Magneto-visual-inertial dead-reckoning : improving estimation consistency by Invariance

David Caruso, Alexandre Eudes, Martial Sanfourche, David Vissière, Guy Le Besnerais

Magneto-visual-inertial Dead-reckoning : Improving Estimation Consistency by Invariance

David Caruso[1,2], Alexandre Eudes[2], Martial Sanfourche[2], David Vissière[1] and Guy Le Besnerais[2]

***Abstract*— Tractable algorithms used for 6DOF visual-inertial odometry have decades-long history of estimation consistency issues. Those arise in particular in two well-studied filters: namely the EKF-SLAM and MSCKF. Recently, strong theoretical works linked the error-state of these filters with their consistency properties; these results led to the synthesis of far more consistent filters. In previous works, we have shown that using similar filter for the fusion of magneto-inertial sensors with optical ones improved classical visual-inertial navigation systems. The consistency of those novel magneto-visual-inertial filters were, however, not addressed until now. This work does. We apply invariance theory findings to the specific case of magneto-inertial odometry and magneto-visual-inertial odometry for the synthesis of a filter with interesting consistency properties. We describe thoroughly such an invariant filter, implement it and conduct experiments on carefully captured data from real sensor. By comparing the results of non-invariant, observability-constrained and invariant versions of the filter, we find that the invariant version (i) shows an error estimate that is consistent with observability of the system, (ii) is applicable in case of unknown heading at initialization, (iii) improves long-term behavior of the filter and (iv) exhibits a lower normalized estimation error. We experiment on challenging scenarios for regular visual-inertial pedestrian navigation systems.**

I. Introduction

We address here the problem of an embedded device position and orientation estimation. The goal is to estimate these quantities with minimum drift with power efficiency and robustness with respect to the various type of motion. We also target an accurate estimation of the uncertainty of these quantities. In one hand, Visual-Inertial Odometry (VIO) has become a fundamental technology for positioning in spatial computing applications such as augmented or virtual reality. In the other hand, Magneto-Inertial Dead-Reckoning (MI-DR) uses the local magnetic perturbations for correcting accelerometers integration drift and recently showed dramatic performance improvement [1]. As failure modes of MI-DR and VIO are disjoint, it was shown that a fused Magneto-Visual-Inertial Navigation System (MVINS) estimator was more robust than either VIO and MI-DR [2, 3]. These latter references use an algorithm based on the Multi-State Constraint Filter (MSCKF) [4] adapted for MVINS problem that we will name here Magneto-Inertial Multi-State Constraint Filter (MI-MSCKF). It is well-known that, in the context of VIO, the MSCKF applied without extra care is not consistent: the state errors are not well represented by the estimated covariance [5]. In this extent, there is no reason why the MI-MSCKF algorithm presented in [3] would be free from the same flaws than its inspirator.

This paper shows that these inconsistencies are significant issues for MVINS, and address them by leveraging the link between the filter's error-state definition and its consistency property. We base our work on theory developped in [6], that was already recently applied to VIO [7]. The contributions of this paper are (i) the description of a Lie group-based parametrization for the MVINS model inspired by invariant theory of [6] (ii) the proof that a MSCKF-like filter based on this parametrization has consistency properties (we use here the *unobservable stochastic transformation* of [7] with some corrections on their mathematical developments), (iii) the comparison of several variants of MVINS filters on a real dataset showing the advantages of the invariant filter.

Furthermore, we prove consistency properties from the MSCKF equations directly. In contrast, all previous works we are aware of, transpose such theoritical results from the associated EKF-SLAM filter.

II. Related work

It has been known for many years that consistency issues can arise from linearization errors in the EKF when using it in a Simultaneous Localization and Mapping (SLAM) context. The authors of [8] demonstrated that this stems in part from the fact that some state-variables are used in several measurements equations with different linearization points. In constrast, full batch optimization of the underlying cost function does not suffer from these sources of spurious observability. However, batch methods are computationally intensive and even if incremental solvers exist [9, 10], truly real-time implementations marginalize past poses and face the same consistency issues [11]. An often used workaround is to freeze the linearization point of all variables already involved in marginalization [12]. This idea was applied to the MSCKF in [13] and to a batch approach in [11]. Yet the potential negative side effects of this practice are unclear.

Another solution is to alter artificially the transition and measurement matrices of a filter in order to enforce the non-observable space of linearized model [14]. This method is referred as Observability Constrained EKF (OC-EKF) in the literature and was applied to MSCKF with

[1] D. Caruso and D. Vissière are with Sysnav
[2] A. Eudes, M. Sanfourche, G. Le Besnerais are with ONERA

success in [15]. This approach is however not theoreticaly strongly founded.

The influence of parametrization of the error-state on the consistency property of EKF has been the subject of research for a long time [16–18]. However, the most promising works regarding this idea arose from the theory of invariance of estimators [19, 20] applied to the EKF in [21] and later to SLAM in [6]. In the latter work, the authors expoloits the symetries of the problem to design the error-state, and synthesize a consistent EKF-SLAM (EKF-SLAM). This "invariant" parametrization ideas disseminated quickly and were used in various contexts such as in [7, 22]. In particular, [23] is very close to our approach, while targeting a totally different application. Very valuable information about the mathematics behind these development can be found in [24, 25].

III. Problem statement

A. Kalman filtering with Non-linear Error

We use a very general formulation of the Kalman filter as in [24, 26], in which uncertainty is represented through the covariance matrix $\mathbf{\Sigma_e}$ of a non-linear error vector $\mathbf{e}$. The latter is defined through an abstract *retraction operator* $\boxplus$:

$$\mathbf{X} = \hat{\mathbf{X}} \boxplus \mathbf{e}, \quad \mathbf{e} \propto \mathcal{N}(0, \mathbf{\Sigma_e}) \tag{1}$$

where $\mathbf{X}$ is the true state value, $\hat{\mathbf{X}}$ is the estimated one and $\mathcal{N}$ the normal distribution. Error's selection is merely a design choice; it defines the space where the uncertainty of the filter is approximated as a Gaussian. It also governs the linearization process of the filter. Therefore it has a major impact on the filter performance.

Loosely speaking, this error operator must verify expected properties for an error such as $\mathbf{X}_0 \boxplus 0 = \mathbf{X}_0$ and have a reciprocal operator $\boxminus$ so that $\mathbf{e} = \mathbf{X} \boxminus \hat{\mathbf{X}}$ around zero. To use it in a EKF, it must also be continuous and differentiable at least in the vicinity of zero.

Let us consider the following generic discrete dynamic model:

$$\begin{aligned} &\text{(propagation)} && \mathbf{X}_{k+1} = f(\mathbf{X}_k, \mathbf{u}_k, \eta_k), \\ &\text{(measurement)} && y = h(\mathbf{X}_k) + \nu_k \end{aligned} \tag{2}$$

with $\eta_k \propto \mathcal{N}(0, \mathbf{\Sigma}_\eta)$ (respectively $\nu_k \propto \mathcal{N}(0, \mathbf{\Sigma}_c)$) the Gaussian noise term of the propagation (respectively measurement) equation.
With this model, the Kalman propagation step writes

$$\hat{\mathbf{X}}_{k+1|k} = f(\hat{\mathbf{X}}_k, \tilde{\mathbf{u}}_k, 0) \tag{3}$$

$$\mathbf{\Sigma_e}_{k+1|k} = \mathbf{\Phi}_k \mathbf{\Sigma_e}_k \mathbf{\Phi}_k^\mathsf{T} + \mathbf{G}_k \mathbf{\Sigma_u} \mathbf{G}_k^\mathsf{T} + \mathbf{C}_k \mathbf{\Sigma}_\eta \mathbf{C}_k^\mathsf{T}, \tag{4}$$

