The 2009 IEEE/RSJ International Conference on
Intelligent Robots and Systems
October 11-15, 2009 St. Louis, USA

On the Nonlinear Observability and the Information form of the SLAM Problem

L.D.L. Perera and Eric Nettleton

Abstract—**The theory of nonlinear observability is an important tool available for the assessment of highly nonlinear estimation problems such as Simultaneous Localization and Mapping (SLAM). It is shown that all the estimated landmarks must be observed and at least two a priori known landmarks be observed for the nonlinear observability of single vehicle SLAM when estimating any number of unknown landmark locations. The relationship between the information form of SLAM and the nonlinear observability is established. It is shown that when the nonlinear observability conditions are satisfied the single vehicle SLAM problem can in theory be initialized with infinitely large initial uncertainties. Simulations and experiments are also provided to substantiate the theoretical results.**

Index Terms—**SLAM, observability, information matrix**

I. Introduction

Simultaneous localization and mapping (SLAM) attempts to estimate the robot pose and the map at the same time and is thought by many ([1] and [2]) as one of the most significant steps towards achieving complete autonomy for mobile robots. However, until recent years the observability, which is fundamental to the very existence of any estimation algorithm has received considerably little attention. Most of this initial work is also based on the linear observability theory. For the first time [3] described the partial observability of the one dimensional linear SLAM problem and attempted to extend the same theory for nonlinear models using linear approximations. [4] and [5] use piecewise constant systems theory [14] for the observability analysis of the SLAM problem. [4] and [5] assume that SLAM is piece wise constant in time segments and use linearized systems in piecewise constant segments to study the observability using the theory of the observability of linear systems. Nonlinear observability theory has been applied for the analysis of observability of SLAM in [6] and [7]. [6] and [7] analyze the local observability of one landmark SLAM problem using the nonlinear observability theory. However, [6] and [7] do not generalize the theory to estimating any number of unknown landmarks or discuss the properties or applications of the nonlinear observability of SLAM in detail. Although the effects of inputs on linear systems are not taken into account in the linear observability analysis, they can significantly affect the behavior of nonlinear systems through their nonlinear models. In this context, nonlinear observability analysis [8] is more appropriate than the linear analysis in describing highly nonlinear and time varying systems such as SLAM. In this paper we present proofs for the nonlinear observability properties of the SLAM problem and show the relationship of the nonlinear observability properties to the information form of the SLAM problem. It has also been shown that the kidnapped robot problem in theory can be solved if the conditions required for the nonlinear observability of SLAM are satisfied.

The paper is organized as follows. Section II describes the theory of the nonlinear observability. Section III provides proofs on the nonlinear observability properties of the single vehicle SLAM problem. Section IV details the implications of the nonlinear observability conditions proved in Section III on the information form of the SLAM problem. It is also shown that the above nonlinear observability properties affect the kidnapped robot problem. Section V provides simulations and experiments to substantiate the theoretical results established. Section VI discusses the results and concludes the work.

II. Nonlinear Observability of The SLAM Problem

A. Theory of Nonlinear Observability

Conceptually, observability determines if there is adequate information in the form of measurements and models of the estimation problem to consistently estimate the state variables in finite time. In general, linear observability theory is applied successfully to linear time invariant systems. For linear time invariant systems the observability tests check whether initial states of the linear time invariant systems can be recovered from a finite number of observations of their outputs and the knowledge of their inputs. In other words the observability by definition is a global phenomenon. It may require a system to go through a long duration of time or distance to distinguish among points.

This work has been supported by the Rio Tinto Centre for Mine Automation and the ARC Center of Excellence program funded by the Australian Research Council (ARC) and the New South Wales State Government.

L.D.L. Perera and Eric Nettleton are at the Australian Centre for Field Robotics, The University of Sydney, NSW 2006, Sydney, Australia (e-mail: l.perera; e.nettleton@acfr.usyd.edu.au).

The linear observability theory has many short comings when applied to the nonlinear systems. It requires linearization of the process and measurement models even though these models are highly nonlinear and do not represent their characteristics in linearlized forms. Furthermore, linear observability theory does not take the effects of inputs on the observability into consideration. Since the SLAM problem is a highly nonlinear state estimation problem, it is more rigorous to take the nonlinear properties of the problem into consideration using an appropriate observability theory.

In the following we review the basics of nonlinear observability theory which can be applicable to highly nonlinear problems. Let $\mathbf{\Sigma}$ be a nonlinear state estimation problem defined by

$$\mathbf{\Sigma}\begin{cases}\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x},\ \mathbf{u}) \\ \mathbf{z} = \mathbf{h}(\mathbf{x})\end{cases} \tag{1}$$

where $\mathbf{x}$ is the state vector estimated, $\mathbf{u}$ the control inputs, $\mathbf{z}$ the measurement vector and $\mathbf{f}(.)$ and $\mathbf{h}(.)$ nonlinear functions designating the process model and the measurement model respectively. The theory of nonlinear observability is based on the following definitions.

Definition 1: The two states $\mathbf{x}_1$, $\mathbf{x}_2 \in \mathbf{x}$ are said to be distinguishable iif for any two points $\{\mathbf{x}_1, t\}$ and $\{\mathbf{x}_2, t\}$ that satisfy $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x},\ \mathbf{u})$ with time $t \in [0, T]$ and the initial conditions $\mathbf{x}_1$ and $\mathbf{x}_2$, there exists at least one value of t such that $\mathbf{h}(\mathbf{x}_1) \neq \mathbf{h}(\mathbf{x}_2)$.

Definition 2: The system $\mathbf{\Sigma}$ is weakly observable at $\mathbf{x}_0$ if there exists a neighborhood of $\mathbf{x}_0$ such that every $\mathbf{x}$ in that neighborhood other than $\mathbf{x}_0$ is distinguishable from $\mathbf{x}_0$.

Definition 3: The system $\mathbf{\Sigma}$ is locally observable at $\mathbf{x}_0$ if for every open neighborhood $\mathbf{N}$ of $\mathbf{x}_0$, $\mathbf{x}_0$ is distinguishable from any other point in $\mathbf{N}$.

Thus, the nonlinear observability according to [8] is a local phenomenon. For a special class of nonlinear problems [13] has shown the following result (Theorem 1).

