

“© 2014 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.”

Model-aided State Estimation for Quadrotor Micro Air Vehicles amidst Wind Disturbances*

Dinuka Abeywardena,¹ Zhan Wang,² Gamini Dissanayake,¹ Steven L. Waslander,³ Sarath Kodagoda¹

Abstract—This paper extends the recently developed Model-Aided Visual-Inertial Fusion (MA-VIF) technique for quadrotor Micro Air Vehicles (MAV) to deal with wind disturbances. The wind effects are explicitly modelled in the quadrotor dynamic equations excluding the unobservable wind velocity component. This is achieved by a nonlinear observability of the dynamic system with wind effects. We show that using the developed model, the vehicle pose and two components of the wind velocity vector can be simultaneously estimated with a monocular camera and an inertial measurement unit. We also show that the MA-VIF is reasonably tolerant to wind disturbances, even without explicit modelling of wind effects and explain the reasons for this behaviour. Experimental results using a Vicon motion capture system are presented to demonstrate the effectiveness of the proposed method and validate our claims.

I. INTRODUCTION

Quadrotor Micro Air Vehicles (MAV) have attracted a considerable interest in the recent past due to their versatility and simple construction. Employing autonomous quadrotor MAVs to perform routine or hazardous tasks in urban and indoor environments is already a much discussed and researched topic. One of the key obstacles that still needs to be overcome to achieve such an objective is the problem of pose estimation. Though established solutions such as Inertial Navigation Systems (INS) combining Inertial Measurement Units (IMU) and GPS exist for outdoor operation, it is well known that these are not suitable for urban or indoor environments where the quality of GPS measurements degrade substantially. Other sensor modalities can be employed to replace GPS in an INS and this has been the focus of a recent string research usually categorized as “sensing in GPS-denied environments”. One promising option is a sensor suite consisting of an IMU and a monocular camera. This is commonly known as Visual-Inertial Fusion (VIF) and has emerged as a promising alternative to INS in the MAV state estimation problem.

This paper analyses and extends a novel VIF formulation for quadrotor MAVs, introduced in our previous work [1].

*This work was supported by the Centre for Autonomous Systems, University of Technology, Sydney

¹Dinuka Abeywardena, Gamini Dissanayake and Sarath Kodagoda are with the Centre for Autonomous Systems, University of Technology, Sydney. dinuka.abeywardena@ieee.org, gamini.dissanayake, sarath.kodagoda@uts.edu.au

²Zhan Wang is with Computer Vision and Active Perception Lab, Centre for Autonomous Systems, CSC, The Royal Institute of Technology KTH, Stockholm 100 44, Sweden. zhanw@kth.se

³Steven Waslander is with the Waterloo Autonomous Vehicles Laboratory, University of Waterloo, Waterloo, ON, Canada N2L 3G1. stevenw@uwaterloo.ca

Termed “Model-Aided Visual-Inertial Fusion” (MA-VIF), this approach employs quadrotor MAV dynamic equations in the state estimator design resulting in improved state estimation accuracy and consistency. The dynamic model used in [1] is only valid when the quadrotor is operating in an environment with no wind. While this is generally the case for spacious indoor environments, for most other indoor and outdoor urban environments, wind disturbances are significant and must be taken into account. As the main contribution in this paper, we show that the MA-VIF can be extended to incorporate the effect of wind on the quadrotor motion while simultaneously estimating those components of wind that affect the quadrotor MAV. We employ a rigorous theoretical analysis as well as real world experiments to illustrate the validity of this claim.

As a secondary contribution we present the reason why the MA-VIF approach as it is in [1] has the capability of coping with slowly varying wind disturbances, albeit with reduced estimation accuracy and over-confident estimates. Through a careful analysis of the process and measurement equations and experimental results, we show that when wind disturbances are not explicitly modelled, the MA-VIF tends to adjust the accelerometer biases in an attempt to reconcile the sensor measurements with its incomplete description of the vehicle dynamics.

II. RELATED WORK

Visual-inertial fusion with a monocular camera has been a much researched topic in the recent past. A tutorial level introduction to VIF along with a summary of applications can be found in [2]. Conventionally VIF algorithms are platform agnostic because they do not make use of platform specific dynamics equations in the filter design. With such a formulation, non-zero accelerations of the camera-IMU rig are a necessary condition for full observability [3].

Martin and Salaun [4] were the first to suggest the possibility of employing the quadrotor dynamic model for improved state estimation. In our previous work [5], we have designed and evaluated IMU based, model-aided state estimators for quadrotors. MA-VIF design presented in [1] extended this by incorporating a monocular camera into the state estimator. There we showed that, in contrast to the conventional VIF algorithms, this formulation is fully observable even when the quadrotor is travelling at a constant velocity. However, MA-VIF design there was based on the assumption that the quadrotor is operating in an enclosed environment with no wind. Here we analyse the case when this assumption is violated.