where $\tilde{\mathbf{u}}_k \propto \mathcal{N}(\mathbf{u}_k, \mathbf{\Sigma_u}_k)$ is a corrupted measurement of the input $\mathbf{u}_k$. Matrices $\mathbf{\Phi}_k$, $\mathbf{G}_k$, $\mathbf{C}_k$ are respectively the Jacobian matrices of the process function f with respect to the state, the input, and the stochastic input of the model. Denoting $\mathbf{e}_\mathrm{b} : \mathbf{e}, \mathbf{u}, \eta \mapsto f(\hat{\mathbf{X}}_k \boxplus \mathbf{e}, \mathbf{u}, \eta) \boxminus f(\hat{\mathbf{X}}_k, \tilde{\mathbf{u}}_k, 0)$, these matrices are formally defined by:

$$\mathbf{\Phi}_k = \left.\frac{\partial \mathbf{e}_\mathrm{b}}{\partial \mathbf{e}}\right|_{0,\tilde{\mathbf{u}}_k,0}, \mathbf{G}_k = \left.\frac{\partial \mathbf{e}_\mathrm{b}}{\partial \mathbf{u}}\right|_{0,\tilde{\mathbf{u}}_k,0}, \mathbf{C}_k = \left.\frac{\partial \mathbf{e}_\mathrm{b}}{\partial \eta}\right|_{0,\tilde{\mathbf{u}}_k,0} \tag{5}$$

The update step with measurement $\tilde{y}_{k+1}$ writes

$$\hat{\mathbf{X}}_{k+1} = \hat{\mathbf{X}}_k \boxplus \left(\mathbf{K}_{k+1}\left(\tilde{y}_{k+1} - h(\hat{\mathbf{X}}_{k+1|k}, 0)\right)\right) \tag{6}$$

$$\mathbf{\Sigma_e}_{k+1} = (\mathbf{I} - \mathbf{K}_{k+1}\mathbf{H}_{k+1})\mathbf{\Sigma_e}_{k+1|k} \tag{7}$$

where the linearized measurement matrix $\mathbf{H}_{k+1}$ is

$$\mathbf{H}_{k+1} = \left.\frac{\partial}{\partial \mathbf{e}}\left(h(\hat{\mathbf{X}}_{k+1|k} \boxplus \mathbf{e})\right)\right|_{\mathbf{e}=0} \tag{8}$$

and Kalman gain $\mathbf{K}_{k+1}$

$$\mathbf{K}_{k+1} = \mathbf{\Sigma}_{e_{k+1|k}} \mathbf{H}_{k+1}^\mathsf{T} \left(\mathbf{H}_{k+1} \mathbf{\Sigma}_{ek+1|k} \mathbf{H}_{k+1} + \mathbf{\Sigma}_c\right)^{-1}. \tag{9}$$

Note that the choice of the error is decoupled from how the mean estimates are kept in memory and propagated.

B. The MI-DR model and the MI-MSCKF

This section describes the MVINS filter of [3].
We will call $\mathtt{R}^w \in \mathrm{SO}(3)$ the attitude of the body frame (body to world convention), $\mathbf{v}^w, \mathbf{p}^w$ respectively the *world* velocity and position of the sensor, $\mathbf{B}^w$ the magnetic field at the current sensor position in world aligned frame, $\boldsymbol{\omega}^b, \mathbf{a}^b, \nabla\mathbf{B}^b$ respectively the instantaneous rotationnal velocity, acceleration and the gradient of magnetic field at sensor position in sensor-aligned frame. $[\boldsymbol{\omega}^b]_\times$ is the skew symmetric matrix associated to the 3D vector $\boldsymbol{\omega}^b$. We start from the continuous model of MI-DR:

$$\dot{\mathtt{R}}^w = \mathtt{R}^w [\boldsymbol{\omega}^b]_\times, \tag{10}$$

$$\dot{\mathbf{v}}^w = \mathtt{R}^w \mathbf{a}^b + \mathbf{g}^w; \quad \dot{\mathbf{p}}^w = \mathbf{v}^w, \tag{11}$$

$$\dot{\mathbf{B}}^w = \mathtt{R}^w \nabla\mathbf{B}^b \mathtt{R}^{w\mathsf{T}} \mathbf{v}^w. \tag{12}$$

The magnetic equation (12) is specific to MI-DR technique. The gradient of magnetic field $\nabla\mathbf{B}^b$ is measured from a network of magnetometer and renders the speed in body frame observable provided it is non-singular. In environments where compass capability is strongly altered by magnetic perturbations, such as indoor, this equation allows to efficiently correct accelerometers integration.

In [3] it is shown how to integrate this model to obtain the discrete dynamic model f:

$$\mathtt{R}_{k+1}^w = \mathtt{R}_k^w \widetilde{\Delta\mathtt{R}}_k, \tag{13}$$

$$\mathbf{v}_{k+1}^w = \mathbf{v}_k^w + \mathbf{g}^w \Delta t_k + \mathtt{R}_k^w \widetilde{\Delta\mathbf{v}}_k, \tag{14}$$

$$\mathbf{p}_{k+1}^w = \mathbf{p}_k^w + \mathbf{v}_k^w \Delta t_k + \frac{1}{2}\mathbf{g}^w \Delta t_k^2 + \mathtt{R}_k^w \widetilde{\Delta\mathbf{p}}_k, \tag{15}$$

$$\begin{aligned} \mathbf{B}_{k+1}^w = \mathbf{B}_k^w &+ \mathtt{R}_k \widetilde{\mathbf{\Delta B}}_{\mathbf{v};k} \mathtt{R}_k^\mathsf{T} \mathbf{v}_k^w + \mathtt{R}_k \widetilde{\mathbf{\Delta B}}_{\mathbf{g};k} \mathtt{R}_k^{w\mathsf{T}} \mathbf{g}^w \\ &+ \mathtt{R}_k \widetilde{\mathbf{\Delta B}}_{\mathbf{a};k}. \end{aligned} \tag{16}$$

Where we introduced abruptly the notations $\widetilde{\Delta\mathtt{R}}_k \in \mathrm{SO}(3)$, $\widetilde{\Delta\mathbf{v}}_k \in \mathbb{R}^3$, $\widetilde{\Delta\mathbf{p}}_k \in \mathbb{R}^3$ and $\widetilde{\mathbf{\Delta B}}_{\mathbf{v};k} \in \mathbb{R}^{3,3}$, $\widetilde{\mathbf{\Delta B}}_{\mathbf{g};k} \in \mathbb{R}^{3,3}$, $\widetilde{\mathbf{\Delta B}}_{\mathbf{a};k} \in \mathbb{R}^3$. They stand for integrated quantities that can

be computed from received measurements. Their exact expressions depend on a choice of integration method. They do not matter here (see [3] for more details).

Within the MSCKF technique, the state of the filter is augmented by a few past states, we define it:

$$\mathbf{X}_k = \left(\mathbf{X}_k^{\mathrm{sc}}, \mathbf{X}_k^{\mathrm{m}}\right) \tag{17}$$

where $\mathbf{X}_k^{\mathrm{m}} = \left(\mathtt{R}_k^w, \mathbf{p}_k^w, \mathbf{v}_k^w, \mathbf{B}_k^w\right)$ is the current state and $\mathbf{X}_k^{\mathrm{sc}} = \left(\mathtt{R}_{k-nc}^w, \mathbf{p}_{k-nc}^w, \ldots, \mathtt{R}_{k-1}^w, \mathbf{p}_{k-1}^w\right)$ is compound of nc "stochastic clones" [4] states (here past camera poses still in the sliding window). In [3], the error-state was simply defined with a substration between real state and estimated state, except for its rotational part where the distance $\log_{\mathrm{SO(3)}}(\mathtt{R}\hat{\mathtt{R}}^{\mathsf{T}})$ is used.

