Parametric Least Squares Estimation for Nonlinear Satellite Channels

Lei Xiao
EURECOM
Mobile Communications Department
BP193, F-06560 Sophia Antipolis, France
Email: lei.xiao@eurecom.fr

Laura Cottatellucci
EURECOM
Mobile Communications Department
BP193, F-06560 Sophia Antipolis, France
Email: laura.cottatellucci@eurecom.fr

***Abstract*—We consider a multiuser MIMO Mobile Satellite System (MSS) and model its channel as a cascade of a slow varying component, directivity vector, and a fast fading component, propagation component. We study the estimation of the slow varying part of the satellite channel at the gateway. Since the channel model is nonlinear, we propose a nonlinear parametric least squares approach. This optimization problem is shown to be equivalent to an eigenvalue complementary problem. The equivalent problem does not require an intermediate estimation of the nuisance (fast fading component) with relevant benefits in terms of computational complexity. The performance of the proposed algorithm is assessed by simulations based on realistic satellite channels.**

I. Introduction

In modern multibeam satellite systems (MSS), thanks to the improvements in the switching speed of beamforming networks (BFN), adaptive beamforming for mobile terminals is nowadays a realistic option.

The knowledge of channel state information (CSI) at the gateway is critical for the design of an adaptive beamformer. Therefore, CSI acquisition becomes a crucial problem and strongly depends on the channel characteristics.

In satellite systems, the fading components of the channel are highly variable, the channel coherence time is much shorter than the round trip delay of the signal, and possible feedback of the CSI is already stale when received. In this situation, channel distribution information (CDI) provides a practical solution since the channel statistics change slower than the CSI.

In a satellite system, the CDI can be estimated at the STs and fed back to the gateway or can be estimated at the gateway if channel reciprocity holds. The latter approach presents well known benefits in terms of system spectral efficiency since a feedback channel is not required. In this contribution, we assume that channel reciprocity holds, at least from a statistical point of view, and the CDI is estimated at the gateway.

The acquisition of the CDI at the gateway, for a satellite system with mobile satellite terminals (ST) equipped eventually with multiple antennas and transmitting in left and right polarization, presents completely new challenges compared to the thoroughly studied field of satellite channel estimation finalized to the coherent detection and decoding of the channel at the receiver side and it is a completely unexplored field. By assuming that the statistics of the propagation coefficients are available at the gateway, the CDI estimation reduces to the estimation of the slow varying components of the fading channel. From a signal processing perspective, this implies the challenging task of estimating parameters observed through multiplicative nuisance.

The estimation of the directivity vectors is intrinsically nonlinear. We consider a parametric model of channels where the directivity vector (slow fading components) is parametrically represented by a linear combination of given known directivity vectors and the varying propagation coefficients (fast fading components) play the role of multiplicative nuisance parameters. In this work, we propose an algorithm to estimate the directivity vector parameters based on a least squares criterion. We show that the estimation problem reduces to an eigenvalue complementary problem. We dub the proposed algorithm Parametric Least Squares Estimation (PLSE). The proposed algorithm does not require the estimation of nuisance parameters and this enables a considerable complexity reduction.

Throughout this article, we adopt the following notations. Vectors are written in boldface lower case letters; matrices in boldface capital letters. Superscripts ^{T}, *, ^{H} denote transposition, elementwise conjugation, conjugate transposition of a matrix, respectively. $\mathfrak{Re}(\cdot)$ denotes the real part operator and $\|\cdot\|_l$ denotes the norm l vector. Shortly, $\|\cdot\|$ denotes the Euclidean norm. $\boldsymbol{A}^{(\backsim i)}$ represents the submatrix of the matrix $\boldsymbol{A}$ obtained by removing the i-th column and the i-th row.

II. System Model

The modeling of a multi-antenna satellite system channels with satellite antenna mobility is currently object of intense research. An updated overview of the ongoing studies and recent results about the channel modeling can be found in [1]. We follow the channel model proposed by [2] and refer to it as Surrey model throughout this work. Thus, the channel is modeled as a cascade, i.e. analytically a multiplication, of two different components: (a) directivity vector between a satellite terminal (SA) and the satellite and (b) propagation coefficients. The directivity vector depends on the radiation patterns of the SAs and the STs' positions. The propagation coefficients model the propagation losses (atmospheric and shadowing) between satellite and ST.

We consider a satellite system consisting of a gateway, a bent-pipe satellite equipped with N antennas (SA) and K STs endowed with R antennas. All antennas transmit in left and right polarizations. The discrete-time baseband received signal at the gateway at time t is given by

$$\boldsymbol{y}[t] = \boldsymbol{D}[t]\boldsymbol{P}[t]\boldsymbol{x}[t] + \boldsymbol{z}[t], \tag{1}$$

where $\boldsymbol{y}[t]$ is the column vector of received signals at the gateway, $\boldsymbol{D}[t]$ is the directivity matrix, $\boldsymbol{P}[t]$ is the propagation matrix, $\boldsymbol{x}[t]$ is the $2RK$ vector of transmitted signals, and $\boldsymbol{z}[t]$ is the additive noise vector introduced at the gateway[1]. The noise vector is a zero mean white Gaussian process with covariance matrix $\sigma_z^2\boldsymbol{I}$.