Estimating the direction and speed of wind can be of critical importance to MAVs operating in outdoor environments. If an estimate of the platforms ground velocity and velocity with respect to the air stream is available, then calculating wind velocity is trivial. Langelaan et al. [6] made use of air speed measurements from a pitot tube and GPS based ground velocity to directly calculate the wind velocity for a fixed wing aircraft. Using simulations, they showed that main source of error in estimating wind velocity is the air speed sensor. Zachariah et al. [7] presented a wind speed estimator suitable for a generic MAV by making use of measurements from three orthogonal anemometers and a monocular camera. There again, only simulation results were used for validation. Moreover, anemometers are hardly suitable for quadrotor MAVs with limited payload capabilities and low flight speeds. Waslander and Wang [8] overcame this issue by employing the quadrotor dynamic model to predict the MAV speed with respect to air. Using simulations they showed that it is possible to estimate wind velocity by only using a GPS and an IMU. This approach is close in spirit to the work presented here. However, to the best of our knowledge, this is the first time where the observability of such an approach has been analysed and the conditions for observability identified.

III. VISUAL-INERTIAL FUSION: METHODOLOGY

A. System Description

The platform under consideration is a quadrotor MAV affixed with an IMU (consisting of a triad of orthogonal accelerometers and gyroscopes) and a monocular camera. The body coordinate frame $\{B\}$ is assumed to be at the center of mass of the quadrotor MAV with its ${}^b\mathbf{x} - {}^b\mathbf{y}$ plane aligned with the propeller plane. Without loss of generality, we assume that the camera is placed at the origin of $\{B\}$ aligned with the same. We assume that the accelerometers are placed at the centre of mass of the quadrotor, as is the case with most commercially available platforms. We also define an earth fixed inertial coordinate frame $\{E\}$, which is defined by the position and heading of $\{B\}$ at the start of the VSLAM estimator but with its vertical axis aligned with gravity. Throughout the paper we use boldface letters to denote vectors and leading superscripts to denote coordinate frame in which the vector is expressed. A trailing subscript denotes individual components of the vector.

The states we wish to estimate are: the position of the origin of $\{B\}$ expressed in $\{E\}$, ${}^e\mathbf{p}$; velocity of the origin of $\{B\}$ measured in $\{E\}$, but expressed in $\{B\}$, ${}^b\mathbf{v}$; scale of the VSLAM estimates λ ; orientation of $\{B\}$ with respect to $\{E\}$ expressed in Z-Y-X Euler angles, $\Theta = \{\phi, \theta, \psi\}$, accelerometer bias β_a and gyroscope bias β_g . The translational and rotational dynamics of the quadrotor MAV are governed by:

$${}^e\dot{\mathbf{v}} = \mathbf{g} + \frac{1}{m} {}^e\mathbf{F}_A \quad (1)$$

$${}^b\dot{\Omega} = J^{-1}({}^b\mathbf{M}_A - {}^b\Omega \times J {}^b\Omega) \quad (2)$$

where \mathbf{g} is the gravity vector, m is the mass of the quadrotor, Ω is the body rotational rate, \mathbf{F}_A , \mathbf{M}_A are

respectively the sum of aerodynamic forces and moments acting on the quadrotor and J is the moment of inertia matrix for the vehicle.

As illustrated in Fig. 1 there are two separate, but coupled, estimators in the system that we are about to describe. First is the monocular SLAM estimator, termed VSLAM, similar to that detailed in [9]. We assume that, given the recent advances in VSLAM algorithms, it is reasonable to consider the VSLAM block as a sensor, which produces unbiased position (up-to-scale) and orientation measurements with an error covariance. This assumption allows us to model the VSLAM estimates as:

$$\begin{cases} \mathbf{h}_{vp} = \lambda {}^e\mathbf{p} + \boldsymbol{\eta}_p \\ \mathbf{h}_{vo} = \Theta + \boldsymbol{\eta}_o \end{cases} \quad (3)$$

where $\boldsymbol{\eta}_p, \boldsymbol{\eta}_o$ are White Gaussian Noise (WGN) terms and $E[\boldsymbol{\eta}_p\boldsymbol{\eta}_o][\boldsymbol{\eta}_p\boldsymbol{\eta}_o]^T] = P_{rr}$ is the VSLAM pose covariance. ($E[\cdot]$ denotes the expectation operator.)

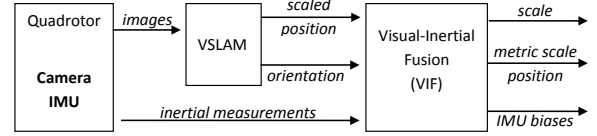


Fig. 1: Visual-inertial fusion for a quadrotor.

The second estimator which fuses VSLAM estimates and inertial measurements is the focus of this paper. This VIF block makes use of accelerometer and gyroscope measurements which are assumed to be corrupted by a fixed, but unknown bias and a zero mean WGN term. These measurements can be modelled as:

$${}^b\mathbf{a} = {}^b\dot{\mathbf{v}} - R^T\mathbf{g} + \beta_a + \boldsymbol{\eta}_a \quad (4)$$

$${}^b\boldsymbol{\omega}_g = {}^b\Omega + \beta_g + \boldsymbol{\eta}_\Theta$$

where R is the rotation matrix that transforms a vector from $\{B\}$ to $\{E\}$, ${}^b\mathbf{a}$ and ${}^b\boldsymbol{\omega}_g$ are, respectively, the accelerometer and gyroscope measurements in $\{B\}$ and $\boldsymbol{\eta}_a$ and $\boldsymbol{\eta}_\Theta$ are zero mean WGN terms.