Theorem 1: If $\mathbf{\Sigma}$ is in control affine form $\mathbf{f}(\mathbf{x},\ \mathbf{u}) = \mathbf{g}^0(\mathbf{x}) + \sum \mathbf{g}^i(\mathbf{x})u_i$ where $\mathbf{x}$ is a vector of n state variables occupying an open subset Ξ of $\mathbb{R}^n$, $\mathbf{g}^0(.)$ **,....,** $\mathbf{g}^i(.)$ are n dimensional vector analytic functions in Ξ, the measurement function $\mathbf{h}(.)$ is an analytic function of $\mathbb{R}^m$ in Ξ and $\mathbf{u}$ is an analytic function of time comprising distinct scalar controls u_i, then $\mathbf{\Sigma}$ is locally weakly observable if the matrix $\mathbf{O}_\Sigma$ (hereinafter we refer to as the nonlinear observability matrix) given below has rank n.

$$\mathbf{O}_\Sigma(\mathbf{d}, \mathbf{f}, \mathbf{h}, n) = \left[(\mathbf{d}L_{\mathbf{f}}^0\mathbf{h})^T \quad (\mathbf{d}L_{\mathbf{f}}^1\mathbf{h})^T \quad \ldots \quad (\mathbf{d}L_{\mathbf{f}}^{n-1}\mathbf{h})^T\right]^T \tag{2}$$

Conversely for systems that are control affine, considering from zero up to n-1[th] order Lie derivatives is adequate to determine the nonlinear observability

A. Nonlinear observability of the SLAM problem

The single vehicle SLAM problem [1] and [2] is a highly nonlinear state estimation problem. Let there be a vehicle moving on a 2D flat surface and estimating its pose $\mathbf{X}_v(t)$ and a map $\mathbf{m}(t)$ with the location states of n landmarks. The estimated states are

$$\mathbf{X}_v(t) = \left[x_v(t) \quad y_v(t) \quad \theta_v(t)\right]^T \tag{3}$$

$$\mathbf{m}(t) = \left[x_1(t) \quad y_1(t) \quad \ldots \quad x_n(t) \quad y_n(t)\right]^T \tag{4}$$

$$\mathbf{X}_n(t) = \left[\mathbf{X}_v^T(t) \quad \mathbf{m}^T(t)\right]^T \tag{5}$$

where $x_v(t)$ is the vehicle's longitudinal coordinate $y_v(t)$ is the vehicle's lateral coordinate, $\theta_v(t)$ is the vehicle heading and $x_i(t)$ and $y_i(t)$ $\forall i$ are the longitudinal and lateral coordinates of the i[th] estimated landmark. The process model assuming a car-like vehicle model and the measurement model assuming a range, bearing sensor are;

$$\dot{\mathbf{X}}_n(t) = \mathbf{f}(\mathbf{X}_n(t),\ \mathbf{u}) + \boldsymbol{\eta}_1(t) \tag{6}$$

$$\mathbf{Z} = \mathbf{h}(\mathbf{X}_n(t)) + \boldsymbol{\eta}_2(t) \tag{7}$$

$$\mathbf{f}(n) = \left[u\cos(\theta(t)) \quad u\sin(\theta(t)) \quad \omega \quad 0 \quad \ldots \quad 0\right]^T \tag{8}$$

$$\mathbf{h}(n) = [(\mathbf{h}_1)^T \quad (\mathbf{h}_2)^T \quad \ldots \quad (\mathbf{h}_n)^T]^T \tag{9}$$

$$\mathbf{h}_i = \begin{bmatrix} \sqrt{(x_i(t) - x_v(t))^2 + (y_i(t) - y_v(t))^2} \\ \tan^{-1}\{(y_i(t) - y_v(t))/(x_i(t) - x_v(t))\} - \theta_v(t) \end{bmatrix} \tag{10}$$

$$\omega = u\tan\gamma / W_b \tag{11}$$

where $\boldsymbol{\eta}_1(t)$ and $\boldsymbol{\eta}_2(t)$ are zero mean noise terms representing the process noise with the covariance matrix $\mathbf{Q}(t)$, and measurement noise with the covariance matrix $\mathbf{R}(t)$, u is the speed input, γ is the steering angle input and W_b is the vehicle wheel base. $\mathbf{h}(\mathbf{X}_n(t))$ is denoted by $\mathbf{h}(n)$ for notational simplicity. Here we choose a new time scale ρ so that ρ is the distance along the vehicle path. Thus, $d\rho/dt$ is the vehicle speed. Hence multiplying (6) by u^{-1} we obtain the n landmark SLAM problem in control affine form ([6]).

III. Analysis of the NonLinear Observability Matrix

Since, the single vehicle SLAM problem is in control affine form we apply the result of Theorem 1 for the observability analysis of the SLAM problem resulting in the following nonlinear observability matrix ($\mathbf{O}_n$).

$$\mathbf{O}_n = \begin{bmatrix} \mathbf{d}_n L_{\mathbf{f}(n)}^0 \mathbf{h}(n) \\ \mathbf{d}_n L_{\mathbf{f}(n)}^1 \mathbf{h}(n) \\ \cdots\cdots\cdots\cdots \\ \mathbf{d}_n L_{\mathbf{f}(n)}^{3+2n-1} \mathbf{h}(n) \end{bmatrix} \tag{12}$$

where $\mathbf{d}_n$ is the gradient operator with respect to $\mathbf{X}_n(t)$ and $L^i_{\mathbf{f}(n)}$ is the Lie derivative of order *i* with respect to $\mathbf{f}(n)$.

A. Nonlinear Observability Matrix Properties

We now investigate the properties of the nonlinear observability matrix of SLAM to gain a deeper insight. Hereinafter we omit the symbol *t* indicating the time varying properties from the equations for clarity. By definition;

$$L^0_{\mathbf{f}(n)}\mathbf{h}_j = \mathbf{h}_j = \begin{bmatrix} h_{j,1,0}((x_j - x_v),(y_j - y_v)) \\ h_{j,2,0}((x_j - x_v),(y_j - y_v),\theta_v) \end{bmatrix} \tag{13}$$

where $h_{j,1,i}$ and $h_{j,2,i}$ denote the i^{th} order Lie derivative of the range measurement and the bearing measurement in functional form of the j^{th} landmark respectively. It is interesting to note that both $h_{j,1,0}$ and $h_{j,2,0}$ can be represented as functions of $(x_j - x_v)$ and $(y_j - y_v)$. Also note that $h_{j,1,0}$ and $h_{j,2,0}$ are not functions of any other landmark states. Now let

$$\mathbf{d}_n L^0_{\mathbf{f}(n)}\mathbf{h}_j = \begin{bmatrix} h_j^{1,0} & h_j^{2,0} & h_j^{3,0} & \mathbf{0}_1 & h_j^{4,0} & h_j^{5,0} & \mathbf{0}_2 \\ h_j^{6,0} & h_j^{7,0} & h_j^{8,0} & \mathbf{0}_3 & h_j^{9,0} & h_j^{10,0} & \mathbf{0}_4 \end{bmatrix} \tag{14}$$

where $\mathbf{0}_1 = \mathbf{0}_3 = \mathbf{0}_{1\times 2(j-1)}$ for $\forall\, j > 1$ and $\mathbf{0}_1$ and $\mathbf{0}_3$ do not exist for $j = 1$. $\mathbf{0}_2 = \mathbf{0}_4 = \mathbf{0}_{1\times 2(n-j)}$ for $\forall$ n > j and $\mathbf{0}_2$ and $\mathbf{0}_4$ do not exist for $j = n$. Thus, when *q* is a positive integer, by recursion we can show that $\mathbf{d}_n L^q_{\mathbf{f}(n)}\mathbf{h}_j$ takes the form similar to (14).