The MI-MSCKF filter uses two measurements. On one hand magnetic measurement is a direct measurement of the field state, in *body* frame: $h_{\mathrm{magn}}(\mathbf{X}_k) = \mathtt{R}^{w\mathsf{T}}\mathbf{B}_k^w$.

On the other hand, a visual measurement is built when a corner feature's track ends in the image stream. For a given track, the measurement equation collects the reprojection errors computed on all images taken at stochastic cloned instants:

$$h_{\mathrm{feat}}(\hat{\mathbf{X}}, \mathbf{l}) = \left[\pi\left(\mathtt{R}_{k\text{-}\mathrm{nc}}^w(\mathbf{l}-\mathbf{p}_k^w)\right)^{\mathsf{T}} \cdots \pi\left(\mathtt{R}_{k\text{-}1}^w(\mathbf{l}-\mathbf{p}_k^w)\right)^{\mathsf{T}}\right]^{\mathsf{T}} \tag{18}$$

where π is the projection function of the camera. The feature measurement is linearized according to:

$$h_{\mathrm{feat}}(\hat{\mathbf{X}} \boxplus \mathbf{e}, \mathbf{l}^* + \delta\mathbf{l}) \simeq h_{\mathrm{feat}}(\hat{\mathbf{X}}, \mathbf{l}^*) + \mathbf{F}\mathbf{e} + \mathbf{E}\delta\mathbf{l} \tag{19}$$

where $\mathbf{F}$ (resp. $\mathbf{E}$) is the Jacobian of h_{feat} with respect to the stochastic cloned part of the state (resp. to the landmark position parameters $\mathbf{l}$).[1] In this expression, the linearization point $\mathbf{l}^*$ comes from triangulation, assuming known past poses. The measurement matrix $\mathbf{H}_{feat}$ is computed by projecting this equation over the nullspace of $\mathbf{E}$ with a QR decomposition:

$$\mathbf{H}_{feat} = \mathbf{O}_0^{\mathsf{T}}\mathbf{F} \qquad \text{with } \mathbf{E} = [\mathbf{O}_1, \mathbf{O}_0]\begin{bmatrix}\mathbf{R}\\ 0\end{bmatrix}, \tag{20}$$

with $[\mathbf{O}_1, \mathbf{O}_0] \in \mathcal{O}$ and $\mathbf{R}$ an upper triangular matrix.

The projection steps (20) eliminates the need to augment the filter state with landmark positions. This is the big improvement of MSCKF-like algorithm compared to traditional EKF-SLAM filter : it saves both computation and memory requirements.

Note: In this section and hereafter, we choose to write all formulas without biases, albeit accelerometer and gyrometers biases are estimated by the filter in all experiments – as commonly done in VIO litterature. This simplify derivations, as bias estimation does not change the property we prove later on regarding the invariance to stochastic transform as noted in [7].

[1]The landmark parametrization can be chosen indifferently: we took as example 3D position of landmark in world frame but could be also 3D position in the coordinate frame of the first camera, inverse depth parametrization on first ray, etc.

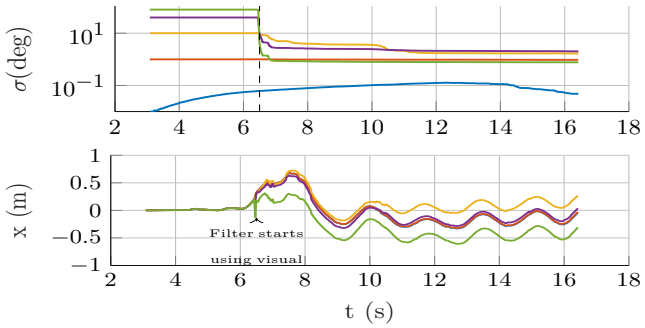


Fig. 1: False observability. Top : Covariance of the estimated heading on a real dataset with various initial covariance for the MI-MSCKF [3] (log scale). bottom: X position. Only magnetic update is employed during the first 6.5 seconds as an initialization. Heading's covariance decreases inconsistently as visual measurement are used, the position inconsistently depends on the initial heading covariance. Note: In the blue curve (with lowest initial uncertainty): the small increase is due to gyroscope noise and is hidden at the higher uncertainty levels and the decreasing at the end is not a sign of inconsistency as the uncertainty never become smaller than the initial value.

IV. Inconsistency and invariance

In this section, we first demonstrate an inconsistent behavior of MI-MSCKF, and then link this inconsistency to invariance properties this filter lacks.

A. An experiment on real data

Five instances of the MI-MSCKF of [3] were run with the same input data but different initialization mean and covariance values. The initialization values only differ from the initial uncertainty of the heading angle. Real data were used for this experiment, more precisely the first few seconds of dataset denoted Traj5 in [3]. The top curves of Fig. 1 depicts the uncertainty of the filter on the heading angle — measured by the corresponding diagonal value of the covariance. We observe that for high initial uncertainty, the heading uncertainty estimated by the filter ends up lower than the initial uncertainty. This is a sign of inconsistency, as, intuitively, the model does not provide information on the absolute heading. Interestingly, this drastic decrease of uncertainty co-occurs with the first use of visual measurements, proving these are a major source of inconsistency.

B. Lack of invariance to unobservable stochastic transform

In order to define consistency, we use in this work the definition of *unobservable stochastic transform* of an EKF which was introduced in [7]. In all this section we assume an EKF built on a non-linear error $\mathbf{e}$ and we let $\mathcal{T}_\theta$ be a transformation of the space defined as a 2-argument function $\mathcal{T}_\theta : (\mathcal{X}, \mathbb{R}^N) \to \mathcal{X}$ in which the second argument argument is seen as a stochastic input drawn from a centered Gaussian distribution of covariance $\mathbf{\Sigma}$. We also assume $\mathcal{T}_\theta$ is such that $\mathcal{T}_\theta(\cdot, 0)$ is the identity function on $\mathcal{X}$. The θ symbol can be seen as a parameter of a family of possible transform functions of interest.

Definition 1. (Stochastic transform of an EKF state) We call a *stochastic transform* of an EKF state at time k the following transformation:

$$\hat{\mathbf{X}}_k \mapsto \mathcal{T}_\theta(\hat{\mathbf{X}}_k, 0), \quad \mathbf{\Sigma}_{\mathbf{e}k} \mapsto \mathbf{M}\mathbf{\Sigma}_{\mathbf{e}k}\mathbf{M}^\mathsf{T} + \mathbf{N}\mathbf{\Sigma}\mathbf{N}^\mathsf{T}, \quad (21)$$

$$\text{with } \mathbf{M} = \frac{\partial}{\partial \mathbf{e}} \left(\mathcal{T}_\theta(\hat{\mathbf{X}} \boxplus \mathbf{e}, 0) \boxminus \mathcal{T}_\theta(\hat{\mathbf{X}}, 0) \right) \Big|_{\mathbf{e}=0} \quad (22)$$

$$\mathbf{N} = \frac{\partial}{\partial \eta} \left(\mathcal{T}_\theta(\hat{\mathbf{X}}, \eta) \boxminus \mathcal{T}_\theta(\hat{\mathbf{X}}, 0) \right) \Big|_{\eta=0} \quad (23)$$

By extension we will call $\mathcal{T}_\theta$ the stochastic transform (function).