Let $\boldsymbol{x}_k[t]$ be the $2R$-dimensional vector of symbols transmitted in left and right polarization by the R antennas of ST k. Then, the vector $\boldsymbol{x}[t]$ of transmitted signals is obtained by stacking together the K vectors $\boldsymbol{x}_k[t]$, i.e.,

$$\boldsymbol{x}[t] = \left(\boldsymbol{x}_1^T[t], \boldsymbol{x}_2^T[t], ..., \boldsymbol{x}_K^T[t]\right)^T . \tag{2}$$

The propagation matrix $\boldsymbol{P}[t]$ is a block diagonal matrix with K independent blocks $\boldsymbol{P}^k[t]$ of size $2 \times 2R$ and form

$$\boldsymbol{P}^k[t] = \begin{pmatrix} P_{k,r}^{(1)}[t] & 0 & \cdots & P_{k,r}^{(R)}[t] & 0 \\ 0 & P_{k,l}^{(1)}[t] & \cdots & 0 & P_{k,l}^{(R)}[t] \end{pmatrix},$$

where $P_{k,o}^{(\ell)}[t]$ denotes the fast fading coefficient affecting the link between the satellite and antenna ℓ at ST k in o-polarization[2]. It is worth noting that this fading component is due to local perturbation of the signals around the ST. Due to the very large distance between SAs and a ST, and propagation in deep space, the same fast fading components affects the signals from all SAs to a single antenna in a certain polarization. We make the realistic assumption that the variations of directivity vectors due to ST movements are negligible in the time interval when the channel is measured for estimation. Thus, we assume that the directivity vectors are constant in our system model and we drop the time index in the matrix $\boldsymbol{D}[t]$. The directivity matrix $\boldsymbol{D}$ can conveniently be structured in KN blocks of form

$$\boldsymbol{D}_n^k = \begin{pmatrix} d_{n,rr}^k & d_{n,rl}^k \\ d_{n,lr}^k & d_{n,ll}^k \end{pmatrix} = \begin{pmatrix} \boldsymbol{d}_{n,r}^k \\ \boldsymbol{d}_{n,l}^k \end{pmatrix}, \tag{3}$$

where $d_{n,ov}^k$, with $o, v \in \{r, l\}$ represents the directivity coefficient of SA n in o polarization in direction of ST k in v polarization; $d_{n,rl}^k$ and $d_{n,lr}^k$ are cross polarizations; $d_{n,rr}^k$ and $d_{n,ll}^k$ are co-polarizations. Then, $\boldsymbol{D}_n^k$ describes the static part of the channel between ST k and SA n and $\boldsymbol{d}_{n,o}^k = (d_{n,or}^k, d_{n,ol}^k)$ is the component in o-polarization at SA n. The block column of size $2N \times 2$, $\boldsymbol{D}^k = (\boldsymbol{D}_1^{kT}, \boldsymbol{D}_2^{kT}, \dots \boldsymbol{D}_N^{kT})^T$ represents the directivity coefficients of ST k. It is common to assume $d_{n,rr}^k = d_{n,ll}^k$ and $d_{n,rl}^k = d_{n,lr}^k$.

The directivity vector corresponding to a certain ST is determined by two factors: the geographic position of the ST and the frequency carrier. Interestingly, the effects of the frequency carrier on the directivity vectors are minor. They can be neglected in a given satellite system, e.g., in Ka band or Ku band. This implies that we can benefit from directivity reciprocity both in Time and Frequency Division Duplex (TDD/FDD) mode, and not only in TDD mode, as in terrestrial mobile communications.

Throughout this work, we make two realistic assumptions: (a) the directivity vectors of some reference STs in a grid are known at the gateway. [3] We denote by $\boldsymbol{G}$ the matrix available at the gateway and containing all the directivity vectors of the points in the grid. The matrix $\boldsymbol{G}$ has a block structure similar to the one of $\boldsymbol{D}$ with blocks $\boldsymbol{G}_n^k$ of form (3); (b) the directivity vector of a ST in an arbitrary position can be determined as a convex combination of the directivity vectors at some reference points. More specifically, let us consider ST k with coordinates $S_k \equiv (x, y)$, and let $G_{\pi(i)} \equiv (a_{\pi(i)}, b_{\pi(i)})$, with $i = 1, 2, 3$, be the three nearest reference points surrounding ST S. The point S_k can be expressed as convex combination of $G_{\pi(1)}, G_{\pi(2)}$, and $G_{\pi(3)}$, i.e.