$$\mathbf{d}_n L^q_{\mathbf{f}(n)}\mathbf{h}_j = \begin{bmatrix} h_j^{1,q} & h_j^{2,q} & h_j^{3,q} & \mathbf{0}_1 & h_j^{4,q} & h_j^{5,q} & \mathbf{0}_2 \\ h_j^{6,q} & h_j^{7,q} & h_j^{8,q} & \mathbf{0}_3 & h_j^{9,q} & h_j^{10,q} & \mathbf{0}_4 \end{bmatrix} \tag{15}$$

Let the columns of $\mathbf{O}_n$ be denoted by $\mathbf{C}^i(n)$ where *i* denotes the i^{th} column of $\mathbf{O}_n$ and $\mathbf{C}^i_j(n)$ denotes the parts of the i^{th} column of $\mathbf{O}_n$ corresponding to the j^{th} landmark. Here (n) in $\mathbf{C}^i_j(n)$ is used to denote $3+2n-1$ Lie derivatives of the observations corresponding to the j^{th} landmark.

Result 1: The nonlinear observability matrix of the SLAM problem estimating n unknown landmarks is rank deficient by 3 when all the estimated landmarks are observed.

Proof: The nonlinear observability matrix for SLAM estimating 1 landmark is;

$$\mathbf{O}_1 = \begin{bmatrix} \mathbf{d}_1 L^0_{\mathbf{f}(1)}\mathbf{h}(1) \\ \mathbf{d}_1 L^1_{\mathbf{f}(1)}\mathbf{h}(1) \\ \cdots\cdots\cdots \\ \mathbf{d}_1 L^4_{\mathbf{f}(1)}\mathbf{h}(1) \end{bmatrix} \tag{16}$$

By row reduction and simplification of (16) and from (13) it follows that the rank of $\mathbf{O}_1$ is 2. It is therefore rank deficient by 3. Hence Result 1 is true for *n*=1. Assume now that Result 1 is true for *n=p* where *p* is a positive integer.

$$\mathbf{O}_p = \begin{bmatrix} \mathbf{d}_p L^0_{\mathbf{f}(p)}\mathbf{h}(p) \\ \mathbf{d}_p L^1_{\mathbf{f}(p)}\mathbf{h}(p) \\ \cdots\cdots\cdots \\ \mathbf{d}_p L^{3+2p-1}_{\mathbf{f}(p)}\mathbf{h}(p) \end{bmatrix} \tag{17}$$

Using the vector properties of $\mathbf{d}_n L^q_{\mathbf{f}(n)}\mathbf{h}_j$ for any positive integer *q*, *n* and *j* to reorder the rows of $\mathbf{O}_p$, we can also express (17) as follows.

$$\mathbf{O}_p = \left[\begin{array}{c|cccc} \mathbf{C}^v_1(p) & \mathbf{C}^m_1(p) & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{C}^v_2(p) & \mathbf{0} & \mathbf{C}^m_2(p) & \cdots & \mathbf{0} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \mathbf{C}^v_p(p) & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{C}^m_p(p) \end{array}\right] \tag{18}$$

where $\mathbf{C}^v_j(p)$ denotes the parts of the first three columns of $\mathbf{O}_p$ associated with the vehicle state and the j^{th} landmark and $\mathbf{C}^m_j(p)$ denotes the parts of the two columns of $\mathbf{O}_p$ associated with the j^{th} landmark state comprising all the Lie derivatives from order 0 to $2+2p$. We now consider the nonlinear observability matrix when one landmark state is added to the state vector;

$$\mathbf{O}_{p+1} = \begin{bmatrix} \mathbf{d}_{p+1} L^0_{\mathbf{f}(p+1)}\mathbf{h}(p+1) \\ \mathbf{d}_{p+1} L^1_{\mathbf{f}(p+1)}\mathbf{h}(p+1) \\ \cdots\cdots\cdots \\ \mathbf{d}_{p+1} L^{4+2p}_{\mathbf{f}(p+1)}\mathbf{h}(p+1) \end{bmatrix} \tag{19}$$

By definition

$$L^0_{\mathbf{f}(p+1)}\mathbf{h}(p+1) = \begin{bmatrix} \mathbf{h}(p) \\ \mathbf{h}_{p+1} \end{bmatrix} \tag{20}$$

$$\begin{aligned} \mathbf{d}_{p+1} L^0_{\mathbf{f}(p+1)}\mathbf{h}(p+1) &= \left(\partial/\partial \mathbf{X}_{p+1}\right) L^0_{\mathbf{f}(p+1)}\mathbf{h}(p+1) \\ &= \begin{bmatrix} \mathbf{d}_p L^0_{\mathbf{f}(p)}\mathbf{h}(p) \quad \mathbf{0} \quad \mathbf{0} \\ \mathbf{d}_{p+1} L^0_{\mathbf{f}(p+1)}\mathbf{h}_{p+1} \end{bmatrix} \end{aligned} \tag{21}$$

$$\mathbf{d}_{p+1} L^1_{\mathbf{f}(p+1)}\mathbf{h}(p+1) = \begin{bmatrix} \mathbf{d}_p L^1_{\mathbf{f}(p)}\mathbf{h}(1,p) \quad \mathbf{0} \quad \mathbf{0} \\ \mathbf{d}_{p+1} L^1_{\mathbf{f}(p+1)}\mathbf{h}_{p+1} \end{bmatrix} \tag{22}$$

Therefore, by recursion it can be shown that;

$$\mathbf{d}_{p+1} L^r_{\mathbf{f}(p+1)}\mathbf{h}(p+1) = \begin{bmatrix} \mathbf{d}_p L^r_{\mathbf{f}(p)}\mathbf{h}(p) \quad \mathbf{0} \quad \mathbf{0} \\ \mathbf{d}_{p+1} L^r_{\mathbf{f}(p+1)}\mathbf{h}_{p+1} \end{bmatrix} \tag{23}$$

for any positive integer *r*. Thus, we can rewrite $\mathbf{O}_{p+1}$ as follows.

