  $S = s_r s_t$ . Introducing  $\mathbf{c}_m \triangleq [c_{m,1}, \dots, c_{m,S}]^t$ ,  $\mathbf{c}'_m \triangleq y_m \mathbf{c}_m^*$  and  $\mathbf{C}_m \triangleq \mathbf{c}_m^* \mathbf{c}_m^t$  we can show that:

$$e^{-\frac{1}{\sigma^2} |y_m - \sum_{s=1}^S c_{m,s} \theta_s|^2} = e^{-\frac{|y_m|^2}{\sigma^2}} e^{-\frac{1}{\sigma^2} [-\mathbf{c}'_m \boldsymbol{\theta} - \boldsymbol{\theta}^\dagger \mathbf{c}'_m + \boldsymbol{\theta}^\dagger \mathbf{C}_m \boldsymbol{\theta}]} \quad (20)$$

In the following analysis we will use the subscript ‘‘all’’ to denote a vector or a matrix that is equal to the sum of all  $M$  vectors or matrices, respectively, which are represented by the same symbol. After substituting the r.h.s. of (20) in  $\mathcal{J}_{\boldsymbol{\theta}}$  we can proceed as follows:

$$\begin{aligned} \mathcal{J}_{\boldsymbol{\theta}} &= \int_{\mathcal{C}^{s_r s_t}} \frac{1}{(\pi\sigma^2)^M} e^{-\frac{1}{\sigma^2} \sum_{m=1}^M |y_m|^2} \\ &\quad e^{-\frac{1}{\sigma^2} \sum_{m=1}^M [-\mathbf{c}'_m \boldsymbol{\theta} - \boldsymbol{\theta}^\dagger \mathbf{c}'_m + \boldsymbol{\theta}^\dagger \mathbf{C}_m \boldsymbol{\theta}]} e^{-\boldsymbol{\theta}^\dagger \boldsymbol{\theta}} d\boldsymbol{\theta} \\ &= \frac{1}{(\pi\sigma^2)^M} e^{-\frac{1}{\sigma^2} \mathbf{y}^\dagger \mathbf{y}} \\ &\quad \int_{\mathcal{C}^{s_r s_t}} e^{-\frac{1}{\sigma^2} [-\mathbf{c}'_{all} \boldsymbol{\theta} - \boldsymbol{\theta}^\dagger \mathbf{c}'_{all} + \boldsymbol{\theta}^\dagger \mathbf{C}_{all} \boldsymbol{\theta} + \sigma^2 \boldsymbol{\theta}^\dagger \boldsymbol{\theta}]} d\boldsymbol{\theta} \end{aligned}$$

where the data vector  $\mathbf{y} = [y_1, \dots, y_M]^t$  contains all the received signal values, over different time and frequency samples in different receiving antennas and for different transmitted training vectors. Also  $\mathbf{C}_{all} = \sum_{m=1}^M \mathbf{C}_m = \sum_{m=1}^M \mathbf{c}_m^* \mathbf{c}_m^t$  is an  $S \times S$  Hermitian matrix of rank  $r = \min\{M, S\}$ . Let  $\mathbf{C}'_{all} = \mathbf{C}_{all} + \sigma^2 \mathbf{I}$ .  $\mathbf{C}'_{all}$  is also a Hermitian matrix of rank  $S$  (full rank). Therefore its inverse exists and is also Hermitian and positive definite. Using this fact we can integrate over  $\boldsymbol{\theta}$  and get an explicit expression for  $\mathcal{J}_{\boldsymbol{\theta}}$ :