Definition 2, introduced by [7], gives a meaning to the concept of invariance to a stochastic unobservable transform.[2]

Definition 2. (Invariance of an EKF output to unobservable stochastic transform) The EKF output is said to be invariant to an unobservable stochastic transformation if both following statements are true:

1) For all $\eta \in \mathbb{R}^N$, the stochastic transform $\mathcal{T}_\theta : (\mathcal{X}, \mathbb{R}^N) \to \mathcal{X}$ describes a unobservable transformation of the model/output on which the EKF is based on. I.e, if at time i we have $\mathbf{X}_i^a = \mathcal{T}_\theta(\mathbf{X}_i^b, \eta)$ then $\forall n > i,\ h(\mathbf{X}_n^a) = h(\mathbf{X}_n^b)$ (noting with indices n the quantities obtained by applying repeatedly the discrete transition function)
2) For any two estimate and covariance of the EKF at time i, say $(\hat{\mathbf{X}}_i^a, \mathbf{\Sigma}_{\mathbf{e}_i}^a)$ and $(\hat{\mathbf{X}}_i^b, \mathbf{\Sigma}_{\mathbf{e}_i}^b)$, so that b-quantities are computed from the stochastic transformation of a-quantities, we have equality of output sequence of the two instances of the filter based on respectively a- and b-quantities as initialization values: $\forall n > i,\ h(\hat{\mathbf{X}}_n^a) = h(\hat{\mathbf{X}}_n^b)$

Intuitively, this last definition states that, whatever are the unobservable quantities values initialized to, and whatever are *the initial uncertainty along the unobservable direction* in the initial covariance, the innovation sequence of the filter should not change: one property we would expect from a consistent filter. For MVINS, respecting Definition 2 means that the mean estimate sequence should not be modified by changing solely the initial heading uncertainty. This was proved wrong for the MI-MSCKF in previous section .

C. Two stochastic transforms of interest

This section exposes two properties that will be used in Section V-B in order to prove the invariance of a new filter to all unobservable stochastic transformations of its model. Sufficient conditions were introduced in [7] to prove by simple computation the invariance of a EKF for two particular types of stochastic transform functions : purely deterministic transform (verifying $\forall \eta, \mathbf{X}, \mathcal{T}_\theta(\mathbf{X}, \eta) = \mathcal{T}_\theta(\mathbf{X}, 0)$) and stochastic identity transform (verifying $\forall \mathbf{X}, \mathcal{T}_{I_d}(\mathbf{X}, 0) = \mathbf{X}$).

1) Invariance of an EKF to a deterministic transform:

Property 1. *Let $\mathcal{T}_\theta$ denotes a stochastic transform function. The output of an EKF is invariant under the deterministic part of this transform, $\mathbf{X}, \eta \mapsto \mathcal{T}_\theta(\mathbf{X}, 0)$, if: (i) the process function f commutes with $\mathcal{T}_\theta(\cdot, 0)$ and, (ii) there exists a constant invertible matrix $\mathbf{W}_\theta$ such that:*

$$\forall \mathbf{e}, \hat{\mathbf{X}}, \quad \mathcal{T}_\theta(\hat{\mathbf{X}} \boxplus \mathbf{e}, 0) = \mathcal{T}_\theta(\hat{\mathbf{X}}, 0) \boxplus \mathbf{W}_\theta \mathbf{e} \quad (24)$$

Proof insight: we consider an EKF built on the model with a choice of error verifying (24) and we assume that two such filters are running concurrently. The first starts from the initial estimate $(\hat{\mathbf{X}}_i, \mathbf{\Sigma}_{\hat{\mathbf{X}}_i})$ at time i. The second starts from the initial estimate $(\hat{\mathbf{Y}}_i, \mathbf{\Sigma}_{\hat{\mathbf{Y}}_i}) = (\mathcal{T}_\theta(\hat{\mathbf{X}}_i, 0), \mathbf{W}_\theta \mathbf{\Sigma}_{\hat{\mathbf{X}}_i} \mathbf{W}_\theta^\mathsf{T})$, a deterministic transform of the first one. It is possible to show by brute force calculus that after one propagation and one update, the two filter estimates and covariances are still related one to the other with the same unobservable deterministic transform. Conclusion can be drawn by induction. The condition (i) is necessary to the proof ([27, p.152]).

2) Invariance of an EKF to a family of stochastic identity transform:

Property 2. *Let $\mathcal{T}_{I_d}$ be a stochastic transform function fulfilling stochastic identity transform properties. The output of an EKF is invariant under $\mathcal{T}_{I_d}$ if:*

$$\forall n \text{ and } i \geq 0, \quad \mathbf{H}_{n+i+1} \Phi_{n+i} \Phi_{n+i-1} ... \Phi_i \mathbf{N}_i = 0 \quad (25)$$

Proof insight: Let us consider an EKF built on the model with an error verifying (25) and suppose two instances of it are started. The first one from estimate $(\hat{\mathbf{X}}_i, \mathbf{\Sigma}_{\hat{\mathbf{X}}_i})$ at time i. The second from estimate $(\hat{\mathbf{X}}_i, \mathbf{\Sigma}_{\hat{\mathbf{X}}_i} + \mathbf{N}_i \mathbf{\Sigma} \mathbf{N}_i^\mathsf{T})$, an unobservable stochastic transform of the first one. By calling the subsequent estimate of the second filter $\hat{\mathbf{Y}}_n$ and $\mathbf{\Sigma}_{\hat{\mathbf{Y}}_n}$, we can show by recursion that:

$$\forall n \geq i, \hat{\mathbf{Y}}_n = \hat{\mathbf{X}}_n \text{ and}$$
$$\mathbf{\Sigma}_{\hat{\mathbf{Y}}_n} = \mathbf{\Sigma}_{\hat{\mathbf{X}}_n} + \Phi_n \Phi_{n-1} ... \Phi_i \mathbf{N}_i \mathbf{\Sigma} \mathbf{N}_i^\mathsf{T} \Phi_i^\mathsf{T} ... \Phi_{n-1}^\mathsf{T} \Phi_n^\mathsf{T} \quad (26)$$

The first equality induces that the filter output is invariant to identity stochastic transform.

V. An invariant parametrization

A. Lie Group embedding of the MI-DR state

We borrow and adapt to our problem the invariant parametrization of IMU states originally presented in [21]. Despite the fact that the miraculous log-linear property of [21] does not hold for MVINS discrete model [3], we show in this paper that this invariant parameterization gives the filter a lower property of invariance to stochastic transformation.

[2]Note that while we have slightly rewritten [7] definitions to give them more context, Definition 2 remains exactly equivalent to their definition of unobservable stochastic transforms

[3]this stems from magnetic equations, even when disregarding biases; a simple calculus using [25, Theorem 2] proves it.

