EURECOM
Department of Mobile Communication
2229, route des Crêtes
B.P. 193
06904 Sophia-Antipolis
FRANCE

Research Report RR-12-267

Parametric Least Squares Estimation for Nonlinear Satellite Channels

June 11^{th}, 2012
Last update June 11^{th}, 2012

Laura Cottatellucci and Lei Xiao

Tel : (+33) 4 93 00 81 00
Fax : (+33) 4 93 00 82 00
Email : {Laura.Cottatellucci,Lei.Xiao}@eurecom.fr

[1]EURECOM's research is partially supported by its industrial members: BMW Group, Cisco, Monaco Telecom, Orange, SAP, SFR, Sharp, STEricsson, Swisscom, Symantec, Thales.

Parametric Least Squares Estimation for Nonlinear Satellite Channels

Laura Cottatellucci and Lei Xiao

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.

Index Terms

Satellite Communication, slow fading component estimation

Contents

List of Figures

1 Introduction

Multibeam satellite systems are widely employed to provide broadband services to geographical areas under-served by terrestrial infrastructure. In modern MSS, satellites are equipped with multiple antennas and beamforming networks (BFN) which allow for multibeam transmissions. Thanks to the improvements in the switching speed of BFNs, 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 in the design of adaptive beamforming systems and strongly depends on the channel characteristics.

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 [**?**]. We follow the channel model proposed by [**?**] 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 direction of the line of sight (LOS) of the signal transmitted from/to the satellite. The propagation coefficients model the propagation losses (atmospheric and shadowing) between satellite and ST. They are highly variable in time and their coherence time is too short compared to propagation delays. Thus, information fed back to the gateway is stale and the estimation of the instantaneous transfer matrix of the channel cannot be utilized for the design of adaptive beamforming.

When the estimation of the full (or instantaneous) CSI is not feasible as in satellite systems where the propagation delay is very long and the instantaneous fed back CSI measurement becomes obsolete, channel distribution information (CDI) can be utilized instead of the instantaneous CSI. CDI provides a practical solution when the acquisition of the instantaneous CSI is not feasible, since the channel statistics change slower than the channel state.

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 (at least from a statistical point of view). The latter approach presents well known benefits in terms of system spectral efficiency since the feedback channel is not required. In this contribution, we assume that channel reciprocity holds 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 and follow the model in [?], the CDI estimation reduces to the estimation of the slow varying components, i.e, the directivity vectors. 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 the channels where the directivity vector is parametrically represented by a linear combination of given known directivity vectors and the varying propagation coefficients 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.

Because of a one-to-one relationship between the parameters representing a directivity vector and the STs' geographic position, the performance of the simulation is assessed in terms of mismatch between actual and estimated location of the satellite terminals. It is worth noticing that all the simulations are based on measurements of a the directivity coefficients of a real satellite system and the generation of the propagation coefficients follows strictly the Surrey model in [?].

The remainder of this contribution is organized as follows. In Section II, we introduce the satellite system model for the proposed estimation and describe a parametric model for the directivity vectors. In section III, we describe the estimation problem and the proposed algorithm, PLSE. In section IV, the application of the PLSE algorithm to a connection-oriented data transmission channel is discussed. In Section V, the performance of the proposed estimation algorithm is assessed through numerical simulations. Section VI provides conclusive remarks.

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. $\mathfrak{Re}$ denotes the real part operator. $\|\cdot\|_l$ denotes the norm l vector. Shortly, $\|\cdot\|$ denotes the Euclidean norm. $\boldsymbol{A}^{(\backsim i)}$ represents the submatrices of the matrix $\boldsymbol{A}$ obtained by removing the i-th column and the i-th row.

2 System Model

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 the 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].

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 $2N \times 2$, $\boldsymbol{D}^k = (\boldsymbol{D}_1^{kT}, \boldsymbol{D}_2^{kT}, \ldots \boldsymbol{D}_N^{kT})^T$ 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

[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.

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 the following two realistic assumptions: (a) the directivity vectors of some reference STs in a grid are known at the gateway. 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. In general, the time slots are nonconsecutive and such that the corresponding propagation channels can be considered statistically independent. However, these Q time slots are sufficiently close such that the directivity matrix can be considered constant in the whole observation time.

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

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

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 $\boldsymbol{\mathcal{Z}}_{n,o}(q) = (z_{n,o}[s_q], z_{n,o}[s_q+1], \ldots, z_{n,o}[s_q+L-1])$.