IV. MODEL-AIDED VISUAL-INERTIAL FUSION

A MA-VIF formulation attempts explicitly model some of the dynamics in equation in (1). A quadrotor MAV in-flight is affected by three forces: gravity, thrust and aerodynamic drag. For quadrotor MAVs with relatively stiff propeller blades, induced drag dominates the aerodynamic drag at low translational velocities [10]. The amount of induced drag is proportional to the MAV translational velocity with respect to the wind, in the propeller plane. Thus for near hover operation \mathbf{F}_A can be approximated by:

$${}^e\mathbf{F}_A = -F_T\mathbf{b}_3 - \alpha \sum_{i=1}^4 \omega_i \tilde{\mathbf{v}}_\infty \quad (5)$$

where F_T is the magnitude of the summation of propeller thrust, ω_i is the rotational velocity of i^{th} propeller, α is a positive coefficient known as rotor drag coefficient, \mathbf{b}_3 is the unit vector along ${}^b\mathbf{z}$ and $\tilde{\mathbf{v}}_\infty$ is the projection of ${}^e\mathbf{v}_\infty$ on the propeller plane. The free stream velocity ${}^e\mathbf{v}_\infty = {}^e\mathbf{v} - {}^e\mathbf{v}_w$ where ${}^e\mathbf{v}_w$ is the velocity of the wind in $\{E\}$. To derive the process equations for the MA-VIF we first combine (5) with (1) and transform the result to $\{B\}$ frame:

$${}^b\dot{\mathbf{v}} \approx R^T \mathbf{g} - [k_1 {}^b v_{\infty x} \quad k_1 {}^b v_{\infty y} \quad \frac{F_T}{m}]^T \quad (6)$$

where the second order terms resulting from the coordinate transformation are neglected and $k_1 = \frac{\alpha}{m} \sum_{i=1}^4 \omega_i$. The validity of this equation for situations where there is no wind has already been analysed in [5]. There, experimental results show that k_1 is approximately a constant, the value of which for a specific quadrotor MAV can be approximated via equation (6) when ground truth state estimates are available. We make the same assumption throughout the rest of this paper. Also, by combining equations (6) and (4), we note that it is possible to replace $\frac{F_T}{m}$ by ${}^b a_z - \beta_{az}$.

It is now possible to incorporate (6) into the state estimator process model given that a dynamic model for wind velocity is available. However, given the complexities involved in developing and adopting such a model, here we make the assumption that the wind velocity is smoothly varying in $\{E\}$. Though we employ this assumption for both the estimator design and analysis, it will be later shown through experiments that this proves to be adequate for coping with real wind patterns.

With the above assumptions, it is now possible to complete the process equation for the MA-VIF:

$$\begin{bmatrix} {}^e\dot{\mathbf{p}} \\ {}^b\dot{\mathbf{v}} \\ \dot{\lambda} \\ \dot{\Theta} \\ \dot{\beta}_a \\ \dot{\beta}_g \\ {}^e\dot{\mathbf{v}}_w \end{bmatrix} = \underbrace{\begin{bmatrix} R^b \mathbf{v} \\ \mathbf{f}_v + \boldsymbol{\eta}_v \\ \eta_\lambda \\ -\Xi \boldsymbol{\beta}_g - \boldsymbol{\eta}_\Theta \\ \boldsymbol{\eta}_{\beta_a} \\ \boldsymbol{\eta}_{\beta_g} \\ \boldsymbol{\eta}_w \end{bmatrix}}_{\mathbf{f}_0} + \underbrace{\begin{bmatrix} \mathbf{0}_{3 \times 1} \\ \mathbf{e}_3 \\ 0 \\ \mathbf{0}_{3 \times 1} \\ \mathbf{0}_{3 \times 1} \\ \mathbf{0}_{3 \times 1} \\ \mathbf{0}_{3 \times 1} \end{bmatrix}}_{\mathbf{f}_1} {}^b a_z + \underbrace{\begin{bmatrix} \mathbf{0}_{3 \times 1} \\ \mathbf{0}_{3 \times 1} \\ 0 \\ \Xi \\ \mathbf{0}_{3 \times 1} \\ \mathbf{0}_{3 \times 1} \\ \mathbf{0}_{3 \times 1} \end{bmatrix}}_{\mathbf{f}_2} {}^b \boldsymbol{\omega}_g. \quad (7)$$

where $\mathbf{f}_v = R^T \mathbf{g} - k_1 \Lambda ({}^b \mathbf{v} - R^T {}^e \mathbf{v}_w) - \Gamma \boldsymbol{\beta}_a$, $\mathbf{e}_3 = [0 \ 0 \ 1]^T$, $\eta_\lambda, \boldsymbol{\eta}_{\beta_a}$ and $\boldsymbol{\eta}_{\beta_g}$ are WGN terms and Ξ is a matrix that converts body rotational rates to Euler rates. The state λ was included to account for the vision measurements. Note that we have used the gyroscope and ${}^b\mathbf{z}$ accelerometer measurements as system inputs. Also:

$$\Lambda = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad \Gamma = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

and $\boldsymbol{\eta}_v, \boldsymbol{\eta}_w$ are zero mean WGN terms denoting uncertainties in process equations due to the assumptions made previously.