We base a new error definition on the following matrix Lie group embedding, named $SE_3(3)$ by the authors of [21], of the MI-DR state (stochastic clones part of the state is treated hereafter):

$$\mathbf{X}^{\mathrm{m}} = \begin{bmatrix} \mathtt{R}^w & \mathbf{v}^w & \mathbf{p}^w & \mathbf{B}^w \\ \mathbf{0}_{3,1} & & \mathbf{I}_{3,3} & \end{bmatrix} \tag{27}$$

$$(\mathbf{X}^{\mathrm{m}})^{-1} = \begin{bmatrix} \mathtt{R}^{w\mathsf{T}} & -\mathtt{R}^{w\mathsf{T}}\mathbf{v}^w & -\mathtt{R}^{w\mathsf{T}}\mathbf{p}^w & -\mathtt{R}^{w\mathsf{T}}\mathbf{B}^w \\ \mathbf{0}_{3,1} & & \mathbf{I}_{3,3} & \end{bmatrix} \tag{28}$$

$$\mathtt{R}^w \in \mathrm{SO}(3), \mathbf{v}^w, \mathbf{p}^w, \mathbf{B}^w \in \mathbb{R}^3,$$

and focus on the "right-invariant" error that writes:

$$\boldsymbol{\epsilon} = \mathbf{X}^{\mathrm{m}}(\hat{\mathbf{X}}^{\mathrm{m}})^{-1} = \tag{29}$$
$$\begin{bmatrix} \mathtt{R}^w\hat{\mathtt{R}}^{w\mathsf{T}} & \mathbf{v}^w - \mathtt{R}^w\hat{\mathtt{R}}^{w\mathsf{T}}\hat{\mathbf{v}}^w & \mathbf{p}^w - \mathtt{R}^w\hat{\mathtt{R}}^{w\mathsf{T}}\hat{\mathbf{p}}^w & \mathbf{B}^w - \mathtt{R}^w\hat{\mathtt{R}}^{w\mathsf{T}}\hat{\mathbf{B}}^w \\ \mathbf{0}_{3,1} & & \mathbf{I}_{3,3} & \end{bmatrix}$$

This error can be associated to a vector error through the vectorized matrix Lie group logarithm $\mathrm{Log}_{\mathbb{SE}_3(3)}(\epsilon)$. Its Lie algebra $\mathfrak{se}_3(3)$ is the set of matrices of the following form:

$$\mathbf{e}^{m\wedge} = \begin{bmatrix} [\mathbf{e}_\theta]_\times & \mathbf{e}_\mathbf{v} & \mathbf{e}_\mathbf{p} & \mathbf{e}_\mathbf{B} \\ \mathbf{0}_{3,3} & \mathbf{0}_{3,1} & \mathbf{0}_{3,1} & \mathbf{0}_{3,1} \end{bmatrix}, \quad \mathbf{e}^{m\wedge} \in \mathfrak{se}_3(3) \tag{30}$$

with $\mathbf{e}^m = \begin{bmatrix} \mathbf{e}_\theta^\mathsf{T} & \mathbf{e}_\mathbf{v}^\mathsf{T} & \mathbf{e}_\mathbf{p}^\mathsf{T} & \mathbf{e}_\mathbf{B}^\mathsf{T} \end{bmatrix}^\mathsf{T} \in \mathbb{R}^{12}$;

For concision, we will write elements dropping the trivial parts of the matrix and write the shortcuts $e^{\mathbf{e}_\mathtt{R}}$ for $\mathrm{Exp}_{\mathrm{SO}(3)}(\mathbf{e}_\mathtt{R})$, $\exp(\mathbf{e}^{m\wedge})$ for $\mathrm{Exp}_{\mathbb{SE}_3(3)}\left(\mathbf{e}^{m\wedge}\right)$. The exponential matrix operator on this matrix Lie group associates an element of its Lie algebra to an element of the group. One has in reduced notations:

$$\exp(\mathbf{e}^{m\wedge}) = \begin{pmatrix} e^{\mathbf{e}_\mathtt{R}} \\ \mathbf{J}_r(-\mathbf{e}_\mathtt{R})\mathbf{e}_\mathbf{v} \\ \mathbf{J}_r(-\mathbf{e}_\mathtt{R})\mathbf{e}_\mathbf{p} \\ \mathbf{J}_r(-\mathbf{e}_\mathtt{R})\mathbf{e}_\mathbf{B} \end{pmatrix}, \tag{31}$$

where $\mathbf{J}_r : \mathbb{R}^3 \to \mathbb{R}^{3\times3}$ is the *right jacobian of* SO(3) [28]:

$$\mathbf{J}_r(\theta) = \mathbf{I}_3 - \frac{1-\cos\|\theta\|}{\|\theta\|^2}[\theta]_\times + \frac{\|\theta\| - \sin\|\theta\|}{\|\theta\|^3}[\theta]_\times^2. \tag{32}$$

We use this exponential to define the error by the following $\boxplus$ and $\boxminus$ operators:

$$\mathbf{X}_k^{\mathrm{m}} = \hat{\mathbf{X}}_k^{\mathrm{m}} \boxplus \mathbf{e}_k^m = \exp(\mathbf{e}_k^{m\wedge})\hat{\mathbf{X}}_k \tag{33}$$

$$= \begin{pmatrix} e^{\mathbf{e}_\mathtt{R}}\hat{\mathtt{R}}^w \\ \mathbf{J}_r(-\mathbf{e}_\mathtt{R})\mathbf{e}_\mathbf{v} + e^{\mathbf{e}_\mathtt{R}}\hat{\mathbf{v}}^w \\ \mathbf{J}_r(-\mathbf{e}_\mathtt{R})\mathbf{e}_\mathbf{p} + e^{\mathbf{e}_\mathtt{R}}\hat{\mathbf{p}} \\ \mathbf{J}_r(-\mathbf{e}_\mathtt{R})\mathbf{e}_\mathbf{B} + e^{\mathbf{e}_\mathtt{R}}\hat{\mathbf{B}}^w \end{pmatrix} \tag{34}$$

$$\mathbf{e}_k^m = \mathbf{X}^{\mathrm{m}} \boxminus \hat{\mathbf{X}}_k^{\mathrm{m}} = \mathrm{Log}_{\mathbb{SE}_3(3)}\left(\mathbf{X}_k^{\mathrm{m}}(\hat{\mathbf{X}}_k^{\mathrm{m}})^{-1}\right) \tag{35}$$

$$= \begin{pmatrix} \mathbf{e}_\mathtt{R} \\ \mathbf{J}_r(-\mathbf{e}_\mathtt{R})^{-1}(\mathbf{v} - \mathtt{R}\hat{\mathtt{R}}^T\hat{\mathbf{v}}) \\ \mathbf{J}_r(-\mathbf{e}_\mathtt{R})^{-1}(\mathbf{p} - \mathtt{R}\hat{\mathtt{R}}^T\hat{\mathbf{p}}) \\ \mathbf{J}_r(-\mathbf{e}_\mathtt{R})^{-1}(\mathbf{B} - \mathtt{R}\hat{\mathtt{R}}^T\hat{\mathbf{B}}) \end{pmatrix} \in \mathbb{R}^{12} \tag{36}$$

$$\text{with } \mathbf{e}_\mathtt{R} = \mathrm{Log}_{\mathrm{SO}(3)}\left(\mathtt{R}\hat{\mathtt{R}}^T\right) \tag{37}$$

For the stochastically cloned part of the state, we use a left multiplication by the natural SE(3) exponential of the error as retraction operator for each stochastic cloned pose:

$$\mathrm{Exp}_{\mathrm{SE}(3)}\left([\mathbf{e}_\mathtt{R}^{\mathrm{sc}\mathsf{T}}, \mathbf{e}_\mathbf{p}^{\mathrm{sc}\mathsf{T}}]\right) = \begin{bmatrix} \mathbf{e}^{\mathbf{e}_\mathtt{R}^{\mathrm{sc}}} & \mathbf{J}_r(-\mathbf{e}_\mathtt{R}^{\mathrm{sc}})\mathbf{e}_\mathbf{p}^{\mathrm{sc}} \\ 0 & 1 \end{bmatrix} \tag{38}$$

The stochastic clone errors are concatenated together with $\mathbf{e}^m$ and respective retraction operators are used to define an error from which we build a new MVINS filter denoted RI-MI-MSCKF, where RI stand for "right-invariant".

Hereafter, we name $\mathcal{M}$ the Lie group structure on the state of the filter as the following direct product: $\mathcal{M} = \mathrm{SE}(3) \times \ldots \times \mathrm{SE}(3) \times \mathbb{SE}_3(3)$.