3 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. [**?**]), we obtain the LSE estimation of $\boldsymbol{h}_{n,r}(q)$ and $\boldsymbol{h}_{n,l}(q)$ given by

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

and

$$\hat{\boldsymbol{h}}_{n,r}(q) = \boldsymbol{\mathcal{Y}}_{n,r}\boldsymbol{X}^H_q(\boldsymbol{X}_q\boldsymbol{X}^H_q)^{-1}, \tag{9}$$

respectively.

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

$$\begin{cases} d^k_{n,rr} P^{(1)}_{k,r}(q) &= \hat{h}^{k,(1)}_{n,rr}(q) + \varepsilon^{k,(1)}_{n,rr}(q) \\ d^k_{n,rl} P^{(1)}_{k,r}(q) &= \hat{h}^{k,(1)}_{n,lr}(q) + \varepsilon^{k,(1)}_{n,lr}(q) \\ d^k_{n,rr} P^{(1)}_{k,l}(q) &= \hat{h}^{k,(1)}_{n,ll}(q) + \varepsilon^{k,(1)}_{n,ll}(q) \\ d^k_{n,rl} P^{(1)}_{k,l}(q) &= \hat{h}^{k,(1)}_{n,rl}(q) + \varepsilon^{k,(1)}_{n,rl}(q) \\ &\vdots \\ d^k_{n,rr} P^{(R)}_{k,l}(q) &= \hat{h}^{k,(R)}_{n,ll}(q) + \varepsilon^{k,(R)}_{n,ll}(q) \\ d^k_{n,rl} P^{(R)}_{k,l}(q) &= \hat{h}^{k,(R)}_{n,rl}(q) + \varepsilon^{k,(R)}_{n,rl}(q). \end{cases} \tag{10}$$

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 (10) 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

$$\widetilde{\boldsymbol{G}}^k_n = \left(\boldsymbol{g}^{\pi(1),T}_{n,r}, \boldsymbol{g}^{\pi(2),T}_{n,r}, \boldsymbol{g}^{\pi(3),T}_{n,r}\right) \tag{11}$$

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,

$$\boldsymbol{d}^{k,T}_{n,r} = \widetilde{\boldsymbol{G}}^k_n \boldsymbol{\alpha}^k. \tag{12}$$

By substituting (12) in (10), we obtain

$$\begin{cases} P^{(1)}_{k,r}(q) \widetilde{\boldsymbol{G}}^k_n \boldsymbol{\alpha}^k &= \hat{\boldsymbol{h}}^{k,(1)}_{n,r}(q) + \boldsymbol{\varepsilon}^{k,(1)}_{n,r}(q) \\ P^{(1)}_{k,l}(q) \widetilde{\boldsymbol{G}}^k_n \boldsymbol{\alpha}^k &= \hat{\boldsymbol{h}}^{k,(1)}_{n,l}(q) + \boldsymbol{\varepsilon}^{k,(1)}_{n,l}(q) \\ &\vdots \\ P^{(R)}_{k,r}(q) \widetilde{\boldsymbol{G}}^k_n \boldsymbol{\alpha}^k &= \hat{\boldsymbol{h}}^{k,(R)}_{n,r}(q) + \boldsymbol{\varepsilon}^{k,(R)}_{n,r}(q) \\ P^{(R)}_{k,l}(q) \widetilde{\boldsymbol{G}}^k_n \boldsymbol{\alpha}^k &= \hat{\boldsymbol{h}}^{k,(R)}_{n,l}(q) + \boldsymbol{\varepsilon}^{k,(R)}_{n,l}(q) \end{cases} \tag{13}$$

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

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

Problem P_0 does not reduce to linear LSE because of the presence of nuisance parameters $P_{k,o}^{(\ell)}(q)$ 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. [**?**] and references therein).

