IDENTIFIABILITY AND PERFORMANCE CONCERNS FOR LOCATION ESTIMATION

Konstantinos Papakonstantinou and Dirk Slock

Mobile Communications Department, Eurecom
Sophia Antipolis, France
Email: papakons@eurecom.fr, slock@eurecom.fr

ABSTRACT

Common localization approaches require a large amount of information to be available in order to achieve identifiability, when the signal propagates in a strictly Non-Line-of-Sight (NLOS) environment. Furthermore, even if they achieve identifiability, they usually perform poorly. In this contribution we investigate the conditions that must be met for identifiability to be feasible and for performance to be adequate. Through basic theorems and simple numerical examples, we study the benefit of exploiting additional information, if such is available and we provide intuitive conclusions that can be proved useful for any localization scheme.

***Index Terms*—** Localization, Positioning, Identifiability, Non Line of Sight , Single Bounce Model, Maximum Likelihood Estimation

1. INTRODUCTION

Geometrical network-based localization techniques are usually implemented in a 2-step procedure: First a set of location-dependent parameters (LDP) is estimated. Common LDP are the Angle of Arrival (AOA), the Angle of Departure (AOD) and the Time of Arrival (TOA) which is equivalent to the delay of any signal path, whether it is the direct Line-of-Sight (LOS) one or any of the NLOS which are available in a multipath environment. Afterward, based on the estimated values of the LDP, a maximum a posteriori (MAP), if information about the priors of the LDP is available, otherwise a maximum likelihood (ML) location estimation can be performed. The latter is the case considered in this paper.

While for the LOS path the geometric relation between any of the above LDP and the coordinates of the Mobile Terminal (MT) is simple, for the NLOS paths such a relation can be derived only with the aid of an appropriate channel model. Thus, in order to exploit the information contained in the LDP of the NLOS paths, we consider the Single Bounce Model (SBM) whose wide applicability stems from the fact in a wireless propagation environment, more bounces imply a larger attenuation of the signal. To further include scenarios in which the MT is moving, we have integrated the SBM with a simple but realistic mobility model. The derived model that can describe dynamic along with static channels, has great advantages compared to the static one, both in terms of identifiability and performance. These advantages are partly due to the fact that a new source of information for the location, the doppler shift, becomes available.

Identifiability in location estimation is the ability to identify (i.e. to estimate) the unique true position of the MT from the pdf of the data (which in our case is the LDP) conditioned on the parameters that need to be estimated. Local identifiability essentially means that position can be identified in an open neighborhood around the true position. In this work we examine local identifiability of the location and the speed and the performance of ML estimation for various cases, using the Fisher Infromation Matrix (FIM).

Eurecom's research is partially supported by its industrial members: BMW Group Research & Technology, Bouygues Telecom, Cisco, Hitachi, ORANGE, SFR, Sharp, STMicroelectronics, Swisscom, Thales. The work presented in this paper has also been partially supported by the European FP7 projects Where and Newcom++ and by the French ANR project Semafor.

2. CHANNEL MODEL

In the following analysis, we consider the Single Bounce Model (SBM) [1]. The SBM is a realistic model for the first few arriving signal components that have non-negligible energy and provides simple geometrical relations between location-dependent parameters and the coordinates of the MT and the scatterers, for static environments. For dynamic environments, i.e. when the MT is moving, a mobility model can be integrated with the SBM. In our work we consider such an environment, as shown in fig. 1 and we assume linear movement of the MT, so that

$$x_i = x_0 + \upsilon_x dt_{i0} \ , \ y_i = y_0 + \upsilon_y dt_{i0} \ , \ dt_{i0} = t_i - t_0 \qquad (1)$$

Let ϕ, ψ, d, f_d denote any AOA, AOD and length of a NLOS path and the doppler shift of the corresponding signal component. With respect to figure 1 and using subscript ij for the parameters at time instant t_i, $0 \leq i < N_t$ and corresponding to path (or scatterer) j, $1 \leq j \leq N_s$, the location-dependent parameters, are given by:

