A NOVEL DUAL ADAPTIVE NEURO-CONTROLLER BASED ON THE UNSCENTED TRANSFORM FOR MOBILE ROBOTS

Marvin K. Bugeja and Simon G. Fabri
Department of Systems and Control Engineering, University of Malta, Msida, MSD2080, Malta

Keywords: Dual adaptive control, Stochastic control, Neural networks, Unscented Kalman filter, Mobile robots.

Abstract: This paper proposes a novel dual adaptive neuro-control scheme based on the unscented transform for the dynamic control of nonholonomic wheeled mobile robots. The controller is developed in discrete time and the robot nonlinear dynamic functions are unknown to the controller. A multilayer perceptron neural network is used to approximate the nonlinear robot dynamics. The network is trained online via a specifically devised unscented Kalman predictor. In contrast to the majority of adaptive control techniques hitherto proposed in the literature, the controller presented in this paper does not rely on the heuristic certainty equivalence assumption, but accounts for the estimates' uncertainty via the principle of *dual adaptive* control. Moreover, the novel dual adaptive control law employs the unscented transform to improve on the first-order Taylor approximations inherent in previously published dual adaptive schemes. Realistic simulations, including comparative Monte Carlo tests, are used to illustrate the effectiveness of the proposed approach.

1 INTRODUCTION

Many publications on the control of nonholonomic wheeled mobile robots (WMRs) (Kanayama et al., 1990; Canudas de Wit et al., 1993) completely ignore the robot dynamics and rely on the assumption that the control inputs, usually motor voltages, instantaneously establish the desired wheel velocities. Others, explicitly account for the robot dynamics due to its mass, friction and inertia (Fierro and Lewis, 1995; Corradini and Orlando, 2001) show that dynamic control leads to an improvement in performance. However, as argued by Fierro and Lewis (1995), perfect knowledge of the robot dynamics is unavailable in practice. In addition, these parameters can also vary over time due to loading, wear and ground conditions. These issues inspired the development of several robust and adaptive WMR controllers over the last decade. These include: pre-trained artificial neural network (ANN) based controllers and robust sliding-mode methods (Corradini and Orlando, 2001), parametric adaptive schemes (Wang and Tsai, 2004), and functional adaptive controllers (Bugeja and Fabri, 2008).

Yet, all these adaptive controllers rely on the heuristic certainty equivalence (HCE) assumption. This means that the estimated functions are used by the controller as if they were the true ones, thereby ignoring completely the inherent uncertainty in the estimations. When the uncertainty is large, for instance during startup or when the functions are changing, HCE often leads to large tracking errors and excessive control actions, which can excite unmodelled dynamics or even lead to instability (Åström and Wittenmark, 1995). In contrast, the so-called dual adaptive controllers based on the *dual control* principle introduced by Fel'dbaum (1965), do not rely on the HCE assumption but account for the estimates' uncertainty in the control design. Specifically, a dual adaptive control law is designed with two aims in mind: (i) to ensure that the output tracks the reference signal with due consideration given to the estimates' uncertainty; (ii) to excite the plant input sufficiently so as to accelerate the estimation process, thereby reducing quickly the uncertainty in future estimates. These two features are known as *caution* and *probing* respectively (Åström and Wittenmark, 1995; Fabri and Kadirkamanathan, 2001).

Of the few dual adaptive controllers proposed in recent years, only our work presented in (Bugeja and Fabri, 2009) focuses on the dynamic control of WMRs. However, the multilayer perceptron (MLP) dual adaptive scheme employed in this work, not only uses the extended Kalman filter (EKF) (which inherently involves a first order approximation) as a neuro-estimator, but the control law itself is based on another

K. Bugeja M. and G. Fabri S. (2009).
A NOVEL DUAL ADAPTIVE NEURO-CONTROLLER BASED ON THE UNSCENTED TRANSFORM FOR MOBILE ROBOTS.
In *Proceedings of the International Joint Conference on Computational Intelligence*, pages 355-362
DOI: 10.5220/0002313503550362

first-order Taylor approximation of the measurement model. In contrast, the novelty of the control scheme presented in this paper comprises: the use of a specifically devised form of the unscented Kalman filter (UKF) (Julier and Uhlmann, 1997; Wan and van der Merwe, 2001) as a recursive weight tuning algorithm instead of the EKF; and more importantly, the development of a novel dual adaptive control law based on the unscented transform (UT) (Julier and Uhlmann, 1997), instead of the first-order Taylor approximation. Together, these novel developments lead to a significant improvement in performance over the EKF-based scheme in (Bugeja and Fabri, 2009). To the best of our knowledge, this is the first time that the UT is being used in the context of dual adaptive control.

The rest of the paper is organized as follows. Section 2 presents the dynamic model of the WMR. In Section 3 we present the novel UT-based dual adaptive dynamic control scheme. Simulation results, including those from a Monte Carlo comparative test, are presented in Section 4, which is followed by a brief conclusion in Section 5.

2 PLANT MODEL

Figure 1 depicts the differentially driven wheeled mobile robot considered in this paper. The following notation is adopted throughout the article:

P_o:	midpoint on the driving axle
P_c:	centre of mass without wheels
d:	distance from P_o to P_c
b:	distance from each wheel to P_o
r:	radius of each wheel
m_c:	mass of the platform without wheels
m_w:	mass of each wheel
I_c:	angular mass of the platform about P_c
I_w:	angular mass of wheel about its axle
I_m:	angular mass of wheel about its diameter

The robot state vector is given by $\boldsymbol{q} = [x \;\; y \;\; \phi \;\; \theta_r \;\; \theta_l]^T$, where (x,y) is the Cartesian coordinate of P_o, ϕ is the robot's orientation with reference to the xy frame, and θ_r, θ_l are the angular displacements of the right and left driving wheels respectively. The *pose* of the robot refers to the vector $\boldsymbol{p} = [x \;\; y \;\; \phi]$.