Theorem 1. *The optimization problem P_0 is equivalent to the problem*

$$\begin{array}{lll} \textbf{\textit{maximize}} & f_k(\boldsymbol{\alpha}) = \dfrac{\boldsymbol{\alpha}^H \mathfrak{Re}(\boldsymbol{H}^k)\boldsymbol{\alpha}}{\boldsymbol{\alpha}^H \mathfrak{Re}(\boldsymbol{\Gamma}^k)\boldsymbol{\alpha}} & \textit{Problem } P_1 \\ \textbf{\textit{subject to}} & \sum_{i=1}^3 \alpha_i = 1 \quad 0 \le \alpha_i \le 1, \qquad i = 1,2,3 & \end{array}$$

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

$$\boldsymbol{H}^k = \widetilde{\boldsymbol{G}}^{k,H} \left(\sum_{q=0}^{Q-1} \sum_{\ell=1}^{R} \left(\hat{\boldsymbol{h}}_r^{k,(\ell)}(q) \hat{\boldsymbol{h}}_r^{k,(\ell)H}(q) + \hat{\boldsymbol{h}}_l^{k,(\ell)}(q) \hat{\boldsymbol{h}}_l^{k,(\ell)H}(q) \right) \right) \widetilde{\boldsymbol{G}}^k, \quad (14)$$

$$\boldsymbol{\Gamma}^k = \widetilde{\boldsymbol{G}}^{k,H} \widetilde{\boldsymbol{G}}^k \quad (15)$$

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

Theorem 1 is proven in Appendix.

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

$$\boldsymbol{\alpha}_{opt} = \frac{\boldsymbol{\alpha}^*}{\|\boldsymbol{\alpha}^*\|_1}. \quad (16)$$

Based on (16), the constraints $\alpha_i \le 1$ are also satisfied if $\alpha_i \ge 0$. Thus, the problem is very similar to a generalized eigenvalue problem (see e.g. [**?**]). However, in general $\boldsymbol{\alpha}$, a solution of the generalized eigenvector problem does not satisfy the constraints $\alpha_i \ge 0$. 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 $\alpha_i \ge 0$.

The global maximum of function $f_k(\boldsymbol{\alpha})$ is achieved by the eigenvector corresponding to the maximum generalized eigenvalue of $\mathfrak{Re}(\boldsymbol{H}^k)$ and $\mathfrak{Re}(\boldsymbol{\Gamma}^k)$. The other generalized eigenvectors of $\mathfrak{Re}(\boldsymbol{H}^k)$ and $\mathfrak{Re}(\boldsymbol{\Gamma}^k)$ achieve local maxima, local minima or saddle points[3] 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 $\mathfrak{Re}(\boldsymbol{H}^k)$ and $\mathfrak{Re}(\boldsymbol{\Gamma}^k)$ 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 $\mathfrak{Re}(\boldsymbol{H}^k)$ and $\mathfrak{Re}(\boldsymbol{\Gamma}^k)$ 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

$$f(\boldsymbol{\alpha}) = \left| \frac{\mathfrak{Re}(\boldsymbol{H}^k)_{ii}}{\mathfrak{Re}(\boldsymbol{\Gamma}^k)_{ii}} \right| \tag{17}$$

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

$$\frac{\boldsymbol{\alpha}^{(\backsim i)H} \mathfrak{Re}(\boldsymbol{H}^k)^{(\backsim i)} \boldsymbol{\alpha}^{(\backsim i)}}{\boldsymbol{\alpha}^{(\backsim i)H} \mathfrak{Re}(\boldsymbol{\Gamma}^k)^{(\backsim i)} \boldsymbol{\alpha}^{(\backsim i)}} \tag{18}$$

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

To summarize, to solve the optimization problem P_1 we analyze all the generalized eigenvectors of $\mathfrak{Re}(\boldsymbol{H}^k)$ and $\mathfrak{Re}(\boldsymbol{\Gamma}^k)$, the generalized eigenvectors of (18) and the values (17). 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. In a practical system for connection oriented communications,

[3] As well known, the optimization of any Rayleigh quotient $\frac{\boldsymbol{x}^T \boldsymbol{A} \boldsymbol{x}}{\boldsymbol{x}^T \boldsymbol{B} \boldsymbol{x}}$, 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 $\boldsymbol{B}$.

such exhaustive search is not required also in case of ST's mobility and the search can be limited to triples adjacent to the area covered by the triplet utilized in the previous estimation.

4 PLSE for Connection-Oriented Channels

In this section, we discuss the application of the PLSE algorithm to a practical system with connection-oriented channels. First, we illustrate the peculiarities of this system, then summarize the steps to apply the proposed PLSE algorithm.