$$\mathbf{O}_{p+1} = \begin{bmatrix} \mathbf{O}_p & \mathbf{0} & \mathbf{0} \\ \mathbf{d}_p L_{\mathbf{f}(p)}^{3+2p}\mathbf{h}(p) & \mathbf{0} & \mathbf{0} \\ \mathbf{d}_p L_{\mathbf{f}(p)}^{3+2p+1}\mathbf{h}(p) & \mathbf{0} & \mathbf{0} \\ \mathbf{d}_{p+1} L_{\mathbf{f}(p+1)}^{0}\mathbf{h}_{p+1} & & \\ \cdots\cdots\cdots & & \\ \mathbf{d}_{p+1} L_{\mathbf{f}(p+1)}^{4+2p}\mathbf{h}_{p+1} & & \end{bmatrix} \quad (24)$$

From (18) we can write $\mathbf{O}_{p+1}$ as follows.

$$\mathbf{O}_{p+1} = \left[\begin{array}{c|ccccc} \mathbf{C}_1^v(p+1) & \mathbf{C}_1^m(p+1) & .. & \mathbf{0} & \mathbf{0} \\ .. & .. & .. & .. & .. \\ \mathbf{C}_p^v(p+1) & \mathbf{0} & .. & \mathbf{C}_p^m(p+1) & \mathbf{0} \\ \hline \mathbf{C}_{p+1}^v(p+1) & \mathbf{0} & .. & \mathbf{0} & \mathbf{C}_{p+1}^m(p+1) \end{array}\right] \quad (25)$$

Now consider $\mathbf{O}_p$, $\mathbf{d}_p L_{\mathbf{f}(p)}^{3+2p}\mathbf{h}(p)$, and $\mathbf{d}_p L_{\mathbf{f}(p)}^{3+2p+1}\mathbf{h}(p)$ terms from (24). From the converse of Theorem 1 and the assumption of the rank condition of $\mathbf{O}_p$ it can be concluded that when considered together as rows of a matrix these terms have 3 null columns.

According to the assumption on rank of $\mathbf{O}_p$, it has three null column vectors. Consider one such null column in $\mathbf{O}_p$ meaning we can find transformations (T_i^v for $i = 1$, 2 and 3 and T_{3+j}^m for $j = 1, 2, ..., 2p$) on the columns of $\mathbf{O}_p$ (from (18)) such that the following expression is true.

$$T_1^v\mathbf{C}^1(p) + T_2^v\mathbf{C}^2(p) + T_3^v\mathbf{C}^3(p) + \sum_{j=1}^{p}\left\{T_{3+2j-1}^m\mathbf{C}^{3+2j-1}(p) + T_{3+2j}^m\mathbf{C}^{3+2j}(p)\right\} = \mathbf{0} \quad (26)$$

By expanding (26) we have the following expression for all $j = 1, 2, 3, ..., p$

$$T_1^v\mathbf{C}_j^1(p) + T_2^v\mathbf{C}_j^2(p) + T_3^v\mathbf{C}_j^3(p) + T_{3+2j-1}^m\mathbf{C}_j^{3+2j-1}(p) + T_{3+2j}^m\mathbf{C}_j^{3+2j}(p) = \mathbf{0} \quad (27)$$

Now consider $\mathbf{O}_{p+1}$ shown in (24). Since it was shown that the terms $\mathbf{O}_p$, $\mathbf{d}_p L_{\mathbf{f}(p)}^{3+2p}\mathbf{h}(p)$, and $\mathbf{d}_p L_{\mathbf{f}(p)}^{4+2p}\mathbf{h}(p)$ have three null columns we can rewrite (27) including Lie derivatives of order $3+2p$ and $4+2p$ as follows.

For $j = 1, 2, 3, ..., p$;

$$T_1^v\mathbf{C}_j^1(p+1) + T_2^v\mathbf{C}_j^2(p+1) + T_3^v\mathbf{C}_j^3(p+1) + T_{3+2j-1}^m\mathbf{C}_j^{3+2j-1}(p+1) + T_{3+2j}^m\mathbf{C}_j^{3+2j}(p+1) = \mathbf{0} \quad (28)$$

Here, (28) covers all the rows from 1 to that corresponding to $\mathbf{C}_p^v(p+1)$ of (25). Let

$$T_1^v\mathbf{C}_{p+1}^1(p+1) + T_2^v\mathbf{C}_{p+1}^2(p+1) + T_3^v\mathbf{C}_{p+1}^3(p+1) + T_{4+2p}^m\mathbf{C}_{p+1}^{4+2p}(p+1) + T_{5+2p}^m\mathbf{C}_{p+1}^{5+2p}(p+1) = \mathbf{n}_1 \quad (29)$$

where $\mathbf{n}_1$ is a column vector to be determined.

It follows from (25) that the structure of $\mathbf{O}_{p+1}$ is such that the rows corresponding to the landmark $p+1$ can be interchanged with the rows corresponding to any other landmark $j = 1, 2, 3, ..., p$. Thus, from (28) and (29) it can be concluded that $\mathbf{n}_1 = \mathbf{0}$ when $j = p+1$. Hence by combining (28) and (29);

$$T_1^v\mathbf{C}^1(p+1) + T_2^v\mathbf{C}^2(p+1) + T_3^v\mathbf{C}^3(p+1) + \sum_{j=1}^{p+1}\left\{T_{3+2j-1}^m\mathbf{C}^{3+2j-1}(p+1) + T_{3+2j}^m\mathbf{C}^{3+2j}(p+1)\right\} = \mathbf{0} \quad (30)$$

Hence (30) results in a null column for $\mathbf{O}_{p+1}$. Therefore, all 3 null columns in $\mathbf{O}_p$ also result in null columns in $\mathbf{O}_{p+1}$. However, $\mathbf{O}_p$ has a maximum of three null columns according to the assumption made on its rank. From (25), if you do any column operations on columns 1 to $5+2p$ using all the columns of $\mathbf{O}_{p+1}$, you cannot have more than 3 null columns because it contradicts with (28)-(30) on the rank of the $\mathbf{O}_p$. Therefore, if there are any more than 3 null columns in (24) they must be obtained only from column operations on columns $4+2p$ and $5+2p$, which are the last two columns of $\mathbf{O}_{p+1}$.