2.1 Kinematics

Assuming that the wheels roll without slipping, the kinematic model of this WMR, detailed by Bugeja and Fabri (2009), is given by:

$$\dot{\boldsymbol{q}} = \boldsymbol{S}(\boldsymbol{q})\boldsymbol{\nu}, \tag{1}$$

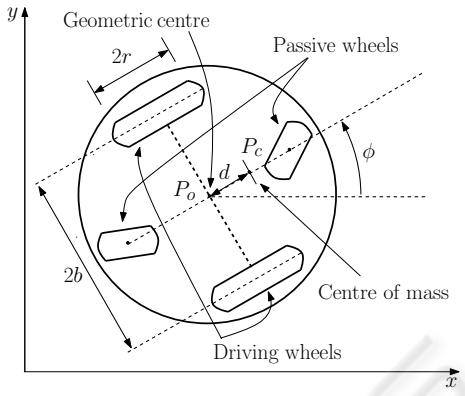

Figure 1: Differentially driven wheeled mobile robot.

where the velocity $\boldsymbol{\nu} \triangleq [\nu_r \;\; \nu_l]^T = \left[\dot{\theta}_r \;\; \dot{\theta}_l\right]^T$ and

$$\boldsymbol{S} = \begin{bmatrix} \frac{r}{2}\cos\phi & \frac{r}{2}\cos\phi \\ \frac{r}{2}\sin\phi & \frac{r}{2}\sin\phi \\ \frac{r}{2b} & -\frac{r}{2b} \\ 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

2.2 Dynamics

The WMR dynamic model, also detailed in (Bugeja and Fabri, 2009), is given by:

$$\bar{\boldsymbol{M}}\dot{\boldsymbol{\nu}} + \bar{\boldsymbol{V}}(\dot{\boldsymbol{q}})\boldsymbol{\nu} + \bar{\boldsymbol{F}}(\dot{\boldsymbol{q}}) = \boldsymbol{\tau}, \tag{2}$$

where:

$$\bar{\boldsymbol{M}} = \begin{bmatrix} \frac{r^2}{4b^2}(mb^2+I) + I_w & \frac{r^2}{4b^2}(mb^2-I) \\ \frac{r^2}{4b^2}(mb^2-I) & \frac{r^2}{4b^2}(mb^2+I) + I_w \end{bmatrix},$$

$$\bar{\boldsymbol{V}}(\dot{\boldsymbol{q}}) = \begin{bmatrix} 0 & \frac{m_c r^2 d\dot{\phi}}{2b} \\ \frac{m_c r^2 d\dot{\phi}}{2b} & 0 \end{bmatrix}, \;\; \bar{\boldsymbol{F}}(\dot{\boldsymbol{q}}) = \boldsymbol{S}^T(\boldsymbol{q})\boldsymbol{F}(\dot{\boldsymbol{q}}),$$

$I = (I_c + m_c d^2) + 2(I_m + m_w b^2)$, $\quad m = m_c + 2m_w$, $\boldsymbol{F}(\dot{\boldsymbol{q}})$ is a vector of frictional forces, and $\boldsymbol{\tau} = [\tau_r \;\; \tau_l]^T$ with τ_r and τ_l being the torques applied to the right and left wheel respectively.

To account for the fact that the controller is implemented on a digital computer, the continuous-time dynamics (2) are discretized through a first order forward Euler approximation with a sampling interval of T seconds, resulting in

$$\boldsymbol{\nu}_k - \boldsymbol{\nu}_{k-1} = \boldsymbol{f}_{k-1} + \boldsymbol{G}_{k-1}\boldsymbol{\tau}_{k-1}, \tag{3}$$

where subscript k denotes that the corresponding variable is evaluated at kT seconds, and vector $\boldsymbol{f}_{k-1}$ and matrix $\boldsymbol{G}_{k-1}$, which together encapsulate the WMR dynamics, are given by

$$\begin{aligned} \boldsymbol{f}_{k-1} &= -T\bar{\boldsymbol{M}}_{k-1}^{-1}\left(\bar{\boldsymbol{V}}_{k-1}\boldsymbol{\nu}_{k-1} + \bar{\boldsymbol{F}}_{k-1}\right), \\ \boldsymbol{G}_{k-1} &= T\bar{\boldsymbol{M}}_{k-1}^{-1}. \end{aligned} \tag{4}$$

The following condition is assumed.

Assumption 2.1. *The control input vector $\boldsymbol{\tau}$ remains constant over a sampling interval of T seconds, which is chosen low enough for the Euler approximation error to be negligible.*

3 CONTROL DESIGN

The trajectory tracking task of a nonholonomic WMR is chosen as a test problem in this paper. In trajectory tracking the robot is required to track a nonstationary kinematically identical *virtual* vehicle, in both pose and velocity at all times, by minimizing the tracking error vector $\boldsymbol{e}_k$ (Kanayama et al., 1990) defined as

$$\boldsymbol{e}_k = \begin{bmatrix} \cos\phi_k & \sin\phi_k & 0 \\ -\sin\phi_k & \cos\phi_k & 0 \\ 0 & 0 & 1 \end{bmatrix} (\boldsymbol{p}_{\boldsymbol{r}k} - \boldsymbol{p}_k), \quad (5)$$

where $\boldsymbol{p}_{\boldsymbol{r}k} = [x_{rk} \;\; y_{rk} \;\; \phi_{rk}]^T$ denotes the virtual vehicle pose vector. Hence, the kinematic control task is to make $\boldsymbol{e}$ converge to zero so that $\boldsymbol{p}$ converges to $\boldsymbol{p}_{\boldsymbol{r}}$.

3.1 Kinematic Control

To address the trajectory tracking problem we employ a discrete-time version of the well-established trajectory tracking controller originally proposed in (Kanayama et al., 1990), given by

$$\boldsymbol{\nu}_{\boldsymbol{c}k} = \begin{bmatrix} \frac{1}{r} & \frac{b}{r} \\ \frac{1}{r} & -\frac{b}{r} \end{bmatrix} \begin{bmatrix} v_{rk}\cos e_{3k} + k_1 e_{1k} \\ \omega_{rk} + k_2 v_{rk} e_{2k} + k_3 v_{rk} \sin e_{3k} \end{bmatrix},$$

where $\boldsymbol{\nu}_{\boldsymbol{c}k}$ is the wheel velocity command vector computed by the kinematic controller, k_1, k_2, and k_3 are *positive* design parameters, v_{rk} and $\boldsymbol{\omega}_{\boldsymbol{r}k}$ are the translational and angular reference velocities respectively corresponding to the desired trajectory, and e_{1k}, e_{2k}, e_{3k} are the elements of $\boldsymbol{e}_k$ in (5).
If one assumes *perfect velocity tracking* (i.e. $\boldsymbol{\nu}_k = \boldsymbol{\nu}_{\boldsymbol{c}k} \;\; \forall \;\; k$), hence ignoring the WMR dynamics expressed in (2), then this kinematic control law alone solves the trajectory tracking problem. However, as pointed out earlier, mere kinematic control rarely suffices and often leads to substantial degradation in performance (Fierro and Lewis, 1995).