In the connection-oriented channel, the gateway has the prior knowledge of the number of active STs and pilot sequences utilized by each ST. Moreover, the gateway is aware of the area where each ST is located from the previous estimation up to some estimation error and mismatches due to the ST's mobility. This gateway may update the estimation of the directivity coefficients based on the previous estimation. It searches the k-th ST in a disc $\mathcal{C}_k(\hat{S}_k, \mathfrak{R}_k)$, where $\hat{S}_k$, the estimated position of ST k at the previous step, is the center of the disc, and $\mathfrak{R}_k$ is the radius. We denote the distance between the actual position of ST k and the center of the disc $\mathcal{C}_k$ as $D_{a,k}$.

The algorithm for connection-oriented channels and based on the PLSE estimation is summarized in Algorithm 1.

for $k = 1, ..., K$ **do**
    for $q = 1, ..., Q$ **do**
        Calculate $\hat{\boldsymbol{h}}^{k}_{n,o}(q)$, $\{o\} = \{r, l\}$ according to (8) and (9).
    end
end
for $k = 1, ..., K$ **do**
    Find all the Π_k adjacent triplets located in $\mathcal{C}_k(\hat{S}_k, \mathfrak{R}_k)$
    for $i = 1, ...\Pi_k$ **do**
        Compute the optimal parametric coefficients $\boldsymbol{\alpha}$ by solving the optimization problem P_1;
    end
    select triplet and parameter $\boldsymbol{\alpha}$ yielding to the minimum least squares.
    Determine position and directivity vector of the ST k corresponding to the optimum triplet and optimum $\boldsymbol{\alpha}$ by applying (4).
end

Algorithm 1: Directivity coefficients estimation based on PLSE for connection-oriented channels.

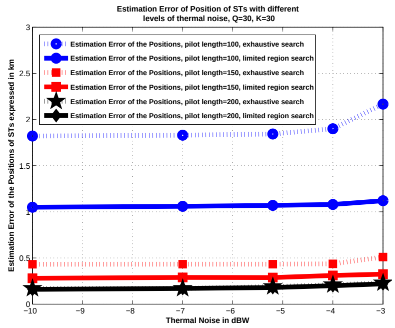

Figure 1: Estimation error of STs' positions in km versus thermal noise

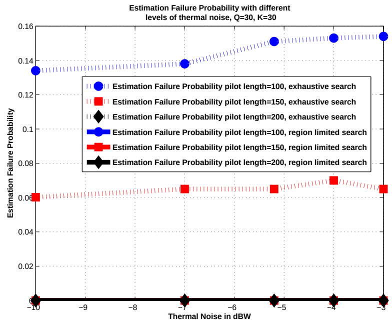

Figure 2: Estimation failure probability versus thermal noise

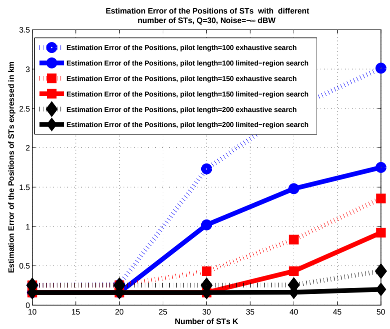

Figure 3: Estimation error of the STs' positions in km versus a varying number of STs

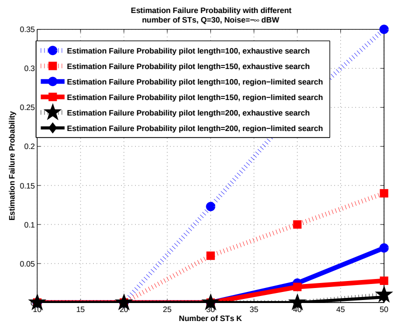

Figure 4: Estimation failure probability versus number of STs

5 Numerical Performance Assessment

In this section, we analyze the performance of the proposed algorithm under two different assumptions, namely, the gateway does not make use of the prior knowledge of the area where ST k is located and a whole map exhaustive search is performed or it does utilize the previous estimation and a limited-region search is performed.

The simulations are performed for satellite terminals equipped with two antennas,i.e., $R = 2$. The satellite is endowed with 163 SA. For the 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 [**?**]. 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 most of Europe. Throughout the whole section, if not differently specified, the following setting is adopted: (1) The pilot sequence length is either 100, 150 or 200 QPSK symbols; (2) In the case of limited-region search, the gateway searches the STts in a circle with a radius of 160 kilometers and the center of the circle is the previous estimated position; (3) In order to initialize the algorithm, the distance D_a between the center of the circle $\hat{S}_k$ and the actual position of ST k is normally distributed with zero mean and variance of 20 kilometers.