The zero order Lie derivatives corresponding to the vehicle observing the landmark $p+1$ can be written as.

$$\mathbf{M}_{p+1}^m = \begin{bmatrix} h_{p+1}^{4,0} & h_{p+1}^{4,0} \\ h_{p+1}^{9,0} & h_{p+1}^{10,0} \end{bmatrix} \quad (31)$$

By simplification we can show that rank of $\mathbf{M}_{p+1}^m$ is 2. Thus, columns $4+2p$ and $5+2p$ are independent. Hence the maximum number of null columns in $\mathbf{O}_{p+1}$ is 3. Therefore, Result 1 is true for $n = p+1$. Since Result 1 is true for $n = 1$ and given it is true for $n = p$ and for $n = p+1$, by the principle of mathematical induction it is true for any positive integer n.

Result 2: The SLAM problem becomes locally weakly observable when (1) Observing two known landmarks or when (2) Observing vehicle's longitudinal and lateral coordinates provided the vehicle observes all the estimated landmarks.

Proof: Let x_1^*, y_1^* x_2^* and y_2^* be the longitudinal and lateral coordinates of the two known landmarks and $\mathbf{C}^i(n)$ be the i^{th} column of the nonlinear observability matrix. Consider the situation of observing two known landmarks first. Let

$$\mathbf{h}_j^* = \begin{bmatrix} h_{j,1,0}^*((x_j^* - x_v), (y_j^* - y_v)) \\ h_{j,2,0}^*((x_j^* - x_v), (y_j^* - y_v), \theta_v) \end{bmatrix} \quad (32)$$

where $\mathbf{h}_j^*$ is the measurement model when the range bearing of the known landmark j is observed and $h_{j,1,i}^*$ and $h_{j,2,i}^*$ are the i^{th} order Lie derivative of the measurement models when observing range and bearing to the j^{th} known landmark. By simplification, the nonlinear observability matrix $\bar{\bar{\mathbf{O}}}_p$ for estimating p number of landmarks now is;

$$\bar{\bar{\mathbf{O}}}_p = \begin{bmatrix} \mathbf{O}_p & \\ \bar{\mathbf{C}}_1^v(p) & \mathbf{0} \\ \bar{\mathbf{C}}_2^v(p) & \mathbf{0} \end{bmatrix} \tag{33}$$

where $\bar{\mathbf{C}}_j^v(p)$ is a three column sub matrix denoting the Lie derivatives of order 0 to $2+2p$ of $\mathbf{h}_j^*$ with respect to $\mathbf{f}(p)$. If we let n=1, it can be shown from (33) that the rank of $\bar{\bar{\mathbf{O}}}_1$ is equal to five (full rank) when observing two known landmarks. Assume that when $j = p$ the nonlinear observability matrix is full rank. That is, the rank of $\bar{\bar{\mathbf{O}}}_p$ is $2p+3$. Now consider $j = p+1$ which corresponds to the addition of one landmark state into the state vector.

$$\bar{\bar{\mathbf{O}}}_{p+1} = \begin{bmatrix} \mathbf{O}_{p+1} & \\ \bar{\mathbf{C}}_1^v(p+1) & \mathbf{0} \\ \bar{\mathbf{C}}_2^v(p+1) & \mathbf{0} \end{bmatrix} \tag{34}$$

From (24) the nonlinear observability matrix $\bar{\bar{\mathbf{O}}}_{p+1}$ takes the following form;

$$\bar{\bar{\mathbf{O}}}_{p+1} = \begin{bmatrix} \bar{\bar{\mathbf{O}}}_p & \mathbf{0} & \mathbf{0} \\ \mathbf{d}_p L_{\mathbf{f}(p)}^{3+2p} \bar{\bar{\mathbf{h}}}(p) & \mathbf{0} & \mathbf{0} \\ \mathbf{d}_p L_{\mathbf{f}(p)}^{4+2p} \bar{\bar{\mathbf{h}}}(p) & \mathbf{0} & \mathbf{0} \\ \mathbf{d}_{p+1} L_{\mathbf{f}(p+1)}^{0} \mathbf{h}_{p+1} & & \\ \cdots\cdots\cdots\cdots\cdots & & \\ \mathbf{d}_{p+1} L_{\mathbf{f}(p+1)}^{4+2p} \mathbf{h}_{p+1} & & \end{bmatrix} \tag{35}$$

where

$$\bar{\bar{\mathbf{h}}}(p) = \begin{bmatrix} (\mathbf{h}_1)^T & (\mathbf{h}_2)^T & .. & (\mathbf{h}_p)^T & (\mathbf{h}_1^*)^T & (\mathbf{h}_2^*)^T \end{bmatrix}^T \tag{36}$$

We now consider the terms $\bar{\bar{\mathbf{O}}}_p$, $\mathbf{d}_p L_{\mathbf{f}(p)}^{3+2p} \bar{\bar{\mathbf{h}}}(p)$, and $\mathbf{d}_p L_{\mathbf{f}(p)}^{4+2p} \bar{\bar{\mathbf{h}}}(p)$ from (35). From Theorem 1, since only zero to $2+2p$ order Lie derivatives of $\bar{\bar{\mathbf{h}}}(p)$ are required to determine the rank of $\bar{\bar{\mathbf{O}}}_p$ it can be concluded that these terms have the same dimensionality when considered together as the rows of a matrix. From (34) and (35) and using Theorem 1, it can be observed that since $\bar{\bar{\mathbf{O}}}_p$ is full rank, any more null columns in (35) can only be obtained from the column operations on columns $4+2p$ and $5+2p$. These are the last two columns of $\bar{\bar{\mathbf{O}}}_{p+1}$ corresponding to the addition of the landmark $p+1$.

Hence from (31) and using the same logic used in proving Result 1, it can be shown that columns $4+2p$ and $5+2p$ of $\bar{\bar{\mathbf{O}}}_{p+1}$ are independent and hence cannot be made null. Therefore, $\bar{\bar{\mathbf{O}}}_{p+1}$ is full rank. Thus, the SLAM problem becomes locally weakly observable when observing two known landmarks and all the estimated landmarks.

Consider now the situation of observing the vehicle's longitudinal and lateral coordinates. When n=1, it can be shown from (33) that $\bar{\bar{\mathbf{O}}}_1$ is full rank. Here it is assumed that $\bar{\mathbf{C}}_1^v(p)$ and $\bar{\mathbf{C}}_2^v(p)$ corresponds to observing the vehicle's longitudinal and lateral coordinates. Now, assuming $\bar{\bar{\mathbf{O}}}_p$ is full rank and using the equations (34)-(36) it follows that $\bar{\bar{\mathbf{O}}}_{p+1}$ is also full rank from the same logic that has been used in proving the full rank property of $\bar{\bar{\mathbf{O}}}_{p+1}$ when observing two known landmarks. This follows from the fact that once $\bar{\bar{\mathbf{O}}}_p$ is full rank $\bar{\bar{\mathbf{O}}}_{p+1}$ becomes full rank irrespective of the structure of the $\bar{\bar{\mathbf{h}}}(p)$.