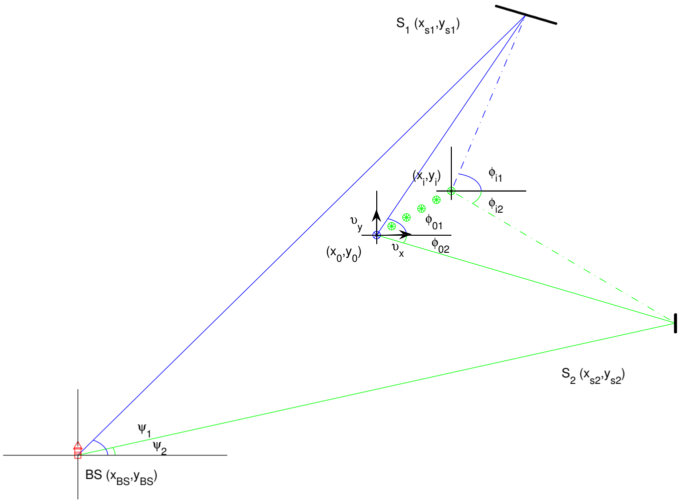

Fig. 1. Single Bounce model

$$\phi_{ij} = \begin{cases} \tan^{-1} \frac{y_{sj}-(y_0+\upsilon_y dt_{i0})}{x_{sj}-(x_0+\upsilon_x dt_{i0})} & , \frac{y_{sj}-(y_0+\upsilon_y dt_{i0})}{x_{sj}-(x_0+\upsilon_x dt_{i0})} > 0 \\ \pi + \tan^{-1} \frac{y_{sj}-(y_0+\upsilon_y dt_{i0})}{x_{sj}-(x_0+\upsilon_x dt_{i0})} & , \frac{y_{sj}-(y_0+\upsilon_y dt_{i0})}{x_{sj}-(x_0+\upsilon_x dt_{i0})} < 0 \end{cases} \tag{2}$$

$$\psi_{ij} = \psi_j = \begin{cases} \tan^{-1} \frac{y_{sj}-y_{BS}}{x_{sj}-x_{BS}} & , \frac{y_{sj}-y_{BS}}{x_{sj}-x_{BS}} > 0 \\ \pi + \tan^{-1} \frac{y_{sj}-y_{BS}}{x_{sj}-x_{BS}} & , \frac{y_{sj}-y_{BS}}{x_{sj}-x_{BS}} < 0 \end{cases} \tag{3}$$

$$d_{ij} = \sqrt{(y_{sj}-(y_0+\upsilon_y dt_{i0}))^2 + (x_{sj}-(x_0+\upsilon_x dt_{i0}))^2} + \sqrt{(y_{sj}-y_{BS})^2 + (x_{sj}-x_{BS})^2} \tag{4}$$

$$f_{d,ij} = \frac{f_c}{c} \frac{\upsilon_x(x_{sj}-(x_0+\upsilon_x dt_{i0})) + \upsilon_y(y_{sj}-(y_0+\upsilon_y dt_{i0}))}{\sqrt{(y_{sj}-(y_0+\upsilon_y dt_{i0}))^2 + (x_{sj}-(x_0+\upsilon_x dt_{i0}))^2}} \tag{5}$$

3. LOCATION ESTIMATION

Location estimation is equivalent to estimating the MT's coordinates at a reference time 0, namely x_0 and y_0. If the MT is moving, its speed components υ_x and υ_y, should be jointly estimated. These two pairs of parameters (parameters of interest) compose a vector which we denote as $\mathbf{p}_{int} = [x_0, y_0, \upsilon_x, \upsilon_y]^t$. The rest of the unknown parameters, which are the coordinates of the scatterers are just nuisance parameters and they compose the vector $\mathbf{p}_{nuis} = [x_{s1}, y_{s1}, \ldots, x_{sN_s}, y_{sN_s}]^t$. The set of all of the above $N_{\mathbf{p}} = 2N_s + 4$ parameters compose the vector:

$$\mathbf{p} = [\mathbf{p}_{int}^t, \mathbf{p}_{nuis}^t]^t \tag{6}$$

Let $\boldsymbol{\theta}$ denote the vector containing all the true values of the LDP. Its entries can be expressed as functions of the entries of $\mathbf{p}$ through eq. (2-5). The location ML estimation will be based on the estimated value $\widehat{\boldsymbol{\theta}}$ of this vector. The size of $\boldsymbol{\theta}$ is $N_{\boldsymbol{\theta}} = (\alpha N_t + \beta)N_s$, with $\alpha \in \{0,1,2,3\}$ being the number of available time-varying LDP and $\beta \in \{0,1\}$ being the number of the non-time-varying ones (only AOD in our case). For example, for $\alpha = 3$ and $\beta = 1$, all possible LDP are available and θ is

$$\boldsymbol{\theta}_{all} = [\boldsymbol{\phi}^t, \boldsymbol{\psi}^t, \mathbf{d}^t, \mathbf{f}_d^t]^t \tag{7}$$