$$S_k = \alpha_1^k G_{\pi(1)} + \alpha_2^k G_{\pi(2)} + \alpha_3^k G_{\pi(3)}$$

with $0 \leq \alpha_i^k \leq 1$, for $i = \{1, 2, 3\}$, and $\sum_{i=1}^3 \alpha_i^k = 1$. If $\boldsymbol{G}^{\pi(i)}$ denotes the $\pi(i)$ block column of $\boldsymbol{G}$ corresponding to point $G_{\pi(i)}$, then, the directivity column block $\boldsymbol{D}^k$ of ST k is given by convex combination of the directivity column vectors with identical coefficients

$$\boldsymbol{D}^k = \alpha_1^k \boldsymbol{G}^{\pi(1)} + \alpha_2^k \boldsymbol{G}^{\pi(2)} + \alpha_3^k \boldsymbol{G}^{\pi(3)}. \tag{4}$$

The estimation of the directivity matrix $\boldsymbol{D}$ is based on the synchronous transmissions of pilot sequences by all active STs. ST k transmits $2R$ pilot sequences of length L, one for each antenna and polarization. They are known by the gateway and differ each other and from the pilot sequences assigned to other STs. The pilot sequences are transmitted during a time slot not longer than the coherence time of the channel. Thus, in a time slot, the propagation matrix is constant and we denote the constant values in time slot q as $\boldsymbol{P}^k(q)$ and $\boldsymbol{P}(q)$ for ST k and all the STs, respectively. Observations over Q different time slots are utilized for the estimation. Under these assumptions, the signal received at SA n in o-polarization, with $o \in \{l, r\}$, is given by

$$y_{n,o}[s_q + s] = \boldsymbol{d}_{n,o}\boldsymbol{P}(q)\boldsymbol{x}[s_q + s] + z_{n,o}[s_q + s], \tag{5}$$

where $\boldsymbol{d}_{n,o} = (\boldsymbol{d}_{n,o}^1, \boldsymbol{d}_{n,o}^2, \dots, \boldsymbol{d}_{n,o}^K)$, s_q is the time offset when the transmission of a pilot sequence for the qth slot starts and $s = 0, \dots, L-1$ is a time index. The observation signal $\boldsymbol{\mathcal{Y}}_{n,o}(q) = (y_{n,o}[s_q], y_{n,o}[s_q+1], \dots, y_{n,o}[s_q+L-1])$ in the coherence time q at SA n and o-polarization, is given by

$$\boldsymbol{\mathcal{Y}}_{n,o}(q) = \boldsymbol{d}_{n,o}\boldsymbol{P}(q)\boldsymbol{X}_q + \boldsymbol{\mathcal{Z}}_{n,o}(q), \tag{6}$$

[1]In this model, the attenuation between satellite and gateway is neglected and the channel link satellite-gateway is modeled as an additive white Gaussian channel. Additional noise introduced at the satellite antenna (e.g. intermodulation noise) is not explicitly considered in this model but it can be taken into account in the additive white noise at the gateway.

[2]In this model we assume that the signal leakage from left to right polarization and vice versa is negligible at the STs.

[3]We recall that they depend on the SAs' radiation patterns.

where $\boldsymbol{X}_q$ is the $2RK \times L$ matrix whose rows are the pilot sequences of the active STs and $\boldsymbol{\mathcal{Z}}_{n,o}(q)$ is the L-dimensional row vector of the noise $\boldsymbol{\mathcal{Z}}_{n,o}(q) = (z_{n,o}[s_q], z_{n,o}[s_q+1], \ldots, z_{n,o}[s_q+L-1])$.

III. Directivity Estimation

In this section, we describe our approach to the estimation of the directivity vectors. It consists of two steps. In the first step, we perform a standard linear estimation of the transfer channel matrix based on standard linear least squares estimation (LSE) in each time slot. The second step consists of a nonlinear estimation of the directivity vectors based on a least squares error criterion.

Let $\boldsymbol{h}_{n,r}(q)$ and $\boldsymbol{h}_{n,l}(q)$ be the transfer vectors from all the ST to SA n at time slot q in left and right polarization, respectively. They consist of K blocks $\boldsymbol{h}^k_{n,r}(q)$ and $\boldsymbol{h}^k_{n,l}(q)$ defined as

$$\begin{aligned}\boldsymbol{h}^k_{n,r}(q) &= \left(h^{k,(1)}_{n,rr}(q),\, h^{k,(1)}_{n,rl}(q), \cdots,\, h^{k,(R)}_{n,rr}(q),\, h^{k,(R)}_{n,rl}(q)\right)\\ &= \left(d^k_{n,rr}P^{(1)}_{k,r}(q),\, d^k_{n,rl}P^{(1)}_{k,l}(q), \cdots,\, d^k_{n,rr}P^{(R)}_{k,r}(q),\, d^k_{n,rl}P^{(R)}_{k,l}(q)\right)\end{aligned}$$

and

$$\begin{aligned}\boldsymbol{h}^k_{n,l}(q) &= \left(h^{k,(1)}_{n,lr}(q),\, h^{k,(1)}_{n,ll}(q), \cdots,\, h^{k,(R)}_{n,lr}(q),\, h^{k,(R)}_{n,ll}(q)\right)\\ &= \left(d^k_{n,lr}P^{(1)}_{k,r}(q),\, d^k_{n,ll}P^{(1)}_{k,l}(q), \cdots,\, d^k_{n,lr}P^{(R)}_{k,r}(q),\, d^k_{n,ll}P^{(R)}_{k,l}(q)\right),\end{aligned}$$

respectively. Then, (6) reduces to

$$\boldsymbol{\mathcal{Y}}_{n,o}(q) = \boldsymbol{h}_{n,o}(q)\boldsymbol{X}_q + \boldsymbol{\mathcal{Z}}_{n,o}(q). \tag{7}$$