Hence $\bar{\bar{\mathbf{O}}}_{p+1}$ is full rank. Thus, from the principle of mathematical induction the SLAM problem becomes locally weakly observable if the vehicle's longitudinal and lateral coordinates and all the estimated landmarks are observed.

IV. SLAM Problem in Information Form

A. Information Filter for State Estimation

If it is assumed that both the process and observation noises of the SLAM problem are Gaussian, the inverse of the state error covariance matrix is the Fisher information matrix. According to [9] the state update equation of the Information form can be expressed as follows.

$$\mathbf{P}^{-1}(t \mid t) = \mathbf{P}^{-1}(t \mid t^-) + \mathbf{I}(t) \tag{37}$$

$$\mathbf{I}(t) = \mathbf{H}^T(t)\mathbf{R}^{-1}(t)\mathbf{H}(t) \tag{38}$$

where $\mathbf{P}(t \mid t)$ is the updated error covariance matrix of $\mathbf{X}_n(t)$, $\mathbf{P}(t \mid t^-)$ is the predicted error covariance matrix of $\mathbf{X}_n(t^-)$, $\mathbf{H}(t)$ is the Jacobian of the measurement model and $\mathbf{R}(t)$ is the measurement noise covariance matrix. $\mathbf{I}(t)$ in (38) is known as the Information Matrix Associated with Observations (IMAO). $\mathbf{P}^{-1}(t \mid t^-)$ can be obtained from the following equation of the Kalman filter prediction.

$$\mathbf{P}(t \mid t^-) = \mathbf{F}(t)\mathbf{P}(t^- \mid t^-)\mathbf{F}^T(t) + \mathbf{Q}(t) \tag{39}$$

where $\mathbf{P}(t^- \mid t^-)$ is the error covariance matrix of $\mathbf{X}_n(t^-)$, $\mathbf{F}(t)$ is the Jacobian of the process model and $\mathbf{Q}(t)$ is the process noise covariance matrix. The information state vector is denoted by $\mathbf{y}(t \mid t) = \mathbf{P}^{-1}(t \mid t)\mathbf{X}_n(t \mid t)$ and $\mathbf{P}^{-1}(t \mid t)$ is the information matrix of the state vector $\mathbf{X}_n(t \mid t)$. A complete derivation of the information form of the Kalman filter can be found in [9].

B. Properties of the Information Matrix Associated with Observations

The prediction (37) of the covariance matrix increases uncertainty. If we start to estimate from an infinitely uncertain system $\mathbf{P}(t\,|\,t^-)$ will also be infinite. Thus, in (37) $\mathbf{P}^{-1}(t\,|\,t^-)$ should be zero. Therefore, in order to have a finite covariance $\mathbf{P}^{-1}(t\,|\,t)$ from the measurement updating we should be able to invert the $\mathbf{I}(t)$. This suggests that when $\mathbf{I}(t)$ is singular it cannot give any information about at least some of the SLAM states.

Consider the proofs of Results 1-2. These 2 results are about the column space of the nonlinear observability matrix given by $\mathbf{O}_n$ and $\bar{\bar{\mathbf{O}}}_n$. The column spaces of $\mathbf{O}_n$ and $\bar{\bar{\mathbf{O}}}_n$ comprise the Jacobians of the rows from zero to appropriate order of Lie derivatives of $\mathbf{h}(n)$ and $\bar{\bar{\mathbf{h}}}(n)$ respectively. Since the Jacobians of zero order Lie derivatives are a subspace of the column space of $\mathbf{O}_n$ and $\bar{\bar{\mathbf{O}}}_n$ it follows that Results 1 and 2 are true for the Jacobians of the Lie derivatives of order zero of $\mathbf{h}(n)$ and $\bar{\bar{\mathbf{h}}}(n)$ respectively.

Since $\mathbf{R}(t)$ is a positive definite matrix $\mathbf{R}^{-1}(t)$ is also a positive definite matrix by Property 1 of the Appendix.

Also since $\mathbf{H}^T(t)\mathbf{R}^{-1}(t)\mathbf{H}(t)$ is Hermitian, rank and null space of $\mathbf{H}^T(t)\mathbf{R}^{-1}(t)\mathbf{H}(t)$ is similar to the rank and the null space of $\mathbf{H}(t)$ from Property 2 of the Appendix. However, by the definition of the nonlinear observability matrix and the Lie derivatives $\mathbf{H}(t)$ is identical to the Jacobians of the Lie derivatives of order zero of the corresponding measurement models. Hence the Results 1 and 2 proved in Section III are true for $\mathbf{H}(t)$ and therefore for $\mathbf{H}^T(t)\mathbf{R}^{-1}(t)\mathbf{H}(t)$. Hence using Results 1 and 2 and the structure of $\mathbf{I}(t)=\mathbf{H}^T(t)\mathbf{R}^{-1}(t)\mathbf{H}(t)$ it follows that the following properties are true for the Information Matrix Associated with Observations.

Property 1: IMAO of the SLAM problem estimating n unknown landmarks when only observing all the estimated landmarks is rank deficient by 3.

Property 2: IMAO of the SLAM problem is full rank when (1) Observing two known landmarks or when (2) Observing vehicle's longitudinal and lateral coordinates provided it observes all the estimated landmarks.

From these 2 properties of the IMAO we can predict the behavior of the SLAM problem when it is initialized with infinite uncertainty. It follows that the SLAM problem initialized with infinite uncertainty (a form of the kidnapped robot problem) can reduce the state uncertainty and consistently estimate the SLAM state vector and its uncertainty using the Information filter if the vehicle observes all the estimated landmarks and at least two a priori known landmarks.

V. Simulations and Experiments

A. Simulation Setup

This section presents results of simulations to evaluate the advantages and disadvantages of ensuring nonlinear observability of the SLAM problem. A 2D simulation environment of 100m square is shown in Fig. 1. Within this environment a car like mobile robot followed a specified trajectory while observing point landmarks in the environment using a range and bearing sensor. An extended Kalman filter based approach to SLAM [11] was used to compare the performance of SLAM in the context of Results 1 and 2 and the properties of the IMAO proved above. A nearest neighbor data association method [2] and a map management method [11] were also used in the simulation of SLAM. In the simulations, the robot observed all the estimated landmarks and two *a priori* known landmarks. The estimated vehicle path and the map are consistent when the nonlinear observability conditions stipulated in Results 1 and 2 are satisfied (Fig. 1).