3.2 UT-based Dual Adaptive Control

If the nonlinear dynamic functions $\boldsymbol{f}_k$ and $\boldsymbol{G}_k$ are assumed to be *perfectly* known, a simple feedback linearizing control law, like the one detailed in (Bugeja and Fabri, 2009), solves the *dynamic control* problem (i.e. assuring that $\boldsymbol{\nu}_k$ tracks $\boldsymbol{\nu}_{\boldsymbol{c}k} \forall k$). However, it is an undeniable fact that in practice the robot dynamics; dependent on mass, inertia, friction and possibly several unmodelled phenomena; are typically unknown and may even change over time. In addition perfect sensor measurements are never available.

To address these complex issues of uncertainty, we propose a novel *dual adaptive* controller employing a MLP ANN trained online via an UKF algorithm in prediction mode. In contrast to the hitherto proposed innovation-based suboptimal dual adaptive laws (Fabri and Kadirkamanathan, 2001; Bugeja and Fabri, 2009), the control law we propose here employs the UT to approximate better the mean and covariance terms arising in the chosen cost function. Hence, the envisaged improvement is not solely due to the superior stochastic estimator employed to train the ANN (the UKF instead of the EKF), but also due to the dual adaptive law itself, as clarified further in the following sections.

3.2.1 Neuro-Stochastic Estimator

To deal with the uncertainty and/or time-varying nature of the dynamic functions $\boldsymbol{f}_k$ and $\boldsymbol{G}_k$, we opt to assume that they are completely unknown to the controller and employ a stochastically trained ANN algorithm for their approximation in real-time.

A sigmoidal MLP ANN with one hidden layer is used to approximate the nonlinear vector $\boldsymbol{f}_{k-1}$, as depicted in Figure 2. Its output is given by

$$\tilde{\boldsymbol{f}}_{k-1} = \begin{bmatrix} \boldsymbol{\phi}^T(\boldsymbol{x}_{k-1}, \hat{\boldsymbol{a}}_k)\hat{\boldsymbol{w}}_{\boldsymbol{1}k} \\ \boldsymbol{\phi}^T(\boldsymbol{x}_{k-1}, \hat{\boldsymbol{a}}_k)\hat{\boldsymbol{w}}_{\boldsymbol{2}k} \end{bmatrix}, \quad (6)$$

in the light of the following statements:

Definition 3.1. *$\boldsymbol{x}_{k-1} = [\boldsymbol{\nu}_{k-1} \;\; 1]$ denotes the ANN input. The augmented constant serves as a bias input.*

Definition 3.2. *$\boldsymbol{\phi}(\cdot,\cdot)$ is the vector of sigmoidal activation functions, whose i^{th} element is given by $\phi_i = 1/\left(1+\exp\left(-\hat{\boldsymbol{s}}_{\boldsymbol{i}}^T\boldsymbol{x}\right)\right)$, where $\hat{\boldsymbol{s}}_{\boldsymbol{i}}$ is i^{th} vector element in the group vector $\hat{\boldsymbol{a}}$; i.e. $\hat{\boldsymbol{a}} = \left[\hat{\boldsymbol{s}}_{\boldsymbol{1}}^T \;\; \cdots \;\; \hat{\boldsymbol{s}}_{\boldsymbol{L}}^T\right]^T$ where L denotes the number of neurons. The time index has been dropped for clarity, and throughout the paper the $\hat{}$ notation indicates that the operand is undergoing estimation.*

Definition 3.3. *$\hat{\boldsymbol{w}}_{\boldsymbol{i}k}$ represents the synaptic weight estimate vector of the connection between the neuron hidden layer and the i^{th} output element of the ANN.*

Assumption 3.1. *The input vector $\boldsymbol{x}_{k-1}$ is contained within a known, arbitrarily large compact set $\boldsymbol{\chi} \subset \mathbb{R}^2$.*

It is known that $\boldsymbol{G}_{k-1}$ is a state-independent matrix with unknown elements (refer to (4)). Hence, its estimation does not require the use of an ANN. Moreover it is a symmetric matrix, a property which is ex-

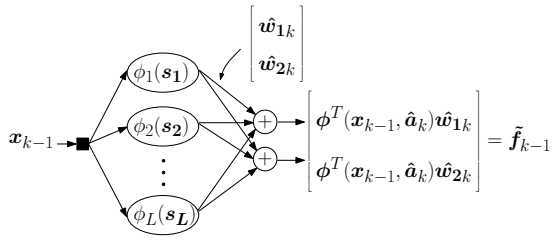

Figure 2: Sigmoidal Multilayer Perceptron neural network.

ploited to construct its estimate as follows

$$\tilde{\boldsymbol{G}}_{k-1} = \begin{bmatrix} \hat{g}_{1k-1} & \hat{g}_{2k-1} \\ \hat{g}_{2k-1} & \hat{g}_{1k-1} \end{bmatrix}, \tag{7}$$

where $\hat{g}_{1k-1}$ and $\hat{g}_{2k-1}$ represent the estimates of the unknown elements in $\boldsymbol{G}_{k-1}$.

The ANN online weight-tuning algorithm is developed next. The following formulation is required in order to proceed.

Definition 3.4. *The unknown parameters requiring estimation are grouped in a single vector* $\hat{\boldsymbol{z}}_k = \begin{bmatrix} \hat{\boldsymbol{r}}_k^T & \hat{\boldsymbol{g}}_k^T \end{bmatrix}^T$, *where* $\hat{\boldsymbol{r}}_k = \begin{bmatrix} \hat{\boldsymbol{w}}_{\mathbf{1}k}^T & \hat{\boldsymbol{w}}_{\mathbf{2}k}^T & \hat{\boldsymbol{a}}_k^T \end{bmatrix}^T$ *and* $\hat{\boldsymbol{g}}_k = \begin{bmatrix} \hat{g}_{1k-1} & \hat{g}_{2k-1} \end{bmatrix}^T$.

Definition 3.5. *The measured output in the dynamic model (3) is denoted by* $\boldsymbol{y}_k = \boldsymbol{\nu}_k - \boldsymbol{\nu}_{k-1}$.