By applying standard results on linear LSE (see e.g. [3]), we obtain the LSE estimation of $\boldsymbol{h}_{n,r}(q)$ and $\boldsymbol{h}_{n,l}(q)$ given by

$$\hat{\boldsymbol{h}}_{n,o}(q) = \boldsymbol{\mathcal{Y}}_{n,o}\boldsymbol{X}^H_q(\boldsymbol{X}_q\boldsymbol{X}^H_q)^{-1} \quad \{o\} = \{r,l\}. \tag{8}$$

The estimation error is $\boldsymbol{\varepsilon}_{n,o}(q) = \hat{\boldsymbol{h}}_{no}(q) - \boldsymbol{h}_{no}(q)$, $o = r,l$. By rearranging the components in $\hat{\boldsymbol{h}}_{n,r}(q)$ and $\hat{\boldsymbol{h}}_{n,l}(q)$ and utilizing the assumptions $d^k_{n,ll} = d^k_{n,rr}$ and $d^k_{n,lr} = d^k_{n,rl}$, we obtain the system of equations

$$\left\{\begin{array}{lcl} d^k_{n,rr}P^{(1)}_{k,r}(q) &=& \hat{h}^{k,(1)}_{n,rr}(q) + \varepsilon^{k,(1)}_{n,rr}(q)\\ d^k_{n,lr}P^{(1)}_{k,r}(q) &=& \hat{h}^{k,(1)}_{n,lr}(q) + \varepsilon^{k,(1)}_{n,lr}(q)\\ d^k_{n,rr}P^{(1)}_{k,l}(q) &=& \hat{h}^{k,(1)}_{n,ll}(q) + \varepsilon^{k,(1)}_{n,ll}(q)\\ d^k_{n,rl}P^{(1)}_{k,l}(q) &=& \hat{h}^{k,(1)}_{n,rl}(q) + \varepsilon^{k,(1)}_{n,rl}(q)\\ &\vdots& \\ d^k_{n,rr}P^{(R)}_{k,l}(q) &=& \hat{h}^{k,(R)}_{n,ll}(q) + \varepsilon^{k,(R)}_{n,ll}(q)\\ d^k_{n,rl}P^{(R)}_{k,l}(q) &=& \hat{h}^{k,(R)}_{n,rl}(q) + \varepsilon^{k,(R)}_{n,rl}(q), \end{array}\right. \tag{9}$$

where the indices of the components of the estimates and the estimation error vectors $\hat{\boldsymbol{h}}^k_{n,o}(q)$ and $\boldsymbol{\varepsilon}^k_{n,o}(q)$ are defined consistently with the ones of vector $\boldsymbol{h}^k_{n,o}(q)$. By making use of (4), we express (9) in a matrix form as function of the channel parameters α^k_1, α^k_2 and α^k_3. Let us define the vector $\boldsymbol{\alpha}^k = \left(\alpha^k_1, \alpha^k_2, \alpha^k_3\right)^T$ and the matrix

$$\widetilde{\boldsymbol{G}}^k_n = \left(\boldsymbol{g}^{\pi(1),T}_{n,r}, \boldsymbol{g}^{\pi(2),T}_{n,r}, \boldsymbol{g}^{\pi(3),T}_{n,r}\right), \tag{10}$$

where $\boldsymbol{g}^{\pi(i)}_{n,r}$ is the first row vector of the block $\boldsymbol{G}^{\pi(i)}_n$ of matrix $\boldsymbol{G}$. Then,

$$\boldsymbol{d}^{k,T}_{n,r} = \widetilde{\boldsymbol{G}}^k_n\boldsymbol{\alpha}^k. \tag{11}$$

By substituting (11) in (9), we obtain

$$\left\{\begin{array}{lcl} P^{(1)}_{k,r}(q)\widetilde{\boldsymbol{G}}^k_n\boldsymbol{\alpha}^k &=& \hat{\boldsymbol{h}}^{k,(1)}_{n,r}(q) + \boldsymbol{\varepsilon}^{k,(1)}_{n,r}(q)\\ P^{(1)}_{k,l}(q)\widetilde{\boldsymbol{G}}^k_n\boldsymbol{\alpha}^k &=& \hat{\boldsymbol{h}}^{k,(1)}_{n,l}(q) + \boldsymbol{\varepsilon}^{k,(1)}_{n,l}(q)\\ &\vdots& \\ P^{(R)}_{k,r}(q)\widetilde{\boldsymbol{G}}^k_n\boldsymbol{\alpha}^k &=& \hat{\boldsymbol{h}}^{k,(R)}_{n,r}(q) + \boldsymbol{\varepsilon}^{k,(R)}_{n,r}(q)\\ P^{(R)}_{k,l}(q)\widetilde{\boldsymbol{G}}^k_n\boldsymbol{\alpha}^k &=& \hat{\boldsymbol{h}}^{k,(R)}_{n,l}(q) + \boldsymbol{\varepsilon}^{k,(R)}_{n,l}(q) \end{array}\right. \tag{12}$$