The ML estimate $\widehat{\mathbf{p}}$ is given by:

$$\widehat{\mathbf{p}} = \underset{\mathbf{p}}{\operatorname{argmin}}\{\mathcal{L}\} \tag{8}$$

Assuming that all the estimated LDP are Gaussian-distributed, the likelihood $\mathcal{L}$ is [2]

$$\mathcal{L} \triangleq \mathcal{L}(\boldsymbol{\theta}(\mathbf{p})) = \frac{1}{2}(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}(\mathbf{p}))^t \mathbf{C}_{\widetilde{\boldsymbol{\theta}}}^{-1}(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}(\mathbf{p})) \tag{9}$$

4. CRAMER-RAO BOUND

The correlation matrix of the parameter estimation errors $\widetilde{\mathbf{p}}$ is bounded below by the inverse of the Fisher Information Matrix (FIM)

$$R_{\widetilde{\mathbf{p}}\widetilde{\mathbf{p}}} = E\{(\widehat{\mathbf{p}} - \mathbf{p})(\widehat{\mathbf{p}} - \mathbf{p})^t\} \geq \mathbf{J}^{-1} \tag{10}$$

where the FIM is given by:

$$\boldsymbol{J} = E\left\{ \left(\frac{\partial \mathcal{L}}{\partial \mathbf{p}}\right) \left(\frac{\partial \mathcal{L}}{\partial \mathbf{p}}\right)^t \right\} = \frac{\partial \boldsymbol{\theta}^t}{\partial \mathbf{p}} \mathbf{C}_{\widetilde{\boldsymbol{\theta}}}^{-1} \frac{\partial \boldsymbol{\theta}}{\partial \mathbf{p}^t} = \mathbf{G}\mathbf{C}_{\widetilde{\boldsymbol{\theta}}}^{-1}\mathbf{G}^t \tag{11}$$

and we have introduced the transformation matrix $\mathbf{G} = \frac{\partial \boldsymbol{\theta}^t}{\partial \mathbf{p}}$.

5. IDENTIFIABILITY CONCERNS

In general, local identifiability of a parameter vector $\mathbf{p}$ can be achieved when the FIM is nonsingular [3, Theorem 1]. The following corollary is an immediate consequence of the above theorem and the asymptotic normality of the ML estimates which leads to eq. (11) for the cases where the estimation of $\mathbf{p}$ is based on a previously estimated vector $\boldsymbol{\theta}$.

Corollary 1: Based on eq.(11), local identifiability of a parameter vector $\mathbf{p}$ can be achieved when the transformation matrix $\mathbf{G}$ is square or wide ($N_{\boldsymbol{\theta}} \geq N_{\mathbf{p}}$) and has full rank $N_{\mathbf{p}}$.

Assuming that the first condition is always met, since it essentially means that the number of equations is greater than or equal to the number of unknowns, we will investigate identifiability in terms of the rank of the transformation matrix. In table 1 we give the rank of $\mathbf{G}$ for environments with just 1 resolvable path and different subsets of the LDP being available. These results were obtained analytically (using Gaussian elimination) and verified by simulations. From the table we can observe that $\mathbf{G}$ can be full rank even when only 2 or 3 different kind of LDP are available. When $N_s \geq 2$, $\mathbf{G}$ and thus also the FIM $\mathbf{J}$, are full rank for all cases given in table 1. Therefore it becomes apparent that the entries of $\mathbf{p}$ are identifiable and localization can be implemented in strictly NLOS scenarios, even when the AOA is not available. Thus the integration of a mobility model with the SBM offers a huge advantage in localization, since in the static model, AOA, AOD and path length (delay) of at least 2 resolvable paths are required for identifying the coordinates.