Assumption 3.2. *By the* Universal Approximation Theorem *of ANN, inside the compact set* $\boldsymbol{\chi}$*, the ANN approximation error is negligibly small when the estimate* $\hat{\boldsymbol{r}}_k$ *is equal to some unknown optimal vector denoted by* $\boldsymbol{r}_k^*$. *The* * *notation denotes optimality.*

In view of the stochastic adaptive approach taken in this work, the unknown optimal parameter vector $\boldsymbol{z}_k^*$ is treated as a random variable, with the initial condition $p(\boldsymbol{z}_0^*) \sim \mathcal{N}(\hat{\boldsymbol{z}}_0, \boldsymbol{P}_0)$, where the covariance $\boldsymbol{P}_0$ reflects the confidence in the initial guess $\hat{\boldsymbol{z}}_0$. Moreover, $\boldsymbol{z}_k^*$ is characterized as a stationary process corrupted by an artificial process noise $\boldsymbol{\rho}_k$, which aids convergence and tracking during estimation. In addition, observation uncertainty is catered for by augmenting a random measurement noise $\boldsymbol{\epsilon}_k$ to $\boldsymbol{y}_k$.

By (6), (7), all previous definitions and assumptions; it follows that the model in (3) can be represented in the following stochastic state-space form

$$\begin{aligned} \boldsymbol{z}_{k+1}^* &= \boldsymbol{z}_k^* + \boldsymbol{\rho}_k \\ \boldsymbol{y}_k &= \boldsymbol{h}\left(\boldsymbol{x}_{k-1}, \boldsymbol{\tau}_{k-1}, \boldsymbol{z}_k^*\right) + \boldsymbol{\epsilon}_k, \end{aligned} \tag{8}$$

where the vector function $\boldsymbol{h}\left(\boldsymbol{x}_{k-1}, \boldsymbol{\tau}_{k-1}, \boldsymbol{z}_k^*\right)$ is nonlinear in the unknown optimal parameter vector $\boldsymbol{z}_k^*$, and is given by

$$\boldsymbol{h}(\cdot) = \tilde{\boldsymbol{f}}(\boldsymbol{x}_{k-1}, \boldsymbol{r}_k^*) + \tilde{\boldsymbol{G}}(\boldsymbol{g}_k^*)\boldsymbol{\tau}_{k-1}. \tag{9}$$

Since the resulting measurement model (8) is nonlinear (due to the MLP network), in a stochastic framework one has to a employ a nonlinear estimator, conventionally the suboptimal EKF, to train the ANN in real-time. However as shown in (Wan and van der Merwe, 2001), the UKF can be a better alternative for stochastic nonlinear estimation. Its benefits over the EKF include, a derivative-free algorithm and superior accuracy in its approximations. For this reason, we opt to employ the UKF in predictive mode for the estimation of $\boldsymbol{z}_{k+1}^*$, as detailed below and in Lemma 3.1.

Definition 3.6. *The* information state *denoted by* I^k*, consists of all output measurements up to instant k and all the previous inputs;* Y^k *and* U^{k-1} *respectively.*

Assumption 3.3. $\boldsymbol{\epsilon}_k$ *and* $\boldsymbol{\rho}_k$ *are both zero-mean white Gaussian processes with covariances* $\boldsymbol{R}_{\boldsymbol{\epsilon}}$ *and* $\boldsymbol{Q}_{\boldsymbol{\rho}}$ *respectively. Moreover,* $\boldsymbol{\epsilon}_k$, $\boldsymbol{\rho}_k$ *and* $\boldsymbol{z}_0^*$ *are mutually independent* $\forall k$.

We propose the use of an unscented Kalman predictor, detailed in Algorithm **??**, to generate recursively estimates for the mean and covariance of $\boldsymbol{z}_{k+1}^*$ conditioned on I^k, denoted respectively by $\hat{\boldsymbol{z}}_{k+1}$ and $\boldsymbol{P}_{k+1}$. This leads to the following lemma:

Lemma 3.1. *In the light of (8), Definition 3.6, and Assumption 3.3, it follows that* $p(\boldsymbol{z}_{k+1}^*|I^k) \approx \mathcal{N}(\hat{\boldsymbol{z}}_{k+1}, \boldsymbol{P}_{k+1})$*, and so* $\hat{\boldsymbol{z}}_{k+1}$ *is considered to be the estimate of* $\boldsymbol{z}_{k+1}^*$ *conditioned on* I^k.

Proof. The proof follows directly by applying a predictive type UKF (additive noise version) on the nonlinear stochastic state-space model in (8). The predictive UKF is effectively the standard UKF algorithm as presented in (Wan and van der Merwe, 2001) for parameter estimation, with the difference that the measurement-update step precedes that for time-update. In addition, the time-update step is advanced by one sample to obtain $\hat{\boldsymbol{z}}_{k+1|k}$ at instant k. □

Lemma 3.2. *On the basis of Lemma 3.1, it follows that* $p(\boldsymbol{y}_{k+1}|I^k)$ *is approximately Gaussian with mean* $\hat{\boldsymbol{y}}_{k+1}$ *and covariance* $\boldsymbol{P}_{\boldsymbol{y}\boldsymbol{y}_{k+1}}$ *given by:*

$$\hat{\boldsymbol{y}}_{k+1} = \hat{\boldsymbol{f}}_k + \hat{\boldsymbol{G}}_k\boldsymbol{\tau}_k, \tag{10}$$

$$\textit{where, } \hat{\boldsymbol{f}}_k = \sum_{i=0}^{2N} W_{mi}\mathbb{F}_{i,k+1|k}, \quad \hat{\boldsymbol{G}}_k = \tilde{\boldsymbol{G}}(\hat{\boldsymbol{g}}_{k+1}) \tag{11}$$

and the covariance

$$\boldsymbol{P}_{\boldsymbol{y}\boldsymbol{y}_{k+1}} = \sum_{i=0}^{2N} W_{ci}\left[\boldsymbol{D}_{\boldsymbol{f}i} + \boldsymbol{D}_{\boldsymbol{G}i}\boldsymbol{\tau}_k\right]\left[\boldsymbol{D}_{\boldsymbol{f}i} + \boldsymbol{D}_{\boldsymbol{G}i}\boldsymbol{\tau}_k\right]^T + \boldsymbol{R}_{\boldsymbol{\epsilon}} \tag{12}$$

$$\textit{where, } \boldsymbol{D}_{\boldsymbol{f}i} = \mathbb{F}_{i,k+1|k} - \hat{\boldsymbol{f}}_k\,, \quad \boldsymbol{D}_{\boldsymbol{G}i} = \mathbb{G}_{i,k+1|k} - \hat{\boldsymbol{G}}_k.$$

Algorithm 3.1. The unscented Kalman predictor algorithm.