where $\widehat{\boldsymbol{h}}^{k,(\ell)}_{n,r}(q) = \left(\widehat{h}^{k,(\ell)}_{n,rr}(q), \widehat{h}^{k,(\ell)}_{n,lr}(q)\right)^T$, $\widehat{\boldsymbol{h}}^{k,(\ell)}_{n,l}(q) = \left(\widehat{h}^{k,(\ell)}_{n,ll}(q), \widehat{h}^{k,(\ell)}_{n,rl}(q)\right)^T$, and $\boldsymbol{\varepsilon}^{k,(\ell)}_{n,r}(q)$ and $\boldsymbol{\varepsilon}^{k,(\ell)}_{n,l}(q)$ are defined similarly.

The directivity estimation reduces to the estimation of the parameters $\boldsymbol{\alpha}$. We estimate these parameters based on a nonlinear least squares error criterion. The optimization problem can be formulated as

$$\begin{array}{ll} \textbf{minimize} & \displaystyle\sum_{\substack{\ell=1,\ldots R\\ q=0,\ldots,Q-1\\ n=1,\ldots,N}} \|\hat{\boldsymbol{h}}^{k,(\ell)}_{n,r}(q) - P^{(\ell)}_{k,r}(q)\widetilde{\boldsymbol{G}}^k_n\boldsymbol{\alpha}\|^2 \\ & \qquad + \|\hat{\boldsymbol{h}}^{k,(\ell)}_{n,l}(q) - P^{(\ell)}_{k,l}(q)\widetilde{\boldsymbol{G}}^k_n\boldsymbol{\alpha}\|^2 \\ \textbf{subject to} & 0 \leq \alpha_i \leq 1, \quad i = 1,2,3 \qquad \textbf{\textit{Problem P}}_0 \\ & \sum_{i=1}^3 \alpha_i = 1 \end{array}$$

with optimization variables $\boldsymbol{\alpha}$ and $P^{(\ell)}_{k,l}$, $\ell = \{1,...,R\}$, $q = \{0,...,Q-1\}$, $n = \{1,...,N\}$.

Problem P_0 does not reduce to linear LSE because of the presence of nuisance parameters $P^{(\ell)}_{k,o}(q)$ and it is in general nonconvex. The following theorem establishes the equivalence of P_0 to a generalized symmetric Eigenvalue Complementarity Problem (EiCP) object of thorough studies in optimization theory (see e.g. [4] and references therein).

Theorem 1. *Problem P_0 is equivalent to the following problem P_1 with optimization variable $\boldsymbol{\alpha}$,*

$$\begin{array}{lll} \textbf{\textit{maximize}} & f_k(\boldsymbol{\alpha}) = \dfrac{\boldsymbol{\alpha}^H\mathfrak{Re}(\boldsymbol{H}^k)\boldsymbol{\alpha}}{\boldsymbol{\alpha}^H\mathfrak{Re}(\boldsymbol{\Gamma}^k)\boldsymbol{\alpha}} & \textbf{\textit{Problem P}}_1 \\ \textbf{\textit{subject to}} & \sum_{i=1}^3 \alpha_i = 1 \quad 0 \leq \alpha_i \leq 1, & i = 1,2,3 \end{array}$$

being $\boldsymbol{H}^k$ and $\boldsymbol{\Gamma}^k$ the 3×3 matrices defined as

$$\boldsymbol{H}^k = \widetilde{\boldsymbol{G}}^{k,H}\left(\sum_{q=0}^{Q-1}\sum_{\ell=1}^{R}\left(\hat{\boldsymbol{h}}^{k,(\ell)}_r(q)\hat{\boldsymbol{h}}^{k,(\ell)H}_r(q) + \hat{\boldsymbol{h}}^{k,(\ell)}_l(q)\hat{\boldsymbol{h}}^{k,(\ell)H}_l(q)\right)\right)\widetilde{\boldsymbol{G}}^k, \tag{13}$$

$$\boldsymbol{\Gamma}^k = \widetilde{\boldsymbol{G}}^{k,H}\widetilde{\boldsymbol{G}}^k \tag{14}$$

with $\hat{\boldsymbol{h}}^{k,(\ell)}_o(q) = \left(\hat{\boldsymbol{h}}^{k,(\ell)H}_{1,o}(q), ..., \hat{\boldsymbol{h}}^{k,(\ell)H}_{N,o}(q)\right)^H$ and $\widetilde{\boldsymbol{G}}^k = \left(\widetilde{\boldsymbol{G}}^{k,H}_1, ..., \widetilde{\boldsymbol{G}}^{k,H}_N\right)^H$.