$$\begin{aligned} \mathcal{J}_{\boldsymbol{\theta}} &= \frac{\det(\mathbf{C}'_{all}^{-1})}{(\pi\sigma^2)^{(M-S)}} e^{-\frac{1}{\sigma^2} \mathbf{y}^\dagger \mathbf{y}} e^{\frac{1}{\sigma^2} (\mathbf{c}'_{all} \mathbf{C}'_{all}^{-1} \mathbf{c}'_{all})} \\ &= \frac{\det((\mathbf{C}_G^* \mathbf{C}_G^t + \sigma^2 \mathbf{I})^{-1})}{(\pi\sigma^2)^{(M-S)}} e^{-(\mathbf{y}^\dagger (\mathbf{C}_G^t \mathbf{C}_G^* + \sigma^2 \mathbf{I})^{-1} \mathbf{y})} \end{aligned} \quad (21)$$

where we have introduced the  $S \times M$  matrix  $\mathbf{C}_G$ :

$$\mathbf{C}_G = \begin{bmatrix} \uparrow & \dots & \uparrow \\ \mathbf{c}_1 & \dots & \mathbf{c}_M \\ \downarrow & \dots & \downarrow \end{bmatrix} \quad (22)$$

and we have applied the matrix inversion lemma. After careful inspection we can write  $\mathbf{C}_G$  as<sup>6</sup>

$$\mathbf{C}_G = [ \mathbf{C}_{L,(1,1)} \quad \dots \quad \mathbf{C}_{L,(N_f, N_t)} ] \quad (23)$$

with each submatrix given by:

$$\mathbf{C}_{L,(i,j)} = \operatorname{diag}(\mathbf{d}_i) (\tilde{\boldsymbol{\Psi}}_j \otimes \tilde{\boldsymbol{\Phi}}_j^t) \quad (24)$$

and  $\mathbf{d}_i = \operatorname{vec}(\mathbf{D}(f_i))$ ,  $\tilde{\boldsymbol{\Psi}}_j = \mathbf{P}_t \boldsymbol{\Psi}(t_j)$  and  $\tilde{\boldsymbol{\Phi}}_j^t = \hat{\boldsymbol{\Phi}}(t_j) \mathbf{P}_r$ . From (23) and (24) the dependency of  $\mathbf{C}_G$  on  $\tilde{\mathbf{p}}_{nuis}$  becomes apparent. Substituting  $\mathcal{J}_{\boldsymbol{\theta}}$  in  $f(\mathbf{Y}|\mathbf{p}_{int})$  we finally obtain:

<sup>6</sup>This representation is not unique. By permuting the rows and/or the columns we get  $S!M!$  equivalent representations but we should permute the elements of  $\mathbf{y}$  as well.

$$f(\mathbf{Y}|\mathbf{p}_{int}) = \int_{\tilde{A}} \mathcal{O}'(s_r, s_t) \det((\mathbf{C}_G^* \mathbf{C}_G^t + \sigma^2 \mathbf{I})^{-1}) e^{-\mathbf{y}^\dagger (\mathbf{C}_G^* \mathbf{C}_G^t + \sigma^2 \mathbf{I})^{-1} \mathbf{y}} d\tilde{\mathbf{p}}_{nuis} \quad (25)$$

Furthermore if we consider the case of  $N_f = 1$ , i.e. the case when we sample the frequency response of the time-varying channel<sup>7</sup> at only one frequency, the r.h.s. of eq. (25) does not depend on  $\tau$  either. This is not surprising, since if  $N_f = 1$  we can replace  $(\Theta_{s_r \times s_t} \odot \mathbf{D}_{s_r \times s_t}(f))$  with a new matrix  $\Theta'_{s_r \times s_t}$  that has the same distribution.