Given the previous prediction $(\hat{\boldsymbol{z}}_{k|k-1}, \boldsymbol{P}_{k|k-1})$, denoted in short-form by $(\hat{\boldsymbol{z}}_k, \boldsymbol{P}_k)$; the following algorithm generates the new prediction $(\hat{\boldsymbol{z}}_{k+1}, \boldsymbol{P}_{k+1})$:

Sigma-points sampling and propagation

$$\begin{aligned}
\mathcal{Z}_{k|k-1} &= \left[\hat{\boldsymbol{z}}_k \;\; \hat{\boldsymbol{z}}_k + \left(\gamma\sqrt{\boldsymbol{P}_k}\right) \;\; \hat{\boldsymbol{z}}_k - \left(\gamma\sqrt{\boldsymbol{P}_k}\right)\right] \\
\mathbb{F}_{k|k-1} &= \tilde{\boldsymbol{f}}(\boldsymbol{x}_{k-1}, \mathcal{R}_{k|k-1}),\ \mathbb{G}_{k|k-1} = \tilde{\boldsymbol{G}}(\mathcal{G}_{k|k-1}) \\
\mathbb{Y}_{k|k-1} &= \mathbb{F}_{k|k-1} + \mathbb{G}_{k|k-1}\boldsymbol{\tau}_{k-1} \\
\hat{\boldsymbol{y}}_k &= \sum_{i=0}^{2N} W_{mi}\mathbb{Y}_{i,k|k-1} \qquad (10)
\end{aligned}$$

Measurement update and estimate prediction

$$\begin{aligned}
\boldsymbol{P}_{\boldsymbol{y}\boldsymbol{y}_k} &= \sum_{i=0}^{2N} W_{ci}\left[\mathbb{Y}_{i,k|k-1} - \hat{\boldsymbol{y}}_k\right]\left[\mathbb{Y}_{i,k|k-1} - \hat{\boldsymbol{y}}_k\right]^T + \boldsymbol{R}_{\boldsymbol{\epsilon}} \\
\boldsymbol{P}_{\boldsymbol{z}\boldsymbol{y}_k} &= \sum_{i=0}^{2N} W_{ci}\left[\mathcal{Z}_{i,k|k-1} - \hat{\boldsymbol{z}}_k\right]\left[\mathbb{Y}_{i,k|k-1} - \hat{\boldsymbol{y}}_k\right]^T \\
\boldsymbol{K}_k &= \boldsymbol{P}_{\boldsymbol{z}\boldsymbol{y}_k}\boldsymbol{P}_{\boldsymbol{y}\boldsymbol{y}_k}^{-1}, \quad \boldsymbol{i}_k = \boldsymbol{y}_k - \hat{\boldsymbol{y}}_k \\
\hat{\boldsymbol{z}}_{k+1} &= \hat{\boldsymbol{z}}_k + \boldsymbol{K}_k\boldsymbol{i}_k \qquad (11) \\
\boldsymbol{P}_{k+1} &= \boldsymbol{P}_k - \boldsymbol{K}_k\boldsymbol{P}_{\boldsymbol{y}\boldsymbol{y}_k}\boldsymbol{K}_k^T + \boldsymbol{Q}_{\boldsymbol{\rho}} \qquad (12)
\end{aligned}$$

where: $\mathcal{Z}^T = \left[\mathcal{R}^T \;\; \mathcal{G}^T\right]^T$, $\gamma = \sqrt{N+\lambda}$, N is the length of $\hat{\boldsymbol{z}}_k$, the scaling parameter $\lambda = \alpha^2(N+\kappa) - N$, constant α determines the spread of the sigma-points, constant κ is a secondary scaling parameter, the UT weights are given by: $W_{m0} = \frac{\lambda}{N+\lambda}$, $W_{c0} = W_{m0} + 1 - \alpha^2 + \beta$, and $W_{mi} = W_{ci} = \frac{1}{2(N+\lambda)}$ $(i = 1, \ldots, 2N)$, and β includes prior knowledge of the estimate's distribution.
Moreover, in the UKF framework the linear algebra operation of adding a column vector to a matrix is defined as the addition of the vector to each column of the matrix. For further details, including guidelines for selecting the UKF scaling parameters, one is referred to (Wan and van der Merwe, 2001).

Proof. The proof for $\hat{\boldsymbol{f}}_k$ in (11) follows directly by applying the UT to estimate the mean of $p\left(\tilde{\boldsymbol{f}}(\boldsymbol{x}_k, \boldsymbol{r}_{k+1}^*)|I^k\right)$. The equation for $\hat{\boldsymbol{G}}_k$ in (11) is simply an application of the basic results in linear probability theory, i.e. $p(\boldsymbol{A}\boldsymbol{x}) = \boldsymbol{A}\bar{\boldsymbol{x}}$. It can be applied since $\tilde{\boldsymbol{G}}$ is linear in the parameters. To derive the equation of $\boldsymbol{P}_{\boldsymbol{y}\boldsymbol{y}_{k+1}}$ in (12) one needs to advance the equation for $\boldsymbol{P}_{\boldsymbol{y}\boldsymbol{y}_k}$ in Algorithm 3.1. by one sampling instant, and substitute for $\mathbb{Y}_{i,k+1|k}$ and $\hat{\boldsymbol{y}}_{k+1}$, using the relations leading to (10) in the same algorithm. □

3.2.2 UT-based Dual Adaptive Control Law

The stochastic formulation in Lemmas 3.1 and 3.2 constitutes the weight adaptation law for the proposed MLP dual adaptive scheme. In addition, it provides a real-time update of the density $p(\boldsymbol{y}_{k+1}|I^k)$. This information is crucial in dual control since unlike HCE schemes, dual adaptive controllers do not ignore the uncertainty of the estimates, but employ it in the development of the control law itself, as follows.

The explicit-type suboptimal innovation-based performance index J_{inn}, adopted from (Fabri and Kadirkamanathan, 2001), and modified to fit the multiple-input-multiple-output (MIMO) nonlinear scenario at hand is given by

$$\begin{aligned}
J_{inn} &= E\Big\{ \left(\boldsymbol{y}_{k+1} - \boldsymbol{y}_{\boldsymbol{d}_{k+1}}\right)^T \boldsymbol{Q}_{\boldsymbol{1}} \left(\boldsymbol{y}_{k+1} - \boldsymbol{y}_{\boldsymbol{d}_{k+1}}\right) \\
&\quad + \left(\boldsymbol{\tau}_k^T \boldsymbol{Q}_{\boldsymbol{2}} \boldsymbol{\tau}_k\right) + \left(\boldsymbol{i}_{k+1}^T \boldsymbol{Q}_{\boldsymbol{3}} \boldsymbol{i}_{k+1}\right) \Big| I^k \Big\}, \qquad (13)
\end{aligned}$$

where $E\left\{\cdot|I^k\right\}$ is the mathematical expectation conditioned on I^k, and the following definitions apply:

Definition 3.7. *$\boldsymbol{y}_{\boldsymbol{d}_{k+1}}$ is the reference vector of $\boldsymbol{y}_{k+1}$ and is given by $\boldsymbol{y}_{\boldsymbol{d}_{k+1}} = \boldsymbol{\nu}_{\boldsymbol{c}k+1} - \boldsymbol{\nu}_k$.*

To obtain $\boldsymbol{\nu}_{\boldsymbol{c}k+1}$ at instant k we advance the kinematic control law by one sampling interval as explained in (Bugeja and Fabri, 2009).