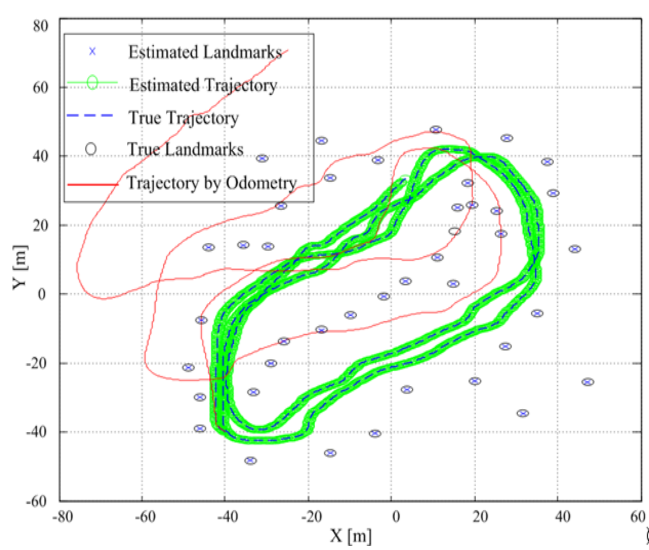


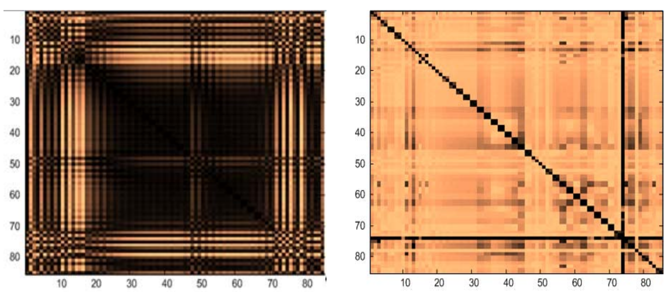
Fig. 1 SLAM simulation with nonlinear observability.

B. Correlations and observability based map management

It has been observed in the simulations that the covariance and the information matrices of the SLAM problem have the structure given by Fig. 2 when the vehicle observes only the estimated landmarks. Here dark colors represent higher values both in covariance and information. The structure of the covariance matrix and the information matrix is given by Fig. 3 when the nonlinear observability is ensured. These figures show that when only the estimated landmarks are observed the covariance matrix is highly dense and information matrix is sparse. It is interesting to note that when SLAM is locally weakly observable, the covariance matrix is lightest in color. That means the correlations among landmarks and the vehicle are at their smallest values when the SLAM problem is locally weakly observable.

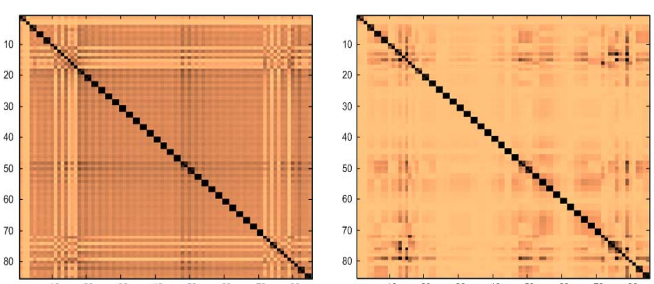
This suggests that when the SLAM is locally weakly observable we can use simplified data association schemes such as nearest neighbor data association and also reduce the number of map states significantly from the estimation as long as the full observability properties of SLAM are maintained. As a consequence, it is possible to implement a map management strategy which takes care of the observability of the SLAM problem. This is because the vehicle and map states of the locally weakly observable SLAM are not highly correlated as in the standard SLAM which only observes estimated landmarks. When any estimated landmark is not observed it can be removed from the state vector and can be stored for future use. In this way it is also possible to keep the size of the map constant or below a certain limit determined by the available computing resources. The key concept is that as long as we maintain the local weak observability property of the SLAM state, we can remove or add any landmark states to the SLAM state vector. Therefore, such a map management strategy together with simple data association methods will be a great benefit for computationally feasible SLAM implementations.

Fig. 4 shows results of a SLAM algorithm initialized with very large errors in longitudinal and lateral coordinates in the same simulation environment. This algorithm consistently estimates the localization information when the SLAM is locally weakly observable, as was shown in Section IV. Initialization with errors of this magnitude is simply not possible in standard unobservable SLAM.



(a) (b)

Fig. 2 Covariance matrix (a) and Information matrix (b) when SLAM is observing estimated landmarks only.



(a) (b)

Fig. 3 Covariance matrix (a) and Information matrix (b) when SLAM is done observing estimated landmarks and two known landmarks.

C. Experiments

Experiments are performed with the car park dataset of the University of Sydney. The car park dataset was obtained by driving a utility vehicle [12] equipped with GPS, wheel and steering encoders and a laser range finder. It was used to check the consistency of the localization error estimates when SLAM is made locally weakly observable by observing at least 2 known landmarks and performing map management according to observability as detailed in Section V. Fig. 5 shows the estimated vehicle trajectory obtained and the map of estimated landmarks. It can be observed that the estimated vehicle path and the landmarks are consistent with the true vehicle path and the landmark locations.

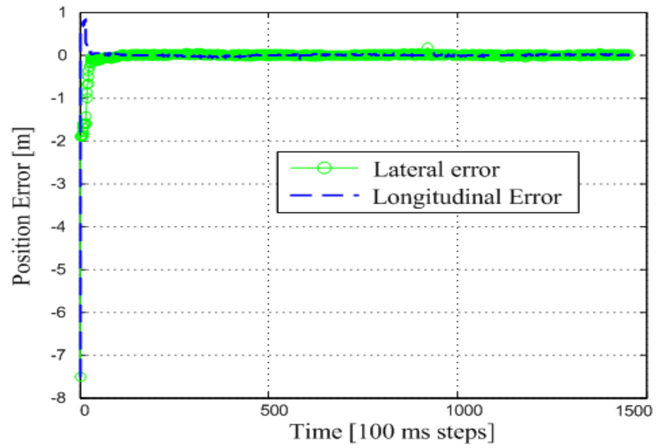


Fig. 4 SLAM initialization with large errors.