Due to space constraint the proof of Theorem 1 is omitted here. It can be found in [6].

The optimal vector $\boldsymbol{\alpha}^*$ provides the desired estimation of the parameter $\boldsymbol{\alpha}^k$ and a PLSE of the directivity column block $\boldsymbol{D}^k$ is given by $\widehat{\boldsymbol{D}}^k = \sum_{i=1}^3 \alpha^*_i\boldsymbol{G}^{\pi(i)}$.

Interestingly, Problem P_1 does not require an explicit estimation of the nuisance parameters, i.e. the propagation coefficients, with consequent computational complexity and numerical error propagation reduction. In the rest of this section we discuss the solution of Problem P_1.

Let us observe that $f_k(\boldsymbol{\alpha})$ assumes the same value on each of the points belonging to the same ray passing through the origin, i.e, $f_k(\boldsymbol{\alpha}) = f_k(\rho\boldsymbol{\alpha})$ for any nonzero real ρ. Therefore, given any vector $\boldsymbol{\alpha}^*$ maximizing $f_k(\boldsymbol{\alpha})$, it is straightforward to derive from it a vector that achieves the optimal value $f(\boldsymbol{\alpha}^*)$ and satisfies the constraint $\sum_i \alpha_i = 1$ by setting

$$\boldsymbol{\alpha}_{opt} = \frac{\boldsymbol{\alpha}^*}{\|\boldsymbol{\alpha}^*\|_1}. \tag{15}$$

Based on (15), the constraints $\alpha_i \leq 1$ are also satisfied if $\alpha_i \geq 0$. Thus, the problem is very similar to a generalized eigenvalue problem (see e.g. [5]). However, in general $\boldsymbol{\alpha}$, a solution of the generalized eigenvector problem does not satisfy the constraints $\alpha_i \geq 0$. In the following, we discuss the utilization of the solutions of a generalized eigenvalue problem to find a solution to P_1, which satisfies also the constraints $\alpha_i \geq 0$.

The global maximum of function $f_k(\boldsymbol{\alpha})$ is achieved by the eigenvector corresponding to the maximum generalized eigenvalue of $\Re\mathfrak{e}(\boldsymbol{H}^k)$ and $\Re\mathfrak{e}(\boldsymbol{\Gamma}^k)$. The other generalized eigenvectors of $\Re\mathfrak{e}(\boldsymbol{H}^k)$ and $\Re\mathfrak{e}(\boldsymbol{\Gamma}^k)$ achieve local maxima, local minima or saddle points[4] of the function $f_k(\boldsymbol{\alpha})$. Moreover, $f_k(\boldsymbol{\alpha})$ is a continuous function of $\boldsymbol{\alpha}$. Therefore, if the generalized eigenvector of $\Re\mathfrak{e}(\boldsymbol{H}^k)$ and $\Re\mathfrak{e}(\boldsymbol{\Gamma}^k)$ yielding the global optimum of the unconstrained problem does not have all components of the same sign, i.e. it cannot be normalized to satisfy the constraint $\alpha_i \geq 0$, the solution of P_1 in the nonnegative orthant is achieved or by the other generalized eigenvectors of $\Re\mathfrak{e}(\boldsymbol{H}^k)$ and $\Re\mathfrak{e}(\boldsymbol{\Gamma}^k)$ or falls on the boundary of the nonnegative orthant. Then, we can compute the solution of P_1 by exhaustive search on the boundary and among the generalized eigenvectors. Among the generalized eigenvectors, we need to analyze the ones that have all nonnegative components. The value of $f_k(\boldsymbol{\alpha})$ is given by the generalized eigenvalue corresponding to the generalized eigenvector.

For searching the solution of P_1 on the boundary, we need to consider two different cases: (a) Two elements of $\boldsymbol{\alpha}$ are 0; (b) One element of $\boldsymbol{\alpha}$ is 0. In the former case, the value of $f(\boldsymbol{\alpha})$ can be easily computed by

$$f(\boldsymbol{\alpha}) = \left| \frac{\Re\mathfrak{e}(\boldsymbol{H}^k)_{ii}}{\Re\mathfrak{e}(\boldsymbol{\Gamma}^k)_{ii}} \right| \tag{16}$$

where $\Re\mathfrak{e}(\boldsymbol{H}^k)_{ii}$ and $\Re\mathfrak{e}(\boldsymbol{\Gamma}^k)_{ii}$ denotes the ith diagonal element of $\Re\mathfrak{e}(\boldsymbol{H}^k)$ and $\Re\mathfrak{e}(\boldsymbol{\Gamma}^k)$, respectively.