Definition 3.8. *Design parameters $\boldsymbol{Q}_{\boldsymbol{1}}$, $\boldsymbol{Q}_{\boldsymbol{2}}$ and $\boldsymbol{Q}_{\boldsymbol{3}}$ are diagonal and $\in \mathbb{R}^{2\times 2}$. Additionally: $\boldsymbol{Q}_{\boldsymbol{1}}$ is positive definite, $\boldsymbol{Q}_{\boldsymbol{2}}$ is positive semi-definite, and $-\boldsymbol{Q}_{\boldsymbol{1}} \leq \boldsymbol{Q}_{\boldsymbol{3}} \leq \boldsymbol{0}$ (element-wise).*

Remark 3.1. *The design parameter $\boldsymbol{Q}_{\boldsymbol{1}}$ is introduced to penalize tracking errors, $\boldsymbol{Q}_{\boldsymbol{2}}$ induces a penalty on large control inputs, and $\boldsymbol{Q}_{\boldsymbol{3}}$ affects the innovation vector so as to induce the dual adaptive feature characterizing this stochastic control law.*

The UT-based dual adaptive control law proposed in this paper is given in the following theorem.

Theorem 3.1. *The control law minimizing performance index J_{inn} in (13), subject to: the WMR dynamic model in (3), all previous definitions, assumptions and lemmas, is given by*

$$\begin{aligned}
\boldsymbol{\tau}_k &= \left(\hat{\boldsymbol{G}}_k^T \boldsymbol{Q}_{\boldsymbol{1}} \hat{\boldsymbol{G}}_k + \boldsymbol{Q}_{\boldsymbol{2}} + \boldsymbol{N}_{\boldsymbol{G}\boldsymbol{G}_{k+1}}\right)^{-1} \\
&\times \left(\hat{\boldsymbol{G}}_k^T \boldsymbol{Q}_{\boldsymbol{1}} \left(\boldsymbol{y}_{\boldsymbol{d}_{k+1}} - \hat{\boldsymbol{f}}_k\right) - \boldsymbol{n}_{\boldsymbol{G}\boldsymbol{f}_{k+1}}\right), \qquad (14)
\end{aligned}$$

$$\begin{aligned}
\textit{where, } \boldsymbol{N}_{\boldsymbol{G}\boldsymbol{G}_{k+1}} &= \sum_{i=0}^{2N} W_{ci} \boldsymbol{D}_{\boldsymbol{G}_i}^T \boldsymbol{Q}_{\boldsymbol{4}} \boldsymbol{D}_{\boldsymbol{G}i}, \\
\boldsymbol{n}_{\boldsymbol{G}\boldsymbol{f}_{k+1}} &= \sum_{i=0}^{2N} W_{ci} \boldsymbol{D}_{\boldsymbol{G}_i}^T \boldsymbol{Q}_{\boldsymbol{4}} \boldsymbol{D}_{\boldsymbol{f}_i}, \\
\boldsymbol{Q}_{\boldsymbol{4}} &= \boldsymbol{Q}_{\boldsymbol{1}} + \boldsymbol{Q}_{\boldsymbol{3}}.
\end{aligned}$$

Proof. Given the approximate Gaussian distribution of $p(\boldsymbol{y}_{k+1}|I^k)$ specified in Lemma 3.2, and several standard results from multivariate probability theory, it follows that cost function (13) can be rewritten as

$$\begin{aligned} J_{inn} &= (\hat{\boldsymbol{y}}_{k+1} - \boldsymbol{y}_{\boldsymbol{d}k+1})^T \boldsymbol{Q_1} (\hat{\boldsymbol{y}}_{k+1} - \boldsymbol{y}_{\boldsymbol{d}k+1}) \\ &+ \boldsymbol{\tau}_k^T \boldsymbol{Q_2} \boldsymbol{\tau}_k + \mathrm{tr}\left(\boldsymbol{Q_4} \boldsymbol{P}_{\boldsymbol{y}\boldsymbol{y}k+1}\right). \end{aligned} \quad (15)$$

By substituting for $\hat{\boldsymbol{y}}_{k+1}$ and $\boldsymbol{P}_{\boldsymbol{y}\boldsymbol{y}k+1}$ in (15), using the relations in (10) and (12) respectively, it is possible to factorize completely in terms of $\boldsymbol{\tau}_k$. The resulting expression is then differentiated with respect to $\boldsymbol{\tau}_k$ and equated to zero, in order to get the dual control law in (14). The resulting second order partial derivative of J_{inn} with respect to $\boldsymbol{\tau}_k$, the Hessian matrix, is given by $2\left(\hat{\boldsymbol{G}}_k^T \boldsymbol{Q_1} \hat{\boldsymbol{G}}_k + \boldsymbol{Q_2} + \boldsymbol{N}_{\boldsymbol{G}\boldsymbol{G}k+1}\right)$. By Definition 3.8 and (15), it is clear that the Hessian matrix is positive definite, meaning that $\boldsymbol{\tau}_k$ in (14) minimizes the dual performance index in (13) uniquely. Moreover, the latter implies that the inverse term in (14) exists without exceptions. □

Remark 3.2. *$\boldsymbol{Q_3}$ which appears in the control law via $\boldsymbol{Q}_4$ acts as a weighting factor, where at one extreme, with $\boldsymbol{Q_3} = -\boldsymbol{Q_1}$, the controller completely ignores the estimates' uncertainty, resulting in HCE control, and at the other extreme, with $\boldsymbol{Q_3} = \boldsymbol{0}$, it gives maximum attention to uncertainty, which leads to* cautious control. *For intermediate settings of $\boldsymbol{Q_3}$, the controller strikes a compromise and operates in dual adaptive mode. It is well known that HCE control leads to large tracking errors and excessive control actions when the estimates' uncertainty is relatively high. On the other hand, cautious control is notorious for sluggish response and control turn-off (Fabri and Kadirkamanathan, 2001). Consequently, dual control exhibits superior performance by striking a balance between the two extremes.*

4 SIMULATION RESULTS

This section presents a number of MATLAB® simulation results demonstrating the effectiveness of the UT-based dual adaptive control scheme proposed in this paper. Given the non-deterministic nature of the stochastic system in question, one cannot rely solely on a single simulation trial to validate the controller under test. Moreover, the analytical proof of strict convergence and stability for a dual adaptive controller for a nonlinear system, is still considered an open problem. For these reasons, a comprehensive Monte Carlo comparative analysis is also presented. This renders the performance evaluation process much more objective and reliable. In this analysis the proposed UT-based dual adaptive controller detailed in Section 3 is compared to the recently proposed EKF-based dual adaptive controller in (Bugeja and Fabri, 2009).