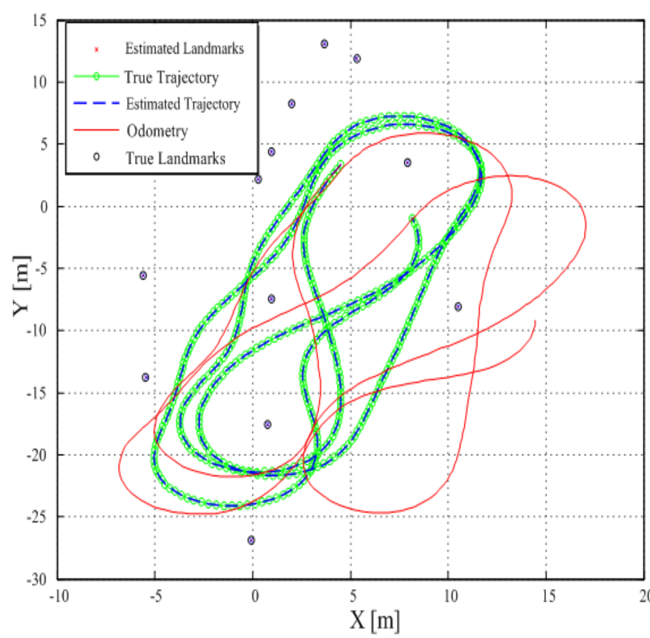


Fig. 5 SLAM results from the car park data set.

VI. Conclusions

The work described in the paper gives a useful insight into the nonlinear properties of the SLAM problem both rigorously and intuitively. It is shown that for the local weak observability of the SLAM problem estimating any number of landmarks, the vehicle must observe all estimated landmarks and at least two known landmarks or the vehicle's longitudinal and lateral coordinates. The properties of the nonlinear observability matrix are vital in understanding the nonlinear observability of SLAM in greater detail. They also indicate future directions of research for designing efficient nonlinear observers for SLAM, localization and mapping.

To the best of our knowledge we have demonstrated the relationship between the nonlinear observability of the SLAM and the kidnapped robot problem for the first time. It is shown here with the aid of the information form of the SLAM formulation that the IMAO has many properties of the nonlinear observability matrix. It has been shown that if the conditions for the nonlinear observability of SLAM are satisfied you can initialize SLAM even with very large uncertainties.

It has also been shown using simulations that the correlations in the SLAM state vector are at a minimum when the nonlinear observability conditions are fully satisfied. Another notable proposal which has been verified both using simulations and experiments is the observability based map management strategy. It has been argued that once the observability of the SLAM state is ensured there is no requirement to maintain a large map state vector and map vehicle correlations in a huge matrix in the state estimation. Therefore, it is interesting to note that if you maintain the full nonlinear observability of the SLAM, the prediction and update of SLAM can be done in almost constant time complexity by utilizing an observability based map management strategy. Such a strategy enables the addition or removal of map states based on the observability of the SLAM state vector.

Finally it is hoped that the nonlinear observability, its relationships with the information form of the SLAM problem and apparent decorrelation of map vehicle states will provide new directions for efficient and improved deployment of SLAM algorithms in several domains.

APPENDIX

***Properties of positive definite matrices* (from [10]).**

1. Every positive definite matrix is invertible and its inverse is also positive definite.
2. Let $\mathbf{A} = ()_{n\times n}$ be positive definite. If $\mathbf{C} = ()_{n\times m}$, then $\mathbf{C}^*\mathbf{AC}$ is positive semi-definite. Furthermore,

$$rank(\mathbf{C}^*\mathbf{AC}) = rank(\mathbf{C}) \tag{40}$$

and

$$null(\mathbf{C}^*\mathbf{AC}) = null(\mathbf{C}) \tag{41}$$

where $\mathbf{C}^*$ is the conjugate transpose of the matrix $\mathbf{C}$.

REFERENCES

[1] R. Smith, M. Self, and P. Cheeseman, “A stochastic map for uncertain spatial relationships”., *Fourth International Symposium of Robotics Research*, pp. 467–474, 1987.

[2] M.W.M.G. Dissanayake, P. Newman, S. Cleark, H.F. Durrant-Whyte and M. Csorba, “A Solution to the Simultaneous Localization and Map Building (SLAM) Problem”, *IEEE Transactions on Robotics and Automation*, Vol 17, No 3, pp 229-241, June 2001.

[3] J. Andrade-Cetto and A. Sanfeliu, “The Effects of Partial Observability When Building Fully Correlated Maps”, *IEEE Trans. on Robotics*. 21(4): pp. 771-777, August 2005.

[4] J. Kim and S. Sukkarieh Autonomous Airborne Navigation in Unknown Terrain Environments. IEEE Transactions on Aerospace and Electronic Systems. 40(3):1031-1045, July 2004.

[5] M. Bryson and S. Sukkarieh Observability Analysis and Active Control for Airborne SLAM. IEEE Transactions on Aerospace and Electronic Systems. 44(1):261-280, January 2008.

[6] K. Lee, W S. Wijesoma and J. I. Guzman “On the Observability and Observability Analysis of SLAM.” Proc. IEEE/RSJ International Conf. on Intelligent Robots and Systems, pages 3569-3574, October 2006.

[7] R.S.A. Martinelli, “Observability analysis for mobile robot localization” in Proc. IEEE/RSJ Int. Conf. Intelligent Robots and Systems, Edmonton, August 2005.

[8] R. Hermann and A.J. Krener, “Nonlinear observability and controllability” IEEE Transactions on Automatic Control. 22(5):728-740., October 1977.

[9] H. Durrant-Whyte, “Multi Sensor Data Fusion”, Lecture notes, Australian Centre for Field Robotics, The University of Sydney, NSW 2006, Australia.

[10] R. A. Horn, C.R. Johnson, “Matrix Analysis”, Cambridge University Press, 1985.

[11] W.S. Wijesoma, L.D.L. Perera and M.D. Adams, “Toward Multi-dimensional Assignment Based Data Association for Robot Localization and Mapping”, IEEE Transactions on Robotics, vol. 22, no. 2, pp. 350-365 April, 2006.

[12] Guivant, J.; Nebot, E., “Implementation of simultaneous navigation and mapping in large outdoor environments.”, In Robotics Research: The Tenth International Symposium, Springer, 2003.

[13] M. Anguelova. *Observability and identifiably of nonlinear systems with applications in biology*. Ph.D. thesis, Department of Mathematical Sciences, Division of Mathematics, Chalmers University of Technology and Göteborg University Göteborg, Sweden 2007.

[14] D. Goshen-Meskin I. Y. Bar-Itzhack. *Observability Analysis of Piece-Wise Constant Systems- Part 1:Theory.* IEEE Trans. on Aerospace and Electronic Systems. 28(4):1056-1068, October 1992.