6. PERFORMANCE CONCERNS

It should be noted that in some of the cases examined, although the FIM is invertible, it is ill-conditioned. This has a huge impact on the performance of the ML estimation, since the CRB becomes very large. To avoid scenarios like these, in which the parameters are "ill-identified" since the estimation error will be so large that their estimated value is meaningless, we will compare the condition number of the FIM for different cases and study its impact on the CRB through a numerical example. In this example we consider a pico-cell so that the distance between the MT and the BS is a few tens of meters ($[x_{BS}, y_{BS}] = [0, 0]$, $[x_0, y_0] = [30, 20]$) while the MT is moving with average walking speed ($[\upsilon_x, \upsilon_y] = [2, -1.5]$). The scatterers' coordinates are drawn from a uniform distribution with support region $[x_{BS}, x_{BS} + 2x_0] \times [y_{BS}, y_{BS} + 2y_0]$. The results are averaged for 10^3 different samples of scatterers' coordinates. The condition number of the FIM $\boldsymbol{J}$ is defined as:

$$c_{\boldsymbol{\theta}} = \frac{\lambda_{max}}{\lambda_{min}} \quad (12)$$

where λ_{max} (λ_{min}) is the maximum (minimum) eigenvalue of $\boldsymbol{J}$. The effect on the CRB is depicted in the increase of the Root Mean Square Error (RMSE) of the position and the speed:

$$RMSE_{p,\boldsymbol{\theta}} = \sqrt{\sigma^2_{\widetilde{x_0}} + \sigma^2_{\widetilde{y_0}}} = \sqrt{tr([\mathbf{J}^{-1}]_{(1:2,1:2)})} \quad (13)$$

$$RMSE_{sp,\boldsymbol{\theta}} = \sqrt{\sigma^2_{\widetilde{\upsilon_x}} + \sigma^2_{\widetilde{\upsilon_y}}} = \sqrt{tr([\mathbf{J}^{-1}]_{(3:4,3:4)})} \quad (14)$$

Since we are not interested in the impact of $\mathbf{C}_{\widetilde{\boldsymbol{\theta}}}$ on the performance, we will assume that $\mathbf{C}_{\widetilde{\boldsymbol{\theta}}} = \boldsymbol{I}$. Any scaling by a positive scalar does not change the condition number, while if the diagonal entries change slightly, the condition number changes but its order of magnitude usually remains the same. In table 2 we give the condition numbers and the position and speed Root Mean Square Errors (RMSE) for cases of interest, normalized with respect to the same quantities for the case when all LDP are available, $c_{\boldsymbol{\theta}_{all}}$, $RMSE_{p,\boldsymbol{\theta}_{all}}$ and $RMSE_{sp,\boldsymbol{\theta}_{all}}$ respectively. All quantities are in dB, i.e. the entries are given by

$$c_n = 10\log_{10}\left(\frac{c_{\boldsymbol{\theta}}}{c_{\boldsymbol{\theta}_{all}}}\right) \quad (15)$$

$$RMSE_{p,n} = 10\log_{10}\left(\frac{RMSE_{p,\boldsymbol{\theta}}}{RMSE_{p,\boldsymbol{\theta}_{all}}}\right) \quad (16)$$

$$RMSE_{sp,n} = 10\log_{10}\left(\frac{RMSE_{sp,\boldsymbol{\theta}}}{RMSE_{sp,\boldsymbol{\theta}_{all}}}\right) \quad (17)$$

The cases shown in table 2 are the ones with the best performance over all the cases for which local identifiability is

Table 1. Rank of the transformation matrix

$N_s = 1 \Leftrightarrow N_{\mathbf{p}} = 6, N_t \geq 4$			
$\boldsymbol{\theta}$	$[\boldsymbol{\phi}^t, \mathbf{d}^t]^t$	$[\mathbf{d}^t, \mathbf{f}_d^t]^t$	$[\boldsymbol{\phi}^t, \mathbf{d}^t, \mathbf{f}_d^t]^t$
$rank(\mathbf{G})$	6	6	6
$\boldsymbol{\theta}$	$[\boldsymbol{\phi}^t, \boldsymbol{\psi}^t, \mathbf{d}^t]^t$	$[\boldsymbol{\phi}^t, \boldsymbol{\psi}^t, \mathbf{f}_d^t]^t$	$[\boldsymbol{\psi}^t, \mathbf{d}^t, \mathbf{f}_d^t]^t$
$rank(\mathbf{G})$	6	5	6