B. Proof of the invariance of this parametrization

In this section we prove the invariance to stochastic unobservable transform of the RI-MI-MSCKF with help from properties defined in Section IV-C.

1) Stochastic unobservable transform for MVINS *model:* In order to work with the observability properties of our model, we define formally the set of unobservable stochastic transform corresponding to composition of rotation around the gravity vector and global translation of world coordinates.

Definition 3. (Unobservable stochastic transform for MI-MSCKF model) We parametrize the family of unobservable stochastic transforms for the model in the following way:

$$\mathcal{T}(\mathbf{X}, \eta, \theta) \overset{\text{def}}{=} \begin{pmatrix} \cdots \\ e^{(\eta_1+\theta_1)\mathbf{g}_{\mathtt{R}^{sc}}} \\ e^{(\eta_1+\theta_1)\mathbf{g}_{\mathbf{p}^{sc}}} + \theta_{2:4} + \eta_{2:4} \\ \cdots \\ e^{(\eta_1+\theta_1)\mathbf{g}_\mathtt{R}} \\ e^{(\eta_1+\theta_1)\mathbf{g}_\mathbf{v}} \\ e^{(\eta_1+\theta_1)\mathbf{g}_\mathbf{p}} + \theta_{2:4} + \eta_{2:4} \\ e^{(\eta_1+\theta_1)\mathbf{g}_\mathbf{B}} \end{pmatrix} \tag{39}$$

with $\mathbf{X} \in \mathcal{M}, \eta \in \mathbb{R}^4, \theta \in \mathbb{R}^4$

Note that, we now put θ as a proper $\mathcal{T}$ argument function as we can parametrize stochastic transform of interests through $\mathbb{R}^4$. We can decompose each element of the family into one *stochastic identity transform* $\mathcal{T}(\hat{\mathbf{X}}, \eta, 0)$ and one *deterministic transform* $\mathcal{T}(\hat{\mathbf{X}}, 0, \theta)$ thanks to the following equality that can be easily verified for MVINS model:

$$\mathcal{T}(\hat{\mathbf{X}}, \eta, \theta) = \mathcal{T}\left(\mathcal{T}(\hat{\mathbf{X}}, 0, \theta), \eta, 0\right) = \mathcal{T}\left(\mathcal{T}(\hat{\mathbf{X}}, 0, \eta), 0, \theta\right). \tag{40}$$

So, to prove that the filter is invariant to all unobservable stochastic transforms, we first prove that it is invariant to all element's deterministic part in the family: $\mathbf{X}, \eta \mapsto \mathcal{T}(\mathbf{X}, 0, \theta)$ then prove that it is invariant to stochastic identity transform $\mathbf{X}, \eta \mapsto \mathcal{T}(\mathbf{X}, \eta, 0)$.

2) Property 1 verification: The condition (24) is verified for our parametrization with:

$$\mathbf{W}_D = \begin{bmatrix} \ddots & 0 & 0 & 0 \\ 0 & \mathbf{w}_D^{\mathrm{sc}_k} & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & \mathbf{w}_D^{\mathrm{m}} \end{bmatrix}, \tag{41}$$

$$\text{with } \mathbf{W}_D^{\mathrm{sc}_k} = \begin{bmatrix} e^{\theta_1 \mathbf{g}} & 0 \\ [\theta_{2:4}]_\times e^{\theta_1 \mathbf{g}} & e^{\theta_1 \mathbf{g}} \end{bmatrix},$$

$$\text{and } \mathbf{W}_D^{\mathrm{m}} = \begin{bmatrix} e^{\theta_1 \mathbf{g}} & 0 & 0 & 0 \\ 0 & e^{\theta_1 \mathbf{g}} & 0 & 0 \\ [\theta_{2:4}]_\times e^{\theta_1 \mathbf{g}} & 0 & e^{\theta_1 \mathbf{g}} & 0 \\ 0 & 0 & 0 & e^{\theta_1 \mathbf{g}} \end{bmatrix}$$

which proves the invariance to $\mathcal{T}(\cdot, 0, \theta)$. The proof consists mainly in calculus with SO(3) properties and is not presented here: it can be found in [27].

3) Property 2 verification: Note that, in the Property 2, the measurement $\mathbf{H}$s and transition Φs matrices are those used effectively by the filter, *ie.* linearized at the estimated values, different of the real unknown values. In contrast with [7], we work here directly on the MSCKF state for proving the property of invariance to unobservable stochastic transform, which is a more straightforward but unusual way to tackle the problem.

The outline of subsequent calculus is the following: first, we will compute $\mathbf{N}_i$, as in Definition 1. Secondly, we will compute the structure of the transition matrix Φ. We will be able to show that $\mathbf{N}_i$ is left unchanged by left multiplication with transition matrix Φ. We will show that for both magnetic and visual measurement update, the linearized measurement matrices $\mathbf{H}$ involved verifies Property 2. We will then conclude.

a) Computation of $\mathbf{N}_i$*:* Taking into account the stochastic cloned part of the state. The computation of $\mathbf{N}_i$ following Definition 1 leads to:

$$\mathbf{N}_i = \left[\begin{array}{cccc|cccc} \cdot & \mathbf{g}^\mathsf{T} & \mathbf{0}_3 & \cdot & \mathbf{g}^\mathsf{T} & \mathbf{0}_3 & \mathbf{0}_3 & \mathbf{0}_3 \\ \cdot & \mathbf{0}_3 & \mathbf{I}_3 & \cdot & \mathbf{0}_3 & \mathbf{0}_3 & \mathbf{I}_3 & \mathbf{0}_3 \end{array}\right]^\mathsf{T} \tag{42}$$

Remarkably, this matrix does not depend on the state estimate, thanks to the choice of the parametrization.

b) Structure of transition matrix Φ*:* We only need some part of the structure of Φ_k to prove our property.

In the MSCKF algorithm, the prediction step can also involve stochastic cloning: we write Φ in the following form:

$$\Phi_k = \begin{bmatrix} \Phi_k^{Sc1} & \Phi_k^{Sc2} \\ 0 & \Phi_k^{\mathrm{m}} \end{bmatrix} \text{ with } \begin{bmatrix} \Phi_k^{Sc1} & \Phi_k^{Sc2} \end{bmatrix} = \tag{43}$$

$$\begin{bmatrix} \mathbf{I}_{6\mathrm{nc}} & \mathbf{0} \end{bmatrix} \text{ or } \left[\begin{array}{cc|cccc} \mathbf{0} & \mathbf{I}_{6(\mathrm{nc}-1)} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{I}_3 & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{I}_3 & \mathbf{0} \end{array}\right] \in \mathbb{R}^{6\mathrm{nc},6\mathrm{nc}+12}$$

depending on if stochastic cloning occurs at time k while Φ_k^{m} structure writes:

$$\Phi_k^{\mathrm{m}} = \begin{bmatrix} \mathbf{I}_{3,3} & * & \mathbf{0}_{3,3} & * \\ \Delta t_k \, [\mathbf{g}]_\times & * & \mathbf{0}_{3,3} & * \\ \frac{\Delta t_k^2}{2} \, [\mathbf{g}]_\times & * & \mathbf{I}_{3,3} & * \\ \mathtt{R}_k \widetilde{\mathbf{\Delta B}}_{\mathbf{g};kk+1} \mathtt{R}_k^T \, [\mathbf{g}]_\times & * & \mathbf{0}_{3,3} & * \end{bmatrix} \tag{44}$$

With this structure, we verify easily that $\Phi_{k+1}\mathbf{N}_k = \mathbf{N}_k$ and that, by recursion, we have:

$$\forall n \text{ and } i \geq 0, \quad \Phi_{n+i}\Phi_{n+i-1}...\Phi_i \mathbf{N}_i = \mathbf{N}_i \tag{45}$$