Figure 1 shows the estimation error of the PLSE in terms of position errors[4] 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 the STs' positions increase only slightly under both assumptions since interference from other STs plays major role. In general, the algorithm depending on prior knowledge (limited-region search) outperforms the algorithm based on the whole map exhaustive search. The performance gap between the two implementations decreases when the pilot sequences length increases. A similar trend appears in Figure 2 where the estimation failure probability of the PLSE for increasing levels of thermal noise is presented. The exhaustive search approach has an estimation failure probability above 10% for pilot length of 100 symbols, while the estimation failure events vanishes for pilot length of 200 symbols. Figure 1 and Figure 2 suggest that the impact of the thermal noise in the performance of PLSE algorithm is not significant. To achieve better estimation, it is more critical to increase the length of pilots.

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 3 and in terms of estimation failure in Figure 4. In this simulation, $Q = 30$, the thermal noise is absent and only co-channel interference is present. When the number of STs is greater than 20, the position's estimation error increases rapidly when the length of the pilot is 100 in both types of estimation, as apparent from Figure 3. On the contrary, when the length of the training pilot is 200, the estimation error of the STs' positions increases very slowly and the PLSE achieves a good estimation. A similar trend is shown in Figure 4 for the estimation failure probability. It is worth noticing that when K STs are transmitting, the channel consists of $2RK = 4K$ links and the performance starts degrading significantly when the training length approaches or is lower than $4K$.

We also analyze the impact of Q, the number of coherence time intervals, on the PLSE estimation of the STs' positions. In our simulations, $K = 40$. The thermal noise is absent. Figure 5 and Figure 6 shows the impact of the number of coherence

[4] 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 to the large length of the directivity vectors (163 elements) and their large range of variation. This choice is adopted throughout all this section.

time intervals on the estimation errors of the STs' locations and the estimation failure probability, respectively. Also in this case, the algorithm based on the prior information outperforms the exhaustive search. Interestingly, the performance is not sensitive to the number of coherence time intervals Q when the pilot length is greater than $2RK$. On the contrary, an increase of Q has a beneficial impact when the training length is short and does not support good performance.

In the following, we study the impact of the distance between adjacent STs for the limited-region search approach. We generate the STs' positions on a square grid. Each ST has the same distance to its adjacent ST. In the simulation, the thermal noise is absent and the length of the pilot sequence is 200. The number of coherence time intervals is 30. The gateway utilizes the information from the previous estimation. The gateway searches the ST in a disc of radius 240 kilometers. The distance between the actual position of the STs and the center of the disc is normally distributed and has a variance of 40 kilometers.

Figure 7 shows the impact of the distance between adjacent STs. As expected, when the distance between adjacent STs increases, the PLSE algorithm achieves better performance. Figure 7 also indicates that, for a given distance between adjacent STs, the estimation errors of the STs' positions also increases slightly as the number of STs increases.

Finally, we study the impact of the radius $\mathfrak{R}_k$ of the search disc and $\mathrm{D}_{a,k}$, the distance between the actual position of a ST and the center of its search disc. In our simulations, $K = 40$, the thermal noise is absent and the pilot length is 200. The number of coherence time intervals is 30. The positions of the STs are randomly and uniformly generated. Additionally, we force the minimal distance between two adjacent STs to be not less than 40 kilometers.

Figure 8 shows the estimation error when different radius lengths are adopted. When the variance of the distance between the center of the search disc and the actual position is 96 kilometers, the position estimation error is more than 2 kilometers if the search is executed only in a disc of radius 80 kilometers. As the radius of the search disc increases, the position estimation errors decreases dramatically: it decreases approximately 0.3 kilometers for a radius of 144 kilometers. When the variance of the distance between the center point of the search area and actual position is 24 kilometers, a search of a ST in a disc of radius 80 kilometers is sufficient, i.e., as the radius increases, the performance of the PLSE algorithm does not improve further.

6 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. Interestingly, the proposed PLSE algorithm does not require estimation of the fast varying channel components which play the role of nuisance parameters in the problem. This enables to keep complexity moderate for real time implementation.

The performance is assessed by numerical simulations based on realistic satellite channels.