It is also possible to augment the accelerometer measurement equation by combining (4) with (6). This leads to:

$$\mathbf{h}_i = \begin{bmatrix} {}^b a_x \\ {}^b a_y \end{bmatrix} = -k_1 \Upsilon ({}^b \mathbf{v} - R^T {}^e \mathbf{v}_w) + \Upsilon \boldsymbol{\beta}_a + \Upsilon \boldsymbol{\eta}_a \quad (8)$$

where Υ is a matrix consisting of the first two rows of Λ . Equations (7), (3) and (8) make up the process and measurement models for the MA-VIF.

V. OBSERVABILITY OF VISUAL-INERTIAL FUSION

Since both process and measurement equations of the MA-VIF are nonlinear, we resort to the ideas presented by Hermann and Krener [11] to analyse the observability of the MA-VIF formulation. They proposed a rank condition test for “locally weak observability” of a non-linear system: *a system is locally weakly observable if it satisfies the observability rank condition generically* [11]. A system described by a process equation of the form $\dot{\mathbf{X}} = \mathbf{f}_0(\mathbf{X}) + \sum \mathbf{f}_i(\mathbf{X})u_i, i = 1 \dots l$, and a measurement equation of the form $\mathbf{y} = \mathbf{h}(\mathbf{X})$ satisfies the observability rank condition if any one of the possible matrices whose rows are of the form

$$\mathcal{O} = \{\nabla L_{f_i \dots f_j}^n h_k(\mathbf{x}) | i, j = 0, \dots, l; k = 1, \dots, m; n \in \mathbb{N}\}$$

is of full column rank. Here m is the number of measurements. Also, $\nabla L_{f_i}^n h_k(\mathbf{x})$ is the gradient of the n^{th} order Lie derivatives of $h_k(\mathbf{x})$ with respect to f_i . Locally weak observability is a necessary condition but not sufficient for strict observability. However, following the approach of [3] and many others, we assume that for most mobile robotic navigational tasks, locally weak observability implies that the considered system contains enough information to perform state estimation. Experimental results presented in sections VII serve to validate the claims made based on this assumption.

If the observability matrix \mathcal{O} is rank deficient, then one or more of the states are not observable. Martinelli [12] introduced a method for identifying any unobservable modes using the null space of \mathcal{O} .

PROPERTY 1: $g(\mathbf{X})$ is an observable mode if and only if its gradient is orthogonal to the null space of \mathcal{O} . This property can be expressed by a system of partial differential equations of the form:

$$\sum_{i=1}^n w_{si}(\mathbf{X}) \frac{\partial g}{\partial X_i} = 0 \quad (9)$$

where $w_{si}(\mathbf{X})$ is the i^{th} component of \mathbf{w}_s which is a non-trivial element of the null space of \mathcal{O} .

A. MA-VIF Estimator

Observability of the MA-VIF formulation with no wind has been analysed previously in [1]. Here we proceed to extend that analysis to the case with non-zero wind. The system under consideration is fully characterised by (7), (3) and (8), with zero noise terms as they do not affect the observability. Note that $\dim(\mathbf{X}) = 19$.

By considering the non-trivial Lie derivatives of the system up to second order, we can construct the following observability matrix for the MA-VIF:

$$\mathcal{O}_1 = [\nabla L^0 h_{vp}(x) \quad \nabla L^0 h_{vo}(x) \quad \nabla L^0 h_i(x) \quad \nabla L_{f_0}^1 h_{vp}(x) \dots \nabla L_{f_0}^1 h_{vo}(x) \quad \nabla L_{f_0}^1 h_i(x) \quad \nabla L_{f_0}^2 h_{vp}(x)]^T \quad (10)$$

$$\begin{bmatrix} \lambda I_3 & \mathbf{0}_{3 \times 3} & {}^e p & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} \\ \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 1} & I_3 & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} \\ \mathbf{0}_{2 \times 3} & -k_1 \Upsilon & \mathbf{0}_{2 \times 1} & D_1 & \Upsilon & \mathbf{0}_{3 \times 3} & -k_1 \Upsilon R^T \\ \mathbf{0}_{3 \times 3} & \lambda R & R^b v & D_2 & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} \\ \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 1} & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} & \Xi & \mathbf{0}_{3 \times 3} \\ \mathbf{0}_{2 \times 3} & k_1^2 \Upsilon & \mathbf{0}_{2 \times 1} & D_3 & \mathbf{0}_{2 \times 3} & \mathbf{0}_{3 \times 3} & k_1^2 \Upsilon R^T \\ \mathbf{0}_{3 \times 3} & -\lambda k_1 R \Lambda & D_4 & D_5 & \lambda R \Gamma & \mathbf{0}_{3 \times 3} & \lambda k_1 R \Lambda R^T \end{bmatrix}$$

where I_3 is the 3×3 identity matrix, $\mathbf{0}_{m \times n}$ is an all zero matrix of size $m \times n$, $D_i (i = 1 \dots 5)$ are functions of $\lambda, \Theta, {}^b v, \beta_a, {}^e v_w$, with:

$$D_4 = g - k_1 R \Lambda {}^b v + k_1 R \Lambda R^T {}^e v_w + R \Gamma \beta_a.$$

It can be shown that all other Lie derivatives of the considered system are either linearly dependent on the rows of \mathcal{O}_1 or are trivially zero.