4.1 Simulation Scenario

The differential WMR under study was simulated using the *continuous-time* dynamic model detailed in (Bugeja and Fabri, 2009). To render the simulations more realistic, a number of model parameters, namely d, m_c, I_c and $\boldsymbol{F}(\dot{\boldsymbol{q}})$, were allowed to vary about a set of nominal values from one simulation trial to another. These variations adhere to the physics of realistic randomly generated scenarios, which represent various load configurations and surface frictional conditions. The nominal parameter values used for simulations correspond to those of *Neurobot*, the real WMR we presented in (Bugeja and Fabri, 2008), with a typical load. These are: $b = 22.95$cm, $r = 6.25$cm, $d = 10$cm, $m_c = 32$kg, $m_w = 1$kg, $I_c = 0.84$kgm^2, $I_w = 0.002$kgm^2, and $I_m = 0.005$kgm^2. Moreover, viscous friction was included in the model by setting $\boldsymbol{F}(\dot{\boldsymbol{q}}) = \boldsymbol{F_c}\dot{\boldsymbol{q}}$, where $\boldsymbol{F_c}$ is a diagonal matrix of coefficients, with nominal diagonal values set to $[2.6, 2.6, 0.35, 0.3, 0.3]$. The control sampling interval T was set to 50ms, and a zero-mean Gaussian measurement noise with covariance $10^{-4}\boldsymbol{I}$, where $\boldsymbol{I}$ denotes the identity matrix, was included.

Each simulation trial consists of eight consecutive simulations. The first six of these correspond to the three modes of operation; i.e. HCE ($\boldsymbol{Q_3} = -\boldsymbol{Q_1}$), cautious ($\boldsymbol{Q_3} = \boldsymbol{0}$) and dual ($\boldsymbol{Q_3} = -0.8\boldsymbol{Q_1}$); for each of the two adaptive schemes being compared. The remaining two trials correspond to: (1) a nominally tuned non-adaptive (NT-NA) controller, which represents a non-adaptive dynamic controller that assumes the model parameters to be equal to their nominal values. This is the best a non-adaptive controller can do when the exact robot parameters are unknown (very realistic); (2) a non-adaptive controller which is perfectly tuned (PT-NA) to the *exact values* of the model parameters. The latter is not realistic, and is used solely for the purpose of relative comparisons. In contrast, the HCE, cautious and dual adaptive controllers assume no preliminary information about the robot whatsoever, since closed-loop control is activated immediately with the initial parameter estimate vector $\hat{\boldsymbol{z}}_0$ selected at random from a zero-mean, Gaussian distribution with variance 0.0025.

For the sake of fair comparison the same noise sequence, reference trajectory, initial conditions, initial filter covariance matrix ($\boldsymbol{P}_0 = 0.1\boldsymbol{I}$), artificial process noise covariance ($\boldsymbol{Q}_{\boldsymbol{\rho}} = 10^{-6}\boldsymbol{I}$), tracking er-

ror penalty ($\boldsymbol{Q_1 = I_2}$), and control input penalty ($\boldsymbol{Q_2 = 0}$), are used in each simulation in a particular trial. In addition, the sigmoidal MLP ANN used in each of the two schemes under test contained five neurons ($L = 5 \Rightarrow N = 27$). Our experiments indicated that adding more neurons did not improve the control performance significantly. In the UT-based scheme, the UKF scaling parameters were set to $\alpha = 1, \kappa = 0$ and $\beta = 2$. The noise sequence is randomly generated afresh for each trial.

4.2 Single Trial Results

A number of simulation results, typifying the performance of the three control modes of the UT-based adaptive scheme are presented in Figure 3.

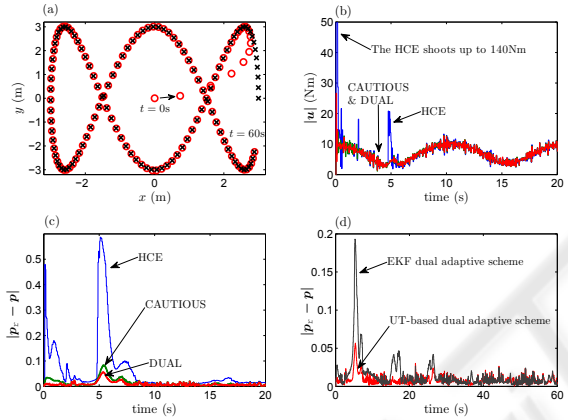

Figure 3: (a): reference ($\times$) and actual ($\bigcirc$) trajectories (UT-based dual); (b): control input (UT-based 3 modes); (c): transient tracking error (UT-based 3 modes); (d): transient tracking error (UT-based dual vs EKF-based dual).

Plot (a) depicts the WMR tracking a demanding reference trajectory, with a non-zero initial tracking error, controlled by the proposed UT-based dual adaptive controller. It depicts the good tracking performance of the proposed scheme, even when the trajectory reaches high speeds of around 1m/s. Plots (b) to (d) correspond to another simulation trial; purposely initiated with zero error conditions, so that any initial transient errors can be fully attributed to the convergence of the estimators. Plot (b) compares the Euclidian norm of the control input vector, for the three modes of the UT-based controller, during the first 20s. The very high transient control inputs of the HCE controller reflect the aggressive and incautious nature of this mode, which ignores completely the high uncertainty in the initial estimates. Plot (c) compares the Euclidian norm of the pose error vector, for the three modes of the UT-based controller, during the first 20s. This plot shows clearly how the HCE mode typically

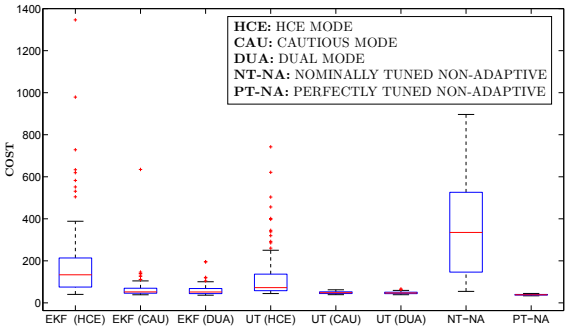

Figure 4: Monte Carlo analysis cost distributions (100runs).

leads to high initial transient errors, while the dual mode exhibits the best transient performance. This is in accordance with Remark 3.2. Plot (d) compares the UT-based dual adaptive controller with the EKF-based dual adaptive controller. The plot indicates that the former has better transient performance, while in steady-state the two controllers lead to the same error.