In the latter case, we examine the maximum value of $f_k(\boldsymbol{\alpha})$ for $\alpha_i = 0, i = 1, 2, 3$ separately. For $\alpha_i = 0$, $\alpha_j > 0$, $i, j = 1, 2, 3$, $i \neq j$, we have

$$\frac{\boldsymbol{\alpha}^{(\backsim i)H} \Re\mathfrak{e}(\boldsymbol{H}^k)^{(\backsim i)} \boldsymbol{\alpha}^{(\backsim i)}}{\boldsymbol{\alpha}^{(\backsim i)H} \Re\mathfrak{e}(\boldsymbol{\Gamma}^k)^{(\backsim i)} \boldsymbol{\alpha}^{(\backsim i)}} \tag{17}$$

and we retain the generalized eigenvectors of $\Re\mathfrak{e}(\boldsymbol{H}^k)^{(\backsim i)}$ and $\Re\mathfrak{e}(\boldsymbol{\Gamma}^k)^{(\backsim i)}$ with components of the same sign.

To summarize, to solve the optimization problem P_1 we analyze all the generalized eigenvectors of $\Re\mathfrak{e}(\boldsymbol{H}^k)$ and $\Re\mathfrak{e}(\boldsymbol{\Gamma}^k)$, the generalized eigenvectors of (17) and the values (16). We compare the values of $f_k(\boldsymbol{\alpha})$ for all the possible cases and choose the maximum one. The corresponding $\boldsymbol{\alpha}^*$ yields the desired estimation.

In order to solve the directivity estimation problem for all the active STs over the full coverage area it is relevant to further observe that (a) Problem P_1 has to be solved for each STs; (b) In the general case, the three nearest points surrounding ST k are not known. Then, an exhaustive search over the whole possible triplets of adjacent reference points is required and the triplet yielding to the least squared error is selected.

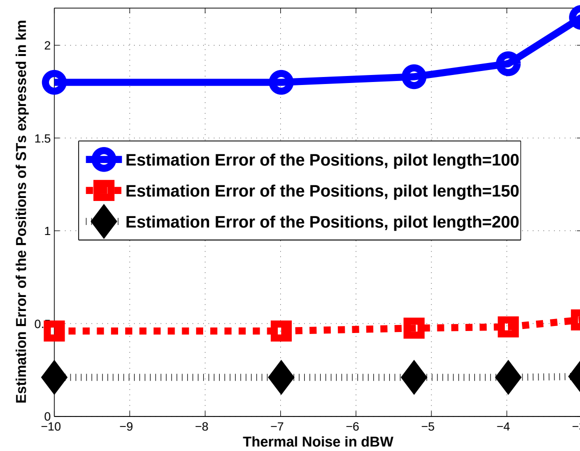

Figure 1. Estimation error of STs positions in km with different levels of thermal noise, $Q = 30$, and $K = 30$

IV. Numerical Performance Assessment

In this section, we analyze the performance of the proposed PLSE algorithm. The simulations are performed for satellite terminals equipped with two antennas, i.e., $R = 2$. The satellite is endowed with 163 SA. For simulations, we utilize the actual directivity vectors of a geostationary system serving the European area. The propagation coefficients are generated according to the Surrey model in [2]. The power of the transmit signals is set to be 0 dBW. The results are obtained by averaging over 100 system realizations, i.e., 100 different groups of STs are randomly generated and the performance of the system is assessed over each realization. The event when the distance between the actual position and the estimated position of a terminal is greater than 40 kilometers is referred to as "estimation failure." The positions of the STs are generated randomly and uniformly in a rectangular region covering the

[4]As well known, the optimization of any Rayleigh quotient $\dfrac{\boldsymbol{x}^T\boldsymbol{A}\boldsymbol{x}}{\boldsymbol{x}^T\boldsymbol{B}\boldsymbol{x}}$, with $\boldsymbol{A}, \boldsymbol{B}$ squared matrices and $\boldsymbol{x}$ vector of consistent dimension, is equivalent to the optimization of $\boldsymbol{x}^T\boldsymbol{A}\boldsymbol{x}$ constrained to $\boldsymbol{x}^T\boldsymbol{B}\boldsymbol{x} = K$. It is straightforward to observe that the gradient of the corresponding Lagrangian vanishes in any $(\lambda, \boldsymbol{v})$, being λ and $\boldsymbol{v}$ respectively a generalized eigenvalue and the corresponding eigenvector of the matrices $\boldsymbol{A}$ and $\boldsymbol{B}$.

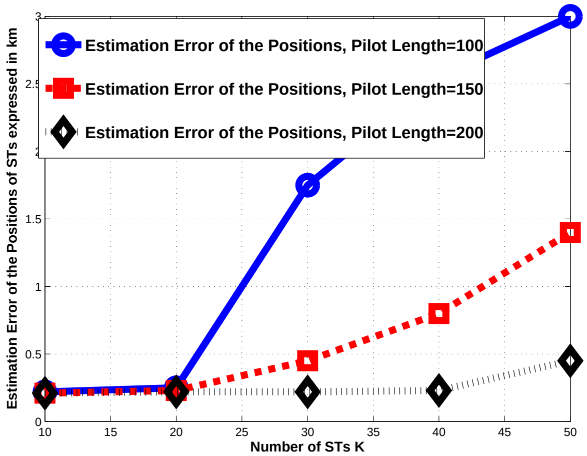

Figure 2. Estimation error of the positions of STs expressed in km with different number of STs, $Q = 30$, and $\sigma_z^2 = -\infty$ dBW.