Lemma 1: The matrix \mathcal{O}_1 is rank deficient. It has a column rank of 18 when at least one of ${}^b \dot{v}_x$ and ${}^b \dot{v}_y$ is non-zero.

Proof: We proceed using block Gaussian elimination to recover the rank of \mathcal{O}_1 . The notation $\mathcal{O}_1(p, q)$ is used to represent the $(p, q)^{th}$ block, and $\mathcal{O}_1(p, :)$ represent the p^{th} row block of \mathcal{O}_1 .

First, given that $\lambda > 0$, we divide row block one by λ . Also since block $\mathcal{O}_1(2, 4)$ is an identity matrix, we eliminate all non-zero entries in $\mathcal{O}_1(:, 4)$. Rotational transformation matrix R is orthogonal and thus is always full rank. This allows us to multiply $\mathcal{O}_1(4, :)$ by $(\lambda R)^{-1}$. Ξ transforms the body rotational rates to Euler rates and is also full rank. Thus can we multiply $\mathcal{O}_1(5, :)$ by Ξ^{-1} . Next, we multiply $\mathcal{O}_1(7, :)$ by $\frac{k_1}{\lambda} \Upsilon R^T$ and add the result to $\mathcal{O}_1(6, :)$ to obtain:

$$\begin{bmatrix} I_3 & \mathbf{0}_{3 \times 3} & {}^e p & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} \\ \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 1} & I_3 & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} \\ \mathbf{0}_{2 \times 3} & -k_1 \Upsilon & \mathbf{0}_{2 \times 1} & \mathbf{0}_{2 \times 3} & \Upsilon & \mathbf{0}_{3 \times 3} & -k_1 \Upsilon R^T \\ \mathbf{0}_{3 \times 3} & I_3 & \frac{{}^b v}{\lambda} & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} \\ \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 1} & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} & I_3 & \mathbf{0}_{3 \times 3} \\ \mathbf{0}_{2 \times 3} & \mathbf{0}_{2 \times 3} & D_7 & \mathbf{0}_{2 \times 3} & \mathbf{0}_{2 \times 3} & \mathbf{0}_{3 \times 3} & \mathbf{0}_{2 \times 3} \\ \mathbf{0}_{3 \times 3} & -\lambda k_1 R \Lambda & D_4 & \mathbf{0}_{3 \times 3} & \lambda R \Gamma & \mathbf{0}_{3 \times 3} & \lambda k_1 R \Lambda R^T \end{bmatrix}$$

where

$$\begin{aligned} D_7 &= \frac{k_1}{\lambda} \Upsilon R^T (g - k_1 R \Lambda {}^b v + k_1 R \Lambda R^T {}^e v_w + R \Gamma \beta_a) \\ &= \frac{k_1}{\lambda} \Upsilon {}^b \dot{v} = \frac{k_1}{\lambda} \begin{bmatrix} {}^b \dot{v}_x & {}^b \dot{v}_y \end{bmatrix}^T \end{aligned}$$

If either one of ${}^b \dot{v}_x$ and ${}^b \dot{v}_y$ is non-zero, then at least one element of the 2×1 vector D_7 is non-zero and we can proceed to eliminate all remaining non-zero elements of $\mathcal{O}_1(:, 3)$. This results in $\mathcal{O}_1(4, :)$ being all zero except in column block two. Thus we can then proceed to eliminate all remaining non-zero elements in $\mathcal{O}_1(:, 2)$. Now consider the following sub-matrix made up of all remaining non-zero elements of $\mathcal{O}_1(3, :)$ and $\mathcal{O}_1(7, :)$.

$$\mathcal{O}_2 = \begin{bmatrix} \Upsilon & -k_1 \Upsilon R^T \\ \lambda R \Gamma & \lambda k_1 R \Lambda R^T \end{bmatrix}$$

Multiplying $\mathcal{O}_2(2, :)$ by $(\lambda R)^{-1}$ and then adding $\Upsilon \mathcal{O}_2(2, :)$ to $\mathcal{O}_2(1, :)$ and finally moving the last row of $\mathcal{O}_2(2, :)$ to

immediately after $\mathcal{O}_2(1, :)$ we obtain:

$$\begin{bmatrix} I_3 & \mathbf{0}_{3 \times 3} \\ \mathbf{0}_{2 \times 3} & k_1 \Upsilon R^T \end{bmatrix}$$

which has a column rank of 5 for any given R . Thus returning to \mathcal{O}_1 we see that its column rank is 18. ■

B. Unobservable Modes

The above analysis proves that the dimension of the null space of \mathcal{O}_1 is 1. It is easy to show that

$$\mathbf{n}_1 = [\mathbf{0}_{3 \times 1}; \mathbf{0}_{3 \times 1}; 0; \mathbf{0}_{3 \times 1}; \mathbf{0}_{3 \times 1}; \mathbf{0}_{3 \times 1}; Re_3]$$

is a basis for the said null space. It is possible to show that $g(\mathbf{X}) = {}^b v_{wz} = \Gamma_3 {}^b v_w = \Gamma_3 R^T {}^e v_w$ is not orthogonal to \mathbf{n}_1 , where Γ_3 is the third column of Γ . Thus, making use of property 1, we can conclude that the unobservable dimension corresponding to \mathbf{n}_1 is the ${}^b z$ component of wind.