This is very handy to prove condition (25) as it is now sufficient to show that $\mathbf{H}\mathbf{N}_i = 0$. We now prove it both for magnetic and visual measurement.

c) Magnetic update: The magnetic update is related directly to the current states :

$$h_{\mathrm{magn}}(\mathbf{X}_k) = \mathtt{R}_k^\mathsf{T} \mathbf{B}_k$$

Computing the first order approximation yields:

$$\begin{aligned} h_{\mathrm{magn}}(\mathbf{X}_k \boxplus \mathbf{e}) &= \mathtt{R}_k^\mathsf{T} e^{-\mathbf{e}_\mathtt{R}} \left(e^{\mathbf{e}_\mathtt{R}} \mathbf{B}_k + \mathbf{J}_r(-\mathbf{e}_\mathtt{R}) \mathbf{e}_\mathbf{B} \right) \\ &= h_{\mathrm{magn}}(\mathbf{X}_k) + \mathtt{R}_k^\mathsf{T} \mathbf{e}_\mathbf{B} + o(\|\mathbf{e}\|) \end{aligned}$$

Thus the measurement matrix to use is:

$$\mathbf{H}_{\mathrm{magn}_k} = \begin{bmatrix} \mathbf{0}_{3,6nc} & \mathbf{0}_{3,9} & \mathtt{R}_k^\mathsf{T} & \mathbf{0}_{3,6} \end{bmatrix}$$

And we have by simple computation: $\mathbf{H}_{\mathrm{magn}_k}\mathbf{N}_i = 0$.

This proves the condition (25) if we had only the magnetic measurement equation. At that point we have thus exhibited a pure magneto-inertial filter that fulfills invariance to stochastic transform. Such a filter was not presented before as far as we know.

d) Visual update: For visual measurements, showing the relation by a direct computation is more cumbersome because of the way the landmark position parameters are eliminated from the measurement function in the MSCKF. Instead, we are going to leverage the invariance of the reprojection function h_{feat} without having to compute $\mathbf{H}_{\mathrm{feat}}$ explicitly to show that $\mathbf{H}_{\mathrm{feat}}\mathbf{N}_i = 0$.

We assume here that landmarks are parameterized by their position $\mathbf{l} \in \mathbb{R}^3$ in *world* frame. We can write the following equality:

$$\begin{aligned} &\forall \hat{\mathbf{X}} \in \mathcal{M}, \eta \in \mathbb{R}^4, \mathbf{l} \in \mathbb{R}^3, \\ &\qquad h_{\mathrm{feat}}(\hat{\mathbf{X}}, \mathbf{l}) = h_{\mathrm{feat}}(\mathcal{T}(\hat{\mathbf{X}}, \eta, 0), e^{\eta_1 \mathbf{g}} \, \mathbf{l} + \eta_{2:4}) \end{aligned} \tag{46}$$

which merely expresses frame invariance of the calibrated monocular reconstruction problem.
Differentiation of the equality with respect to η gives:

$$\mathbf{F}\boldsymbol{N}_i + \mathbf{E}\partial_\eta(e^{\eta_1 \mathbf{g}}\mathbf{l} + \eta_{2:4}) = 0 \tag{47}$$

$$\mathbf{F}\boldsymbol{N}_i = \mathbf{E} \left[[\mathbf{l}]_\times \mathbf{g}, \mathbf{I}_3 \right]. \tag{48}$$

Starting from (20) and substituting previous equation one has

$$\mathbf{H}_{\mathrm{feat}}\mathbf{N}_i \overset{(20)}{=} \boldsymbol{O}_0^\mathsf{T}\mathbf{F}\mathbf{N}_i \overset{(48)}{=} \boldsymbol{O}_0^\mathsf{T}\mathbf{E} \left[[\mathbf{l}]_\times \mathbf{g}, \mathbf{I}_3 \right]. \tag{49}$$

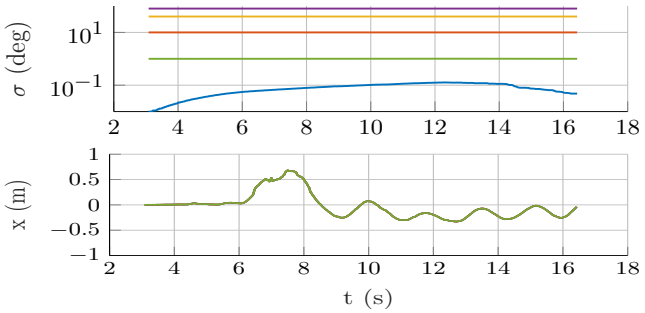


Fig. 2: top: heading uncertainty propagated by the invariant filter; bottom: the x position estimated by the filters. Both are with five different initial heading uncertainties. The five curves on bottom figure cannot be distinguished. With the new parametrization, the initial heading uncertainty does not influence the position estimate of the filter, and the estimation of heading uncertainty stays constant as expected. Please compare to Fig. 1 depicting the behavior of the filter of [3].

And by definition (20) of $\mathbf{O}_0$: $\mathbf{O}_0^\top \mathbf{E} = 0$ and finally the condition (25) $\mathbf{H}_{\text{feat}} \mathbf{N}_i = 0$ holds[4].

We thus have proven that the output of RI-MI-MSCKF for the MVINS system is invariant under $\mathcal{T}(\cdot, \cdot, 0)$.

Combined with the invariance to deterministic transform proved in (41) and using the equality noticed in (40), we prove that the RI-MI-MSCKF is invariant to any stochastic unobservable transform of the MVINS system.

Note: OC-EKF method can be applied to the MVINS problem [3]. The resulting estimator also verifies Definition 2. However, such approach is less theoretically satisfying and it is not clear that it does not introduce errors, as it is quite arbitrarily build from a Frobenius norm minimization. We name this estimator OC-MI-MSCKF in the next section.

VI. Experiment and comparison on real data

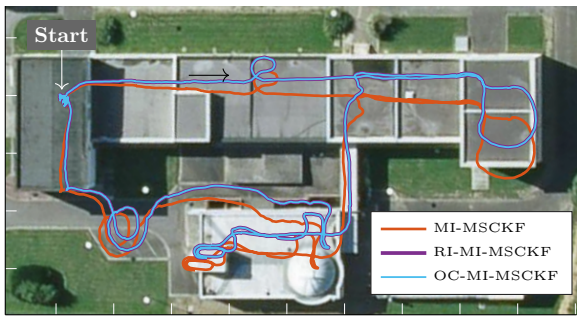
We implemented the two alternate versions, RI-MI-MSCKF and OC-MI-MSCKF of the MI-MSCKF and conducted a comparative study on real datasets described in [3]. In these implementations, we parameterize landmarks with an inverse depth in the first ray, and use Harris corner and KLT for features detection and tracking. Visual outliers are rejected with a two views gyro-aided RANSAC and a χ_2 gating test. Stochastic cloning occurs at least at 10Hz.

A. Improved behaviour with inaccurate heading

First, we verify on the data of Section IV that the RI-MI-MSCKF heading uncertainty and position outputs are in line with the proven invariance results (compare Fig. 2 with Fig. 1).

For further analysis, we run the three filters equally initialized with the *true* heading and a large heading covariance, while taking care to use the same parameters and measurements (we bypass non deterministic outlier rejection for this experiment). Thus, we expect the three

[4]if we were to use an *inverse depth in first ray* parametrization of features, the condition is also true, and can be demonstrated similarly. $\forall \hat{\mathbf{X}}, \eta, \mathbf{l}, h(\hat{\mathbf{X}}, \mathbf{l}) = h(\mathcal{T}(\hat{\mathbf{X}}, \eta, 0), d)$ so that by differentiating with respect to η one directly has: $0 = \mathbf{F}\boldsymbol{N}_i$.