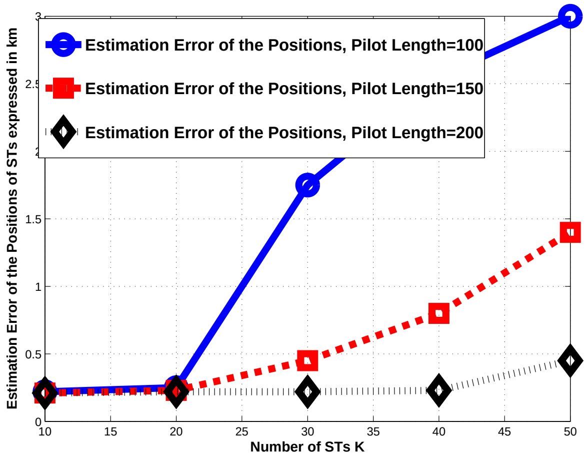

Figure 3. Estimation error of the positions of STs expressed in km with different number of coherence time intervals, $K = 40$, and $\sigma_z^2 = -\infty$ dBW.

most of Europe. The pilot sequences' length is either 100, 150 or 200. They consist of QPSK symbols randomly generated.

Figure 1 shows the estimation error of the PLSE in terms of positions errors[5] for increasing levels of thermal noise. In the system there are 30 active STs. The number of coherence time intervals in the simulation is 30. As apparent in Figure 1, when the thermal noise increases, the estimation error of STs' positions increase only slightly since interference from other STs plays major role.

The impact of the number of active STs in the system on the PLSE estimation is shown in terms of distance estimation error in Figure 2. In this simulation, $Q = 30$, the thermal noise is absent and only co-channel interference is present.

[5]The choice to show the performance in terms of error on the distance instead of the error on the directivity vectors is due to the fact that the average error on the directivity vector is not very representative because of the large length of the directivity vectors (163 elements) and their large range of variation. This choice is adopted throughout all this section.

When the number of STs is greater than 20, the position's estimation error increases rapidly when the length of the pilot is 100, as apparent from Figure 2 . On the contrary, when the length of the training pilot is 200, the estimation error of the positions of STs increases very slowly and the PLSE achieves a good estimation of the positions. It is worth noticing that when K STs are transmitting, the channel consists of $2RK = 4K$ links. The performance starts degrading significantly when the training length approaches or is lower than $4K$.

Finally, we analyze the impact of Q, the number of coherence time intervals on the PLSE estimation. In our simulations, $K = 40$. The thermal noise is absent. Figure 3 shows the impact of the number of coherence time intervals on the estimation errors of the STs' locations. Interestingly, the algorithm's performance are not sensitive to the number of coherence time intervals when the pilot length is greater than 2RK. On the contrary, it has a beneficial impact when the training length is short and does not guarantee good performance.

V. Conclusions

We provided an estimation algorithm of the slow varying component of a satellite channel at the gateway. We propose a nonlinear parametric least squares estimation that can be expressed as a nonconvex constrained optimization problem. We show that the constrained optimization reduces to an eigenvalue complementary problem and does not require estimation of the fast varying channel components. This enables to keep complexity moderate for real time implementation.

A numerical analysis of the performance shows the dependence of the proposed algorithm on thermal noise, number of coherence time intervals used for the estimation, number of users and training length.

Acknowledgment

The authors thank Gael Scot and Marie Robert for thorough discussions and detailed information on the satellite channel modeling. This research work was partially funded by the French Space Agency, Centre National d'Études Spatiales (CNES).

References

[1] P. Arapoglou, K. Liolis, M. Bertinelli, A. Panagopoulos, P. Cottis, R. Gaudenzi *MIMO over Satellite: A Review* IEEE Communication Surveys Tutorials, vol. 13, pp. 27-51, 2011.

[2] P. R. King, *Modelling and Measurement of the Land Mobile Satellite MIMO Radio Propagation Channel*, PhD Thesis, 2007.

[3] S. M. Kay, *Fundamentals of Statistical Signal Processing, Volume 1: Estimation Theory*, Prentice Hall, 1993, vol. 1, Prentice Hall Signal Processing Series.

[4] Marcelo Queiroz and Joaquim Judice and Carlos Humes,*The Symmetric Eigenvalue Complementarity Problem*, Mathematics of Computation, vol. 73, n. 248, pp. 1849–1863, August 2003

[5] Stephen Boyd and Lieven Vandenberghe, *Convex Optimization*, Cambridge University Press, New York, 2007.

[6] Xiao Lei and Laura Cottatellucci, *Technical Report RR-12-267: Parametric Least Squares Estimation for Non-linear Satellite Channels*, `[online] Available: www.eurecom.fr/fr/publication/list/type/report`, June 2012.