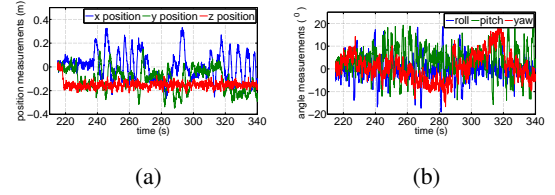


Fig. 2: Simulated VSLAM position(a) and orientation (b) measurements. Note the scale of position measurements.

VI. ESTIMATOR DESIGN

Considering the non-linearities in the process and measurement equations, we propose an Extended Kalman Filter (EKF) based estimator for the MA-VIF. To proceed with the estimator design, we need to remove the unobservable mode, namely the ${}^b z$ component of wind velocity from the state vector. Note that the quadrotor dynamics in (6) is independent of this component. However, as the assumption about slow varying wind velocity in $\{E\}$ meant that wind velocity had to be expressed in $\{E\}$ in the state vector, removing the ${}^b z$ component of wind velocity is not straightforward. To achieve this, consider the sub-matrix $Q_w = E[\eta_w \eta_w^T]$ of the process noise matrix required for the EKF design. If we assume that all three components of η_w are uncorrelated and of the same magnitude, then it can be shown that Q_w remains diagonal even after a frame transformation. This shows that, for the purpose of the estimator design, the assumption about smoothly varying wind velocity can be made either in $\{E\}$ or $\{B\}$ with equivalent results. Having expressed wind velocity in $\{B\}$, it is now straightforward to remove the unobservable component ${}^b \dot{v}_{wz}$ from the state vector. The resulting process and measurement equations are then:

$$\mathbf{h}_i = \begin{bmatrix} {}^b a_x \\ {}^b a_y \end{bmatrix} = -k_1 \Upsilon ({}^b v - {}^b v_w) + \Upsilon \beta_a + \Upsilon \eta_a \quad (11)$$

$$\begin{bmatrix} {}^e\dot{\mathbf{p}} \\ {}^b\dot{\mathbf{v}} \\ \dot{\lambda} \\ \dot{\Theta} \\ \dot{\beta}_a \\ \dot{\beta}_g \\ \dot{\tilde{\mathbf{v}}}_w \end{bmatrix} = \begin{bmatrix} R^T \mathbf{g} - k_1 \Lambda ({}^b\mathbf{v} - {}^b\mathbf{v}_w) + \Gamma ({}^b\mathbf{a} - \beta_a) + \boldsymbol{\eta}_v \\ \eta_\lambda \\ \Xi ({}^b\boldsymbol{\omega}_g - \beta_g - \boldsymbol{\eta}_\Theta) \\ \boldsymbol{\eta}_{\beta_a} \\ \boldsymbol{\eta}_{\beta_g} \\ \tilde{\boldsymbol{\eta}}_w \end{bmatrix} \quad (12)$$

where, $\tilde{\mathbf{v}}_w = [{}^b v_{wx} \ {}^b v_{wy}]^T$. VSLAM measurement equations remain the same. Note that we do not make the assumption that the ${}^b z$ component of wind velocity is zero. The removal of the ${}^b z$ component of wind velocity from the state vector is made possible by the fact that when posed in this formulation, the process equation for ${}^b\dot{\mathbf{v}}$ is independent of that component. An observability analysis similar to that carried out in section V can be employed to show that same condition in *Lemma 1* is sufficient to ensure that the system described by (12), (11) and (3) is locally-weakly observable. This allows us to proceed with an EKF design to estimate \mathbf{X} in the standard manner. We omit the details for brevity.

Additionally, it is possible to rewrite (11) by combining β_a and ${}^b\mathbf{v}_w$ as they both have the same process dynamics.

$$\mathbf{h}_i = \begin{bmatrix} {}^b a_x \\ {}^b a_y \end{bmatrix} = -k_1 \Upsilon {}^b\mathbf{v} + \Upsilon \tilde{\beta}_a + \Upsilon \boldsymbol{\eta}_a \quad (13)$$

where $\tilde{\beta}_a = \beta_a + k_1 {}^b\mathbf{v}_w$. This implies that it is not possible to distinguish between the effect of wind and bias by only analysing the accelerometer measurements. Wind velocity and accelerometer biases are only made observable as the same replacement cannot be done in the process dynamics for ${}^b\dot{\mathbf{v}}$ in (12), where β_{ax} and β_{ay} do not appear. However, in practise this implies that when the wind effects are not explicitly modelled, as is the case in [1], it is possible for a state estimator to account for the difference between true and predicted measurements by adjusting the accelerometer biases. Experimental results presented in the next section shows that this is indeed the case.