Table 2. Normalized Condition Number and RMSE

$N_s = 3, N_t = 50$			
$\boldsymbol{\theta}$	$[\boldsymbol{\phi}^t, \mathbf{d}^t, \mathbf{f}_d^t]^t$	$[\boldsymbol{\phi}^t, \boldsymbol{\psi}^t, \mathbf{d}^t]^t$	$[\boldsymbol{\psi}^t, \mathbf{d}^t, \mathbf{f}_d^t]^t$
c_n	1.2	10.1	6.4
$RMSE_{p,n}$	1.6	7.7	6.4
$RMSE_{sp,n}$	0.6	5.2	9.9

feasible. For the cases not shown in this table, the condition number is increased by at least an order of magnitude and that leads to a similar degradation in performance, when fewer LDP are available. The only exceptions are actually the 3 cases shown in this table. However for 2 of these cases the RMSE for both the position and the speed is more than doubled and only for the case when the AOD is not available, we notice that the degradation is very small.

So far we have consider the improvement in performance of the ML estimation of an identifiable parameter vector $\mathbf{p}$, which occurs when more LDP are exploited in the location estimation process, but the number of nuisance parameters remains constant. It is extremely useful to study the impact on the performance also in situations where exploiting new LDP comes at the cost of jointly estimating a new set of nuisance parameters. This happens when considering a dynamic rather than a static environment and thus the speed of the MT needs to be jointly estimated as mentioned above. This can also happen when the set of unexploited LDP depends deterministically on the entries of $\mathbf{p}$ but also on an unknown error term. For example there might be an unknown synchronization offset that needs to be taken into account for the delays or an orientation/calibration offset that needs to be taken into account for the AOA and/or the AOD. The following theorem applies to all of the above cases and proves when the location ML estimation will be more accurate:

Theorem 1: Introducing and exploiting new LDP $\boldsymbol{\theta}_2$ (data) that depend on the entries of the $N_{\mathbf{p}_1} \times 1$ parameter vector $\mathbf{p}_1$ that needs to be estimated due to the problem formulation (which might consist of parameters of interest and possibly some nuisance parameters as well) but also on the entries of new vector of nuisance parameters $\mathbf{p}_2$, will lead to an enhancement of the performance of the ML estimation only if the transformation matrix $\mathbf{G}_{22} = \frac{\partial \boldsymbol{\theta}_2^t}{\partial \mathbf{p}_2}$ is wide ($N_{\boldsymbol{\theta}_2} > N_{\mathbf{p}_2}$) and has full rank $N_{\mathbf{p}_2}$.

Proof: Let $\boldsymbol{\theta}_1$ be the $N_{\boldsymbol{\theta}_1} \times 1$ vector containing the data that are already used in the estimation process and define $\boldsymbol{\theta} \triangleq$

$[\boldsymbol{\theta}_1^t \boldsymbol{\theta}_2^t]^t$, $\mathbf{p} \triangleq [\mathbf{p}_1^t \mathbf{p}_2^t]^t$, $\mathbf{G} = \frac{\partial \boldsymbol{\theta}^t}{\partial \mathbf{p}}$ and $\mathbf{G}_{ij} = \frac{\partial \boldsymbol{\theta}_i^t}{\partial \mathbf{p}_j}$. The FIM for the new problem is given by :