Notice that so far no explicit categorization of the parameters has been made (except from the unknown amplitudes contained in  $\theta$  and the delays contained in  $\tau$ ). One can define  $\mathbf{p}_{int}$  to contain just the parameters that need to be estimated for tracking the mobile or even choose to include some nuisance parameters and jointly estimate all. Whether the Bayesian estimation, based on the marginal pdf, or the joint estimation will yield better results is not trivial to show analytically. Marginalization would require integration over a subspace with many dimensions and is not guaranteed to always result in high accuracy. On the other hand joint estimation would lead to an algorithm with very high computational complexity, since we would need to keep track of a multivariate density. Thus, we chose to sacrifice optimality for efficiency and do a step-by-step Bayesian estimation by recognizing that the DoAs  $\phi$  and the DoDs  $\psi$  must be estimated prior to or jointly with  $v_r$  and  $\alpha_r$  for the tracking method to give accurate results. The proposed algorithm for a MIMO system is summarized below:

- Set  $\mathbf{p}_{int} = [s_r, \phi]$ ,  $\tilde{\mathbf{p}}_{nuis} = [s_t, \psi, \mathbf{p}^r, \mathbf{p}^t]$  and use (25) to find the DoAs, using only one observation,  $N_f = N_t = 1$ .
- Set  $\mathbf{p}_{int} = [s_t, \psi]$ ,  $\tilde{\mathbf{p}}_{nuis} = [s_r, \phi, \mathbf{p}^r, \mathbf{p}^t]$  and use (25) to find the DoDs, using only one observation,  $N_f = N_t = 1$ .
- Set  $\mathbf{p}_{int} = [v_r, \alpha_r]$ ,  $\tilde{\mathbf{p}}_{nuis} = [v_t, \alpha_t, \mathbf{p}^r, \mathbf{p}^t]$ <sup>8</sup> and use (25) to estimate our true parameters of interest, using all observations.

The reason for using only spatial (and no temporal) information to estimate DoAs and DoDs is that the terms due to the doppler frequency shift cancel out of the expression (this is why  $v_r, \alpha_r, v_t, \alpha_t$  are not contained in  $\tilde{\mathbf{p}}_{nuis}$ ) leading to fewer nuisance parameters and higher estimation accuracy. In steps one and two of the above method, the algorithms in [2], [3], [4] could be employed instead. However that would require the implementation of a separate algorithm and the whole method is not guaranteed to yield higher accuracy, especially if a multi-dimensional -instead of a 1-dimensional- maximum likelihood estimation is adopted. For a system with a single transmit and a single receive antenna (SISO), only the third step can be implemented.

<sup>7</sup>Obtained by taking the Fourier transform with respect to the delay  $\tau$ .

<sup>8</sup>The discretization error (if any) of  $\phi, \psi$  could be modeled as uniformly distributed in  $[-\epsilon, \epsilon]$  and included in  $\tilde{\mathbf{p}}_{nuis}$ .