Appendices

A Proof of Theorem 1

By utilizing the definition of $\hat{\boldsymbol{h}}_o^{k,(\ell)}$ and $\widetilde{\boldsymbol{G}}^k$ in Theorem 1, the objective function of problem P_0 can be rewritten as

$$\begin{aligned} f(\boldsymbol{\alpha}, \boldsymbol{P}^k(q)) &= \sum_{\ell=1}^{R}\sum_{q=0}^{Q-1} \|\hat{\boldsymbol{h}}_r^{k,(\ell)}(q) - P_{k,r}^{(\ell)}(q)\widetilde{\boldsymbol{G}}^k\boldsymbol{\alpha}\|^2 + \|\hat{\boldsymbol{h}}_l^{k,(\ell)}(q) - P_{k,l}^{(\ell)}(q)\widetilde{\boldsymbol{G}}^k\boldsymbol{\alpha}\|^2 \\ &= \sum_{\ell=1}^{R}\sum_{q=0}^{Q-1} |P_{k,r}^{(\ell)}|^2\boldsymbol{\alpha}^H\widetilde{\boldsymbol{G}}^{k,H}\widetilde{\boldsymbol{G}}^k\boldsymbol{\alpha} - \mathfrak{Re}\left(P_{k,r}^{(\ell)}\hat{\boldsymbol{h}}_r^{k,(\ell)H}(q)\widetilde{\boldsymbol{G}}^k\boldsymbol{\alpha}\right) + \hat{\boldsymbol{h}}_r^{k,(\ell)H}(q)\hat{\boldsymbol{h}}_r^{k,(\ell)}(q) \\ &+ \sum_{\ell=1}^{R}\sum_{q=0}^{Q-1} |P_{k,l}^{(\ell)}|^2\boldsymbol{\alpha}^H\widetilde{\boldsymbol{G}}^{k,H}\widetilde{\boldsymbol{G}}^k\boldsymbol{\alpha} - \mathfrak{Re}\left(P_{k,l}^{(\ell)}\hat{\boldsymbol{h}}_l^{k,(\ell)H}(q)\widetilde{\boldsymbol{G}}^k\boldsymbol{\alpha}\right) + \hat{\boldsymbol{h}}_l^{k,(\ell)H}(q)\hat{\boldsymbol{h}}_l^{k,(\ell)}(q) \end{aligned} \tag{19}$$

The application of rules for complex gradient operators (see e.g. [**?**, **?**]) yields

$$\frac{\partial f(\boldsymbol{\alpha}, \boldsymbol{P}^k(q))}{\partial P_{k,o}^{(\ell)}(q)} = 2P_{k,o}^{(\ell)}(q)\boldsymbol{\alpha}^H\widetilde{\boldsymbol{G}}^{k,H}\widetilde{\boldsymbol{G}}^k\boldsymbol{\alpha} - 2\boldsymbol{\alpha}^H\widetilde{\boldsymbol{G}}^{k,H}\hat{\boldsymbol{h}}_o^{k,(\ell)}(q). \tag{20}$$

Since $f(\boldsymbol{\alpha}, \boldsymbol{P}^k(q))$ is convex in $P_{k,o}^{(\ell)}(q)$, for any given $\boldsymbol{\alpha}$, it is minimized by the value of $P_{k,o}^{(\ell)}(q)$ where (20) vanishes, i.e.

$$P_{k,o}^{(t)}(q) = \frac{\boldsymbol{\alpha}^H\widetilde{\boldsymbol{G}}^{k,H}\hat{\boldsymbol{h}}_o^{k,(\ell)}(q)}{\boldsymbol{\alpha}^H\widetilde{\boldsymbol{G}}^{k,H}\widetilde{\boldsymbol{G}}^k\boldsymbol{\alpha}}, \quad o = r, l. \tag{21}$$