$$\begin{aligned} \boldsymbol{J}_{new} &= E\Big\{ \left(\frac{\partial \mathcal{L}}{\partial \mathbf{p}}\right) \left(\frac{\partial \mathcal{L}}{\partial \mathbf{p}}\right)^t \Big\} = \mathbf{G}\mathbf{C}^{-1}\mathbf{G}^t \\ &= \begin{bmatrix} \mathbf{G}_{11} & \mathbf{G}_{21} \\ \mathbf{0} & \mathbf{G}_{22} \end{bmatrix} \begin{bmatrix} \mathbf{C}_{\boldsymbol{\theta}_1|\mathbf{p}} & \mathbf{0} \\ \mathbf{0} & \mathbf{C}_{\boldsymbol{\theta}_2|\mathbf{p}} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{G}_{11}^t & \mathbf{0} \\ \mathbf{G}_{21}^t & \mathbf{G}_{22}^t \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{G}_{11}\mathbf{C}_{\boldsymbol{\theta}_1|\mathbf{p}}^{-1}\mathbf{G}_{11}^t + \mathbf{G}_{21}\mathbf{C}_{\boldsymbol{\theta}_2|\mathbf{p}}^{-1}\mathbf{G}_{21}^t & \mathbf{G}_{21}\mathbf{C}_{\boldsymbol{\theta}_2|\mathbf{p}}^{-1}\mathbf{G}_{22}^t \\ \mathbf{G}_{22}\mathbf{C}_{\boldsymbol{\theta}_2|\mathbf{p}}^{-1}\mathbf{G}_{21}^t & \mathbf{G}_{22}\mathbf{C}_{\boldsymbol{\theta}_2|\mathbf{p}}^{-1}\mathbf{G}_{22}^t \end{bmatrix} \\ &\triangleq \begin{bmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{C} & \boldsymbol{D} \end{bmatrix} \end{aligned} \quad (18)$$

The first term of the sum composing $\boldsymbol{A}$ can be recognized as the FIM $\boldsymbol{J}_1$ of the original estimation problem while $\boldsymbol{D}$ is the FIM $\boldsymbol{J}_2$ for the estimation problem of $\mathbf{p}_2$.

If both of the conditions for $\boldsymbol{G}_{22}$ are met, we can use the inversion formula for 2×2 block matrices to obtain the $N_{\mathbf{p}_1} \times N_{\mathbf{p}_1}$ upper left submatrix of the inverse of $\boldsymbol{J}_{new}$ as follows:

$$\begin{aligned} &[\boldsymbol{J}_{new}^{-1}]_{(1:L,1:L)} = (\boldsymbol{A} - \boldsymbol{B}\boldsymbol{D}^{-1}\boldsymbol{C})^{-1} \\ &(\ \boldsymbol{J}_1 + \mathbf{G}_{21}\mathbf{C}_{\boldsymbol{\theta}_2|\mathbf{p}}^{-1}\mathbf{G}_{21}^t \\ &- \mathbf{G}_{21}\mathbf{C}_{\boldsymbol{\theta}_2|\mathbf{p}}^{-1}\mathbf{G}_{22}^t\boldsymbol{J}_2^{-1}\mathbf{G}_{22}\mathbf{C}_{\boldsymbol{\theta}_2|\mathbf{p}}^{-1}\mathbf{G}_{21}^t\ )^{-1} \end{aligned} \quad (19)$$

To show that the performance is improved with the addition of new data, it suffices to show that $[\boldsymbol{J}_{new}]_{(1:L,1:L)} > \boldsymbol{J}_1$. This is true because the sum of the other two matrices on the r.h.s. of (19) results in a positive semidefinite matrix. This can be proved by defining $\mathbf{a} \triangleq \mathbf{C}_{\boldsymbol{\theta}_2|\mathbf{p}}^{-1/2}\mathbf{G}_{22}^t\mathbf{u}$ and $\mathbf{b} \triangleq \mathbf{G}_{21}\mathbf{C}_{\boldsymbol{\theta}_2|\mathbf{p}}^{-1}\mathbf{G}_{22}^t(\mathbf{G}_{22}\mathbf{C}_{\boldsymbol{\theta}_2|\mathbf{p}}^{-1}\mathbf{G}_{22}^t)^{-1/2}\mathbf{u}$, where $\mathbf{u}$ is any non-zero vector and applying Cauchy-Schwartz inequality $||\mathbf{a}||^2||\mathbf{b}||^2 \geq (\mathbf{a}^t\mathbf{b})^2$, to get:

$$\begin{aligned} &\mathbf{u}^t\ (\ \mathbf{G}_{21}\mathbf{C}_{\boldsymbol{\theta}_2|\mathbf{p}}^{-1}\mathbf{G}_{21}^t - \\ &\mathbf{G}_{21}\mathbf{C}_{\boldsymbol{\theta}_2|\mathbf{p}}^{-1}\mathbf{G}_{22}^t\boldsymbol{J}_2^{-1}\mathbf{G}_{22}\mathbf{C}_{\boldsymbol{\theta}_2|\mathbf{p}}^{-1}\mathbf{G}_{21}^t\ )\ \mathbf{u} \geq 0 \end{aligned} \quad (20)$$