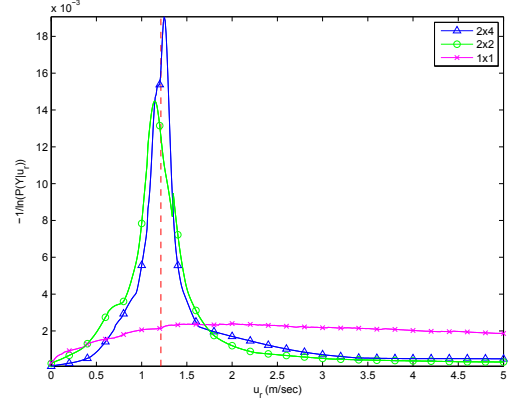


Fig. 2. Log-likelihood for estimating the speed  $v_r$  at 20dB

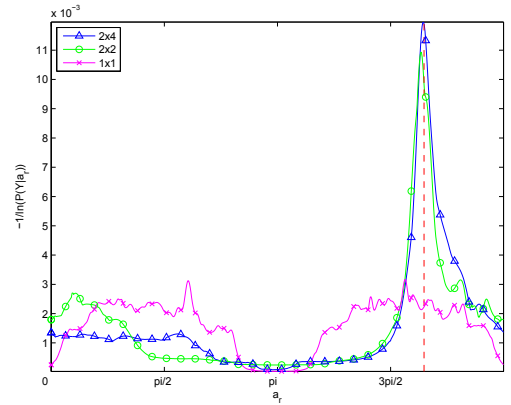


Fig. 3. Log-likelihood for estimating the direction  $\alpha_r$  at 20dB

#### IV. SIMULATION RESULTS

For sake of simplicity, we assume that the BS is fixed, i.e.  $v_t = \alpha_t = 0^\circ$ . We further assume that the main lobe of the transmitting and the receiving antenna array is steered to a direction perpendicular to the array and has a beamwidth of  $180^\circ$ . The energy of the signal components transmitted to or received from other directions is negligible. This implies that the DoAs and the DoDs are either in  $[-\pi/2, \pi/2]$  or in  $[\pi/2, 3\pi/2]$ . The power gains of the steering directions are also assumed to be known. To compute the value of the multidimensional integral in (25), Monte Carlo simulations have been performed. Normally 100 iterations are enough for the algorithm to converge to the true density. To make our graphs more clear and emphasize our results, we have plotted the 1-dimensional normalized log-likelihoods  $-\frac{1}{\ln f(\mathbf{Y}|v_r)}$  and  $-\frac{1}{\ln f(\mathbf{Y}|\alpha_r)}$  as a function of  $v_r$  and  $\alpha_r$ , respectively, for different  $n_r \times n_t$  systems. The vertical dashed line depicts the true value.

In figures 2 and 3 the advantage of MIMO over SISO at high SNR (20dB) is clearly illustrated. With just one antenna

<sup>9</sup>Or  $v_r = \alpha_r = 0$  if we consider transmission in the uplink.



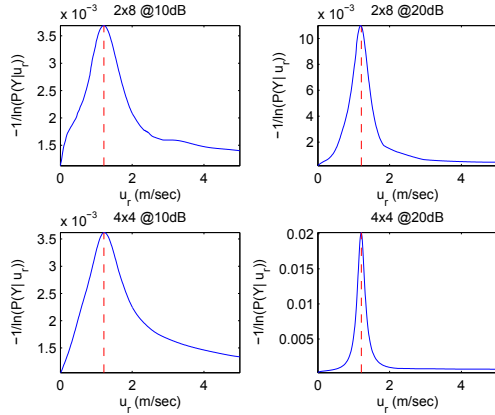


Fig. 4. Log-likelihood for estimating the speed  $v_r$ .

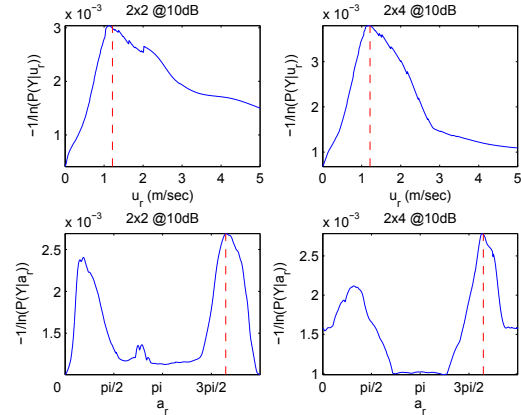


Fig. 6. Performance at low SNR

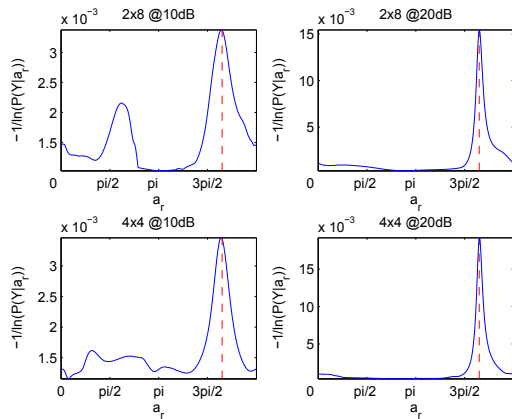


Fig. 5. Log-likelihood for estimating the direction  $\alpha_r$ .