VII. EXPERIMENTS

A. Experimental Setup

The quadrotor platform used for the experiments was the Parrot AR Drone II. It is equipped with two cameras facing forward and downward as well as a triad of accelerometers and gyroscopes. Accelerometer and gyroscope measurements captured at 200Hz were timestamped on-board the MAV and wirelessly transmitted to a ground station computer. We collected several flight data sets by manually piloting the AR Drone II in an environment with a Vicon motion capture system which used to produce real-time estimates for the vehicle states at 120Hz. An industrial grade fan situated at one side of the Vicon capture area was used to generate wind predominately along the x axis of $\{\mathbf{E}\}$ frame. A hand held anemometer was used to obtain a rough measurement of the wind speed which at the centre of the arena was about 1.5m/s.

The purpose of the experiments presented here is to verify the ability of the MA-VIF to produce accurate vehicle pose estimates amidst wind disturbances, given the measurements from a suitable VSLAM algorithm and an IMU. For this reason, we chose to simulate the VSLAM algorithm using the Vicon motion estimates. This was achieved by first sub-sampling the Vicon position and orientation estimates at 10Hz and then scaling and adding suitable levels of noise according to equation (3) with $\lambda = 1/5$, $\eta_p \in \mathcal{N}(0, (1/5)^2)$ (1 m standard deviation in metric scale) and $\eta_o \in \mathcal{N}(0, 4)$ (2 degrees standard deviation). The resulting position and orientation measurements are illustrated in Fig. 2.

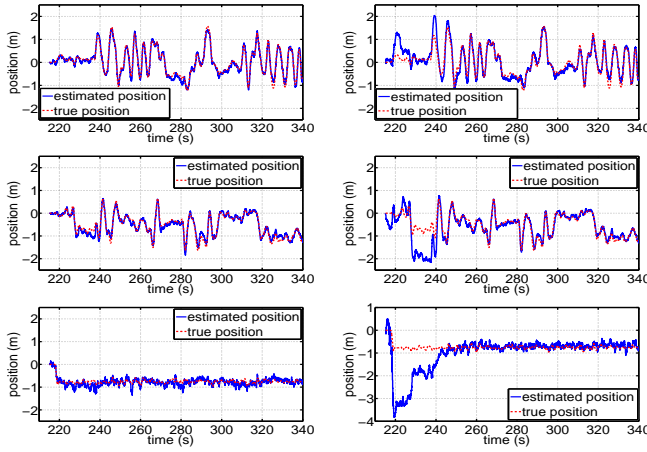
B. Results

We performed two different experiments with each gathered data set. First the IMU and simulated VSLAM measurements were processed using a MA-VIF implementation designed for the case of no wind, originally presented in [1]. For ease of reference, we term this the MA-VIF-NW. Second, the same set of data were processed by the MA-VIF design presented here. For both experiments, position and scale estimates are presented in Fig. 3 and 4 respectively.

First, we see that the MA-VIF is capable of quickly recovering the scale λ thus producing accurate position and velocity estimates in metric scale. The scale estimate closely matches the true value. It can also be seen that the scale estimate begins to deviate from the true value when the both ${}^b v_x$, ${}^b v_y$ are approximately zero. This agrees with the theoretical observability results derived in section V. The wind velocity estimates for the MA-VIF are presented in Fig. 4. Note how the wind is predominantly in the x direction as is expected. Also note how when the quadrotor MAV moves away from the origin along y axis, both the x and y components of wind estimates rapidly decrease (for example 290s - 300s and 320s - 340s). This indicates that the wind estimates are qualitatively accurate.

Focusing on the MA-VIF-NW implementation, we note that it required careful filter tuning to obtain the results presented here and the estimator performance was much more sensitive to noise parameters. Considering the estimates of MA-VIF-NW, we see that the scale estimate takes longer to converge and as a result produce substantial errors in position estimates. The advantage of the MA-VIF can be seen when the estimation accuracies are quantitatively analysed. Velocity estimation errors of both estimators presented in Fig. 5 illustrates this fact clearly. RMS velocity estimation errors for the MA-VIF were 0.21, 0.15 and 0.17 for the x,y,z axes respectively and 0.38, 0.29 and 0.57 for MA-VIF-NW. This is as expected, since the MA-VIF-NW makes use of an incorrect model to describe the evolution of body frame velocity, thus resulting in over-confident and less accurate estimates.

The accelerometer bias estimates of both estimators are presented in Fig. 6. Note the similarity in MA-VIF-NW bias estimates for the ${}^b x$, ${}^b y$ axes and the wind velocity estimates for the same axes for the MA-VIF. In fact, as expected, it can be seen that the accelerometer biases of MA-VIF-



(a) (b)

Fig. 3: Position estimates with ground truth of MA-VIF - (a) and MA-VIF-NW - (b). Top, middle and bottom are respectively x, y, z axes estimates.

NW agrees closely with the values predicted for $\tilde{\beta}_a$ by (13). We remark that this is also the reason why MA-VIF-NW estimates are not substantially erroneous, despite using an insufficient process model.