In addition, with minimal code optimization the computation time for the proposed UT-based dual controller is around 30% more than that of the EKF-based dual controller. This is not unexpected, mainly due to the time-intensive sigma-points propagation within the UKF algorithm. Yet, the computation time of the UT-based controller is still around 12% of the whole sampling period.

4.3 Monte Carlo Comparative Results

A Monte Carlo simulation involving 100 simulation trials was performed. Each of the eight simulations in a trial corresponds to a trajectory time horizon of one minute in real time under the simulation conditions specified earlier (and with zero error initial conditions). After each simulation the following cost $\text{COST} = \sum_{k=1}^{k_{fin}} |\boldsymbol{p_r} - \boldsymbol{p}|$ is calculated. This serves as a performance measure for each of the eight controllers operating under the same conditions, where lower values of COST are naturally preferred.

The salient statistical features of the eight cost distributions resulting from the Monte Carlo analysis, are depicted in the boxplot of Figure 4. Additionally, the mean and variance of each of these cost distributions are listed and ranked in Table 1. These results indicate clearly that *in general* the UT-based dual adaptive controller brings about a significant improvement in tracking performance, not only over non-adaptive controllers which assume nominal values for the robot parameters, but also over the EKF-based dual controller presented in (Bugeja and Fabri, 2009). More-

Table 1: Mean and Variance of the cost distributions.

Controller	Mean cost	Cost variance	Rank
EKF-HCE	192	40225	6
EKF-CAU	67	3847	4
EKF-DUA	61	731	3
UT-HCE	140	32813	5
UT-CAU	48	30	2
UT-DUA	47	26	1
NT-NA	372	59614	7
PT-NA	39	5	-na-

over, it is just as evident that within each of the two schemes, the dual control mode is even better than the cautious mode, as anticipated in Remark 3.2. This complies with the dual control philosophy that a balance between *caution* and *probing* yields the best performance in adaptive control. It is also not surprising that the performance of the HCE modes is characterized by a high cost variance and several extreme outliers. This is the result of the complete lack of caution in the presence of high initial uncertainty, leading to high transient errors. An important observation is that each mode in the UT-based scheme is superior to the corresponding mode in the EKF-based scheme. We associate this to the superior (second order) approximations introduced by the UT when compared to the EKF (first order).

5 CONCLUSIONS

The novelty in this paper comprises the use of the UT to improve on the EKF-based dual-adaptive dynamic controller recently proposed in (Bugeja and Fabri, 2009). Specifically the proposed UT-based dual-adaptive scheme employs the UKF (in predictive mode) as a recursive weight tuning algorithm, and in addition includes a novel dual-adaptive control law that uses the UT to propagate nonlinear mappings of distributions, rather that the first order approximations involved in the EKF-based law.

The presented results show clearly that the proposed novel dual controller exhibits significant improvements in performance, not only over the EKF-based dual scheme, but also on all other non-dual and non-adaptive controllers tested in this paper.

Recently we have also implemented this novel controller successfully on *Neurobot*. The obtained experimental results validate the proposed scheme on a real mobile robot for the first time and will soon be published elsewhere.

REFERENCES

Åström, K. J. and Wittenmark, B. (1995). *Adaptive Control*. Addison-Wesley, Reading, MA, 2nd edition.

Bugeja, M. K. and Fabri, S. G. (2008). Multilayer perceptron adaptive dynamic control of mobile robots: experimental validation. In H. Bruyninckx, L. P. and Kulich, M., editors, *European Robotics Symposium 2008 (EUROS'08),Prague*, Springer Tracts in advanced Robotics (STAR), pages 165–174. Springer.

Bugeja, M. K. and Fabri, S. G. (2009). Dual adaptive dynamic control of mobile robots using neural networks. *IEEE Trans. Syst., Man, Cybern. B*, 39(1):129–141.

Canudas de Wit, C., Khennoul, H., Samson, C., and Sordalen, O. J. (1993). Nonlinear control design for mobile robots. In Zheng, Y. F., editor, *Recent Trends in Mobile Robots*, Robotics and Automated Systems, chapter 5, pages 121–156. World Scientific.

Corradini, M. L. and Orlando, G. (2001). Robust tracking control of mobile robots in the presence of uncertainties in the dynamic model. *Journal of Robotic Systems*, 18(6):317–323.

Fabri, S. G. and Kadirkamanathan, V. (2001). *Functional Adaptive Control: An Intelligent Systems Approach*. Springer-Verlag, London, UK.

Fierro, R. and Lewis, F. L. (1995). Control of a nonholonomic mobile robot: Backstepping kinematics into dynamics. In *Proc. IEEE 34th Conference on Decision and Control (CDC'95)*, pages 3805–3810, New Orleans, LA.

Julier, S. J. and Uhlmann, J. K. (1997). A new extention of the Kalman filter to nonlinear systems. In *Proc. of AeroSense: The 11th Int. Symp. on Aerospace/Defence Sensing, Simulation and Controls*.

Kanayama, Y., Kimura, Y., Miyazaki, F., and Noguchi, T. (1990). A stable tracking control method for an autonomous mobile robot. In *Proc. IEEE International Conference of Robotics and Automation*, pages 384–389, Cincinnati, OH.

Wan, E. A. and van der Merwe, R. (2001). The unscented kalman filter. In Haykin, S., editor, *Kalman Filtering and Neural Networks*, Adaptive and Learning Systems for Signal Processing, Communications, and Control, chapter 7, pages 221–280. John Wiley & Sons, Inc.

Wang, T.-Y. and Tsai, C.-C. (2004). Adaptive trajectory tracking control of a wheeled mobile robot via lyapunov techniques. In *Proc. 30th Annual Conference of the IEEE Industrial Electronics Society*, pages 389–394, Busan, Korea.