at each side of the communication link, it is almost impossible to track the mobile. With as many as 2 antennas at each side, the estimation error becomes already very small. In figures 4 and 5 we show that even with noisy measurements (SNR=10dB),  $v_r$  and  $\alpha_r$  can be estimated correctly in a  $4 \times 4$  and a  $2 \times 8$  system. On the other hand in a  $2 \times 2$  or a  $2 \times 4$  system our parameters of interest are slightly misestimated at low SNR as shown in figure 6. Our results indicate that decreasing the SNR or the number of antennas leads to an increase in the variance of the estimated parameter (and thus of the estimation error). It further results in a second peak in the log-likelihood corresponding to  $\alpha_r$ . This stems from the fact that our expression (r.h.s. of 25) depends on  $\alpha_r$  only through its cosine and  $\cos(\phi_l - \alpha_r) = \cos(\alpha_r - \phi_l)$ ,  $\forall l$ . Thus if most of the DoAs cannot be estimated (SISO case or MIMO with just a few antennas) or if their effect cannot be removed successfully by integration, this ambiguity cannot be resolved.

## V. CONCLUSIONS - REMARKS

In this contribution we have pointed out how the speed and the direction of a MT can be estimated accurately, by using an

appropriate, for NLOS conditions, channel model and exploiting the coefficients corresponding to the Doppler frequency shift with the aid of the Bayesian framework. We have shown the enhanced performance of a MIMO over a SISO system, in terms of estimation accuracy. This enhancement mainly stems from the fact that with a MIMO system, information contained in the space dimension can be exploited to estimate nuisance parameters (like the DoAs and the DoDs) and then use their values as prior knowledge. In future work we intend to show how the presence of a LoS path can be integrated in our method.

## REFERENCES

- [1] I. E. Telatar, "Capacity of multi-antenna gaussian channels," *European Transactions on Telecommunications*, vol. 10, no. 6, pp. 585 – 595, 1999.
- [2] R. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Trans. Antennas Propagat.*, vol. 34, no. 3, pp. 276 – 280, 1986.
- [3] R. Roy and T. Kailath, "Esprit-estimation of signal parameters via rotational invariance techniques," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 37, no. 7, pp. 984 – 995, 1989.
- [4] H. Miao, K. Yu, and M. J. Juntti, "2-d unitary esprit based joint aoa and aod estimation for mimo system," in *Proc. IEEE 17th International Symposium on Personal, Indoor and Mobile Radio Communications*, Sept. 2006.
- [5] L. Xiong, "A selective model to suppress nlos signals in angle-of-arrival(aoa) location estimation," in *Proc. 9th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications*.
- [6] Y. Qi, H. Kobayashi, and H. Suda, "Analysis of wireless geolocation in a non-line-of-sight environment," *IEEE Trans. Wireless Commun.*, vol. 5, no. 3, pp. 672 – 681, 2006.
- [7] S. Al-Jazzar and J. J. Caffery, "ML and bayesian toa location estimators for nlos environments," in *Proc. 56th IEEE Vehicular Technology Conference*, vol. 2, Sept. 2002, pp. 1178– 1181.
- [8] H. Miao, K. Yu, and M. J. Juntti, "Positioning for nlos propagation: Algorithm derivations and cramer-rao bounds," *IEEE Trans. Veh. Technol.*, vol. 56, no. 5, pp. 2568 – 2580, 2007.
- [9] M. Debbah and R. R. Muller, "Capacity complying mimo channel models," in *37th Annual Asilomar Conference on Signals, Systems and Computers, California, USA, 2003*.
- [10] —, "Mimo channel modeling and the principle of maximum entropy," *IEEE Trans. Inform. Theory*, vol. 51, no. 5, pp. 1667 – 1690, 2005.
- [11] X. Ma, L. Yang, and G. B. Giannakis, "Optimal training for mimo frequency-selective fading channels," *IEEE Trans. Wireless Commun.*, vol. 4, no. 2, pp. 453 – 466, 2005.
- [12] E. Jaynes, *Probability Theory: The logic of science*. Cambridge University Press.