By substituting (21) in $f(\boldsymbol{\alpha}, \boldsymbol{P}^k(q))$ and neglecting the constant terms, we obtain the optimization problem

$$\begin{aligned} &\textbf{minimize} \quad -\sum_{\ell=1}^{R}\sum_{q=0}^{Q-1}\left(\frac{\boldsymbol{\alpha}^H\widetilde{\boldsymbol{G}}^{k,H}\hat{\boldsymbol{h}}_r^{k,(\ell)}(q)\hat{\boldsymbol{h}}_r^{k,(\ell)H}(q)\widetilde{\boldsymbol{G}}^k\boldsymbol{\alpha}}{\boldsymbol{\alpha}^H\widetilde{\boldsymbol{G}}^{k,H}\widetilde{\boldsymbol{G}}^k\boldsymbol{\alpha}} + \frac{\boldsymbol{\alpha}^H\widetilde{\boldsymbol{G}}^{k,H}\hat{\boldsymbol{h}}_l^{k,(\ell)}(q)\hat{\boldsymbol{h}}_l^{k,(\ell)H}(q)\widetilde{\boldsymbol{G}}^k\boldsymbol{\alpha}}{\boldsymbol{\alpha}^H\widetilde{\boldsymbol{G}}^{k,H}\widetilde{\boldsymbol{G}}^k\boldsymbol{\alpha}}\right) \\ &\textbf{subject to} \quad \sum_{i=1}^{3}\alpha_i = 1 \quad 0 \le \alpha_i \le 1, \qquad i = 1,2,3 \end{aligned}$$

which is equivalent to

$$\begin{array}{ll} \textbf{maximize} & f_k(\boldsymbol{\alpha}) = \dfrac{\boldsymbol{\alpha}^H \boldsymbol{H}^k \boldsymbol{\alpha}}{\boldsymbol{\alpha}^H \boldsymbol{\Gamma}^k \boldsymbol{\alpha}} \\ \textbf{subject to} & \sum_{i=1}^{3} \alpha_i = 1 \quad 0 \leq \alpha_i \leq 1, \qquad i = 1,2,3. \end{array}$$

By observing that $\boldsymbol{H}^k$ and $\boldsymbol{\Gamma}^k$ are Hermitian, $\boldsymbol{\alpha}$ is a vector of reals, and the quadratic forms are real, the equalities $\boldsymbol{\alpha}^H \boldsymbol{H}^k \boldsymbol{\alpha} = \boldsymbol{\alpha}^H \mathfrak{Re}(\boldsymbol{H}^k)\boldsymbol{\alpha}$ and $\boldsymbol{\alpha}^H \boldsymbol{\Gamma}^k \boldsymbol{\alpha} = \boldsymbol{\alpha}^H \mathfrak{Re}(\boldsymbol{H}^k)\boldsymbol{\alpha}$ hold. This concludes the proof of Theorem 1.

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).

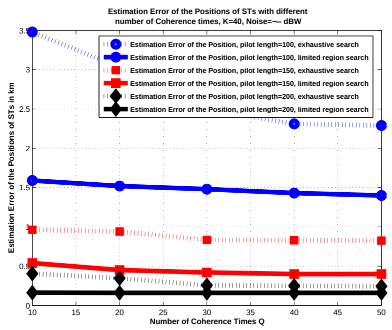

Figure 5: Estimation error of the STs' positions in km versus number of coherence time intervals

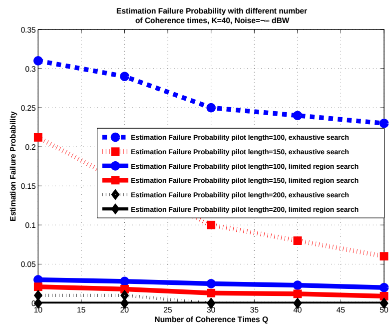

Figure 6: Estimation failure probability versus number of coherence time intervals

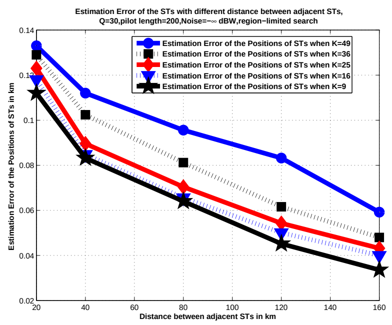

Figure 7: Estimation error of the STs' positions in km versus distance between adjacent STs

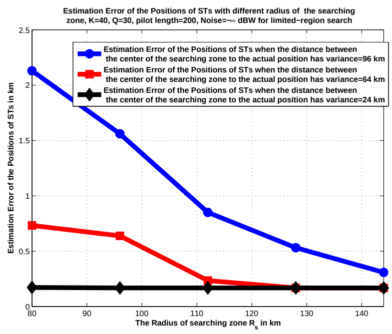

Figure 8: Estimation error of the STs' positions in km versus radius of the search disc.