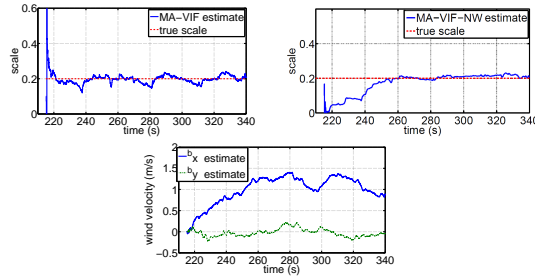


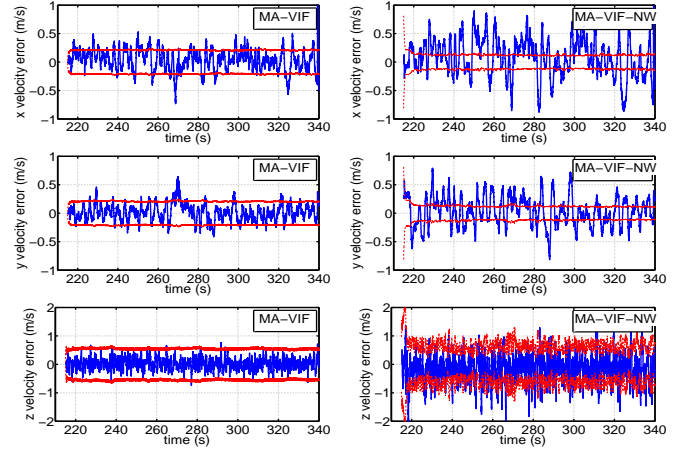
Fig. 4: Scale estimates and MA-VIF wind velocity estimate

VIII. CONCLUSION

This paper extended the model-aided visual-inertial fusion for quadrotor MAV state estimation by considering the effects of wind disturbances. By explicitly modelling the effects of wind on quadrotor MAV dynamics, we showed that it is possible to simultaneously estimate both the vehicle pose and two components of the wind velocity vector, only using a monocular camera and an inertial measurement unit. Experiments conducted in an indoor environment with light wind disturbances illustrate that this approach also improves the pose estimation accuracy of the MAV. In addition to the results presented here, a quantitative analysis of the wind velocity estimates in rapidly varying wind gusts was performed with the use of a high-fidelity quadrotor simulator and will be presented in a future article.

REFERENCES

- [1] D. Abeywardena, Z. Wang, S. Kodagoda, and G. Dissanayake, “Visual-inertial fusion for quadrotor Micro Air Vehicles with improved scale



(a) (b)

Fig. 5: Velocity estimation with 3σ bounds of MA-VIF - (a) and MA-VIF-NW - (b).

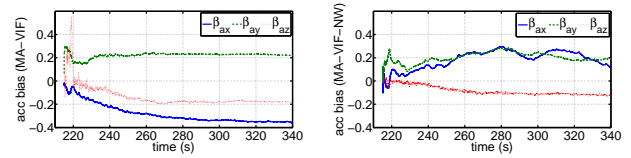


Fig. 6: Accelerometer bias estimates.

- observability,” in *Proc. IEEE Int. Conf. Robot. Autom.*, May 2013, pp. 3133–3138.
- [2] P. Corke, J. Lobo, and J. Dias, “An introduction to inertial and visual sensing,” *Int. J. Robot. Res.*, vol. 26, no. 6, pp. 519–535, 2007.
- [3] J. Kelly and G. S. Sukhatme, “Visual-inertial sensor fusion: Localization, mapping and sensor-to-sensor self-calibration,” *Int. J. Robot. Res.*, vol. 30, no. 1, pp. 56–79, 2011.
- [4] P. Martin and E. Salaun, “The true role of accelerometer feedback in quadrotor control,” in *Proc. IEEE Int. Conf. Robot. Autom.*, May 2010, pp. 1623–1629.
- [5] D. Abeywardena, S. Kodagoda, G. Dissanayake, and R. Munasinghe, “Improved state estimation in quadrotor MAVs: A novel drift-free velocity estimator,” *IEEE Robot. Autom. Mag.*, vol. PP, no. 99, pp. 1–1, 2013.
- [6] J. W. Langelaan, N. Alley, and J. Neidhoefer, “Wind field estimation for small unmanned aerial vehicles,” *Journal of Guidance, Control, and Dynamics*, vol. 34, no. 4, pp. 1016–1030, 2011.
- [7] D. Zachariah and M. Jansson, “Self-motion and wind velocity estimation for small-scale uavs,” in *Robotics and Automation (ICRA), 2011 IEEE International Conference on*. IEEE, 2011, pp. 1166–1171.
- [8] S. L. Waslander and C. Wang, “Wind disturbance estimation and rejection for quadrotor position control,” in *AIAA Infotech@ Aerospace Conference and AIAA Unmanned... Unlimited Conference*, Seattle, WA, 2009.
- [9] J. Civera, A. Davison, and J. Montiel, “Inverse depth parametrization for monocular SLAM,” *IEEE Trans. Robot.*, vol. 24, no. 5, pp. 932–945, Oct. 2008.
- [10] R. Mahony, V. Kumar, and P. Corke, “Multirotor aerial vehicles: Modeling, estimation, and control of quadrotor,” *IEEE Robot. Autom. Mag.*, vol. 19, no. 3, pp. 20–32, 2012.
- [11] R. Hermann and A. Krener, “Nonlinear controllability and observability,” *IEEE Trans. Autom. Control*, vol. 22, no. 5, pp. 728–740, Oct. 1977.
- [12] A. Martinelli, “State estimation based on the concept of continuous symmetry and observability analysis: The case of calibration,” *IEEE Trans. Robot.*, vol. 27, no. 2, pp. 239–255, April 2011.