1 tick = 10 meters

Fig. 3: Results on Traj1 when initialized with the correct heading value with a large covariance. OC- and MI- trajectories can hardly be distinguished. MI-MSCKF output is rotated.

%/deg	Traj1	Traj2	Traj3	Traj4	Traj5
MI-MSCKF	0.35/0.40	0.33/2.44	0.54/0.05	0.32/3.46	0.23/1.96
MI-MSCKF-LCOV	0.33/2.58	0.38/7.17	0.56/5.15	0.30/12.05	0.20/8.23
RI-MI-MSCKF	0.38/0.72	0.33/2.51	0.55/0.06	0.30/3.91	0.23/2.00
RI-MI-MSCKF-LCOV	0.38/0.72	0.34/2.51	0.55/0.06	0.30/3.91	0.23/2.00
OC-MI-MSCKF	0.38/0.72	0.33/2.49	0.55/0.01	0.30/3.85	0.23/1.99
OC-MI-MSCKF-LCOV	0.38/0.72	0.33/2.49	0.55/0.01	0.30/3.85	0.23/1.99

TABLE I: Final translational (% of trajectory length) and angular drift (degree) for various filters. The LCov suffix denotes try with a high covariance of the initial heading. For MI-MSCKF of [3], the angular drift strongly depends on the initial heading uncertainty, while it does not for the others. For this experiment, every filter uses exactly the same visual information, non-deterministic outlier detection was bypassed.

trajectories to register nicely on a satellite map on Fig. 3. We observe MI-MSCKF does not superimpose with the two others and the map: its output trajectory is rotated. In fact, because of lack of invariance of the MI-MSCKF, the heading gets corrupted near the start, while the heading covariance decreases as in Fig. 1.

Table I shows that this behavior of MI-MSCKF is general on the entire dataset of [3]. For the two other methods RI-MI-MSCKF and OC-MI-MSCKF both the final translational and rotational drifts are not strongly affected by the magnitude of initial heading covariance. This improvement opens the way to delayed heading initialization, a process that could be invoked in our indoor/outdoor dataset as outdoor the north direction is observable with magnetometers. In that specific scenario, it is likely that the standard MI-MSCKF would suffer badly from linearization error.

B. Better long-term behaviour

Traditionally in VIO, one want to solve for the relative pose with respect to initialization and the true value of heading is not of interest. In that case the filter is just initialized with a small heading covariance. We found the consistency improvement also helps in this case.

We compare the three filters on Traj1 and run them 10 times to sample the non deterministic outlier visual features detection. Fig. 4 displays a close-up of a selection of the estimated trajectories in the vicinity of the starting point, where all the trajectories also end. At the end of the trajectory, the sensor comes back to the starting point and is moved around while fixing a calibration

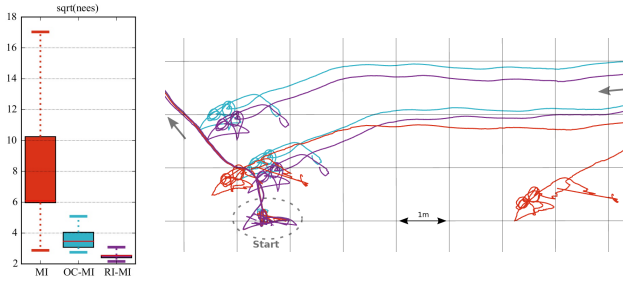


Fig. 4: Left: Normalized Estimation Error at the end of the trajectory for position and heading; statistics over the 10 run. Right: Zoom over the starting point of TRAJ1 for the three filters MI-MSCKF, RI-MI-MSCKF and OC-MI-MSCKF initialized with a zero heading uncertainty. Non deterministic outlier removal of visual features is activated: for each algorithm we display only the best and worst trajectories of ten runs.

pattern to improve the accuracy of estimating the final location. We observe two interesting facts from the final part of the trajectories. First, the MI-MSCKF trajectory is more chaotic than RI-MI-MSCKF and OC-MI-MSCKF which stay smooth the entire trajectory long. Secondly, Fig. 4 shows statistics of the Normalized Estimation Errors measured at the end of trajectory; a low value is a proof of consistency, it reveals that RI-MI-MSCKF is slightly more consistent than OC-MI-MSCKF. We also unexpectedly observe that the final locations of MI-MSCKF are more dispersed than the two others, MI-MSCKF seems to be more sensitive to the input visual features than the others. An explanation could be that consistency improves in return the χ_2 outlier gating test.

VII. Conclusion

We derived an original invariant parametrization of the MI-MSCKF and demonstrated its invariance to unobservable stochastic transform, a property naturally expected from an EKF when filtering with partially unobservable state. We have shown on real data that the proposed parametrization gives to the filter the ability to handle correctly the case of inaccurate heading initialization, as well as it significantly improves the smoothness of the estimation over long trajectories. According to our tests the proposed invariant of MI-MSCKF is more consistent than the observability constrained version. Also, we believe that the use of the Lie Group could be beneficial in the case of delayed or/and temporary observability of the heading. This specific case would deserve further analysis though.

References

[1] C.-I. Chesneau, M. Hillion, J.-F. Hullo, G. Thibault, and C. Prieur, "Improving magneto-inertial attitude and position estimation by means of magnetic heading observer," in *IPIN*, 2017.

[2] D. Caruso, A. Eudes, M. Sanfourche, D. Vissiere, and G. Le Besnerais, "Robust Indoor/Outdoor Navigation through Magneto-visual-inertial Optimization-based Estimation," in *IROS*, 2017.

[3] D. Caruso, A. Eudes, M. Sanfourche, D. Vissière, and G. Le Besnerais, "A Robust Indoor/Outdoor Navigation Filter Fusing Data from Vision and Magneto-Inertial Measurement Unit," *MDPI Sensors*, vol. 17, no. 12, p. 2795, 2017.

[4] A. I. Mourikis and S. I. Roumeliotis, "A Multi-State Constraint Kalman Filter for Vision-aided Inertial Navigation," in *ICRA*, 2007.

[5] J. A. Hesch, D. G. Kottas, S. L. Bowman, and S. I. Roumeliotis, "Camera-IMU-based localization: Observability analysis and consistency improvement," *IJRR*, vol. 33, no. 1, pp. 182–201, 2014.

[6] A. Barrau and S. Bonnabel, "An EKF-SLAM algorithm with consistency properties," *arXiv:1510.06263 [cs]*, 2015.

[7] T. Zhang, K. Wu, D. Su, S. Huang, and G. Dissanayake, "An Invariant-EKF VINS Algorithm for Improving Consistency," *IROS*, 2017.

[8] G. P. Huang, A. I. Mourikis, and S. I. Roumeliotis, "Analysis and improvement of the consistency of extended Kalman filter based SLAM," in *ICRA*, 2008.

[9] M. Kaess, H. Johannsson, R. Roberts, V. Ila, J. J. Leonard, and F. Dellaert, "iSAM2: Incremental smoothing and mapping using the Bayes tree," *IJRR*, vol. 31, no. 2, pp. 216–235, 2012.

[10] N. Keivan, A. Patron-Perez, and G. Sibley, "Asynchronous adaptive conditioning for visual-inertial slam," in *Experimental Robotics*, 2016.

[11] J. Engel, V. Koltun, and D. Cremers, "Direct Sparse Odometry," *TPAMI*, vol. 40, no. 3, pp. 611–625, 2018.

[12] G. P. Huang, A. I. Mourikis, and S. I