If $N_{\boldsymbol{\theta}_2} = N_{\mathbf{p}_2}$ and $\boldsymbol{G}_2$ has full rank, (19) reduces to :

$$[\boldsymbol{J}_{new}^{-1}]_{(1:L,1:L)} = \boldsymbol{J}_1^{-1} \quad (21)$$

thus the performance of the ML estimation is exactly the same, while the complexity of the method increases.

If $N_{\boldsymbol{\theta}_2} < N_{\mathbf{p}_2}$ and $\boldsymbol{G}_2$ is full rank, $\boldsymbol{D}$ is singular and thus non-invertible. The parameter vector $\mathbf{p}_2$ is not identifiable, however $\mathbf{p}_1$ is still identifiable and $[\boldsymbol{J}_{new}^{-1}]_{(1:L,1:L)}$ can be derived by replacing the inverse of $\boldsymbol{D}^{-1}$ with its pseudo-inverse $\boldsymbol{D}^+ = \mathbf{G}_{22}(\mathbf{G}_{22}^t\mathbf{G}_{22})^{-1}\mathbf{C}_{\boldsymbol{\theta}_2|\mathbf{p}}(\mathbf{G}_{22}^t\mathbf{G}_{22})^{-1}\mathbf{G}_{22}^t$ in (19). This results again in (21) and thus in no improvement in performance.

Finally, if $\boldsymbol{G}_{22}$ has rank $k < \min\{N_{\boldsymbol{\theta}_2}, N_{\mathbf{p}_2}\}$, $N_{\boldsymbol{\theta}_2} - k$ of its columns contain no additional information for $\mathbf{p}_2$. If they contain additional information only for $\mathbf{p}_1$ then, the information they contain for $\mathbf{p}_2$ (if any) could be removed by elementary column operations so that the corresponding entries of $\boldsymbol{\theta}_2$ can be included in $\boldsymbol{\theta}_1$, leading to a different partioning of the FIM. That will improve performance, since new LDP that depend only on $\mathbf{p}_1$ and no new nuisance parameters, are exploited. If, however, they contain no additional information on any of the entries of $\mathbf{p}$ then they should be discarded. Either way, $\boldsymbol{G}_2$ becomes a $N_{\mathbf{p}_2} \times k$ full rank matrix and the last of the above cases applies. This completes the proof.

Theorem 1 can be applied in any ML localization approach. For example, proposition 1 in [4] can be derived using this theorem. The fact that the performance of a TOA method does not improve by including the TOA of the NLOS components, if these are modeled as the TOA corresponding to the LOS component plus an unknown error term, is a direct application of theorem 1 with $N_{\boldsymbol{\theta}_2} = N_{\mathbf{p}_2}$.

7. CONCLUSIONS

In this paper we have thoroughly investigated identifiability and performance for a ML location estimation technique which is suitable for strictly NLOS propagation environments. The impact of exploiting various location-dependent parameters, like the AOA, the AOD and the path lengths (delays) along with the location and motion dependent doppler shift has been depicted. In general, considering a dynamic instead of a static channel can lead not only to an enhancement in performance, but also to local identifiability of the location in cases when fewer subsets of LDP are available. In these cases, the condition number of the FIM can be used as an indicator for the performance. The results presented herein can be easily extended to include cases when prior statistics of the LDP are available or can be obtained by inference. The information will just have to be added, by adding a corresponding matrix $\mathbf{J}_{prior}$ to the FIM.

8. REFERENCES

[1] 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.

[2] K. Papakonstantinou and D. Slock, "NLOS Mobile Terminal Position and Speed Estimation," in *Proc. 3rd International Symposium on Communications, Control and Signal Processing*, Mar. 2008.

[3] Thomas J. Rothenberg, "Identification in Parametric Models," *Econometrica*, vol. 39, no. 3, pp. 577 – 591, 1971.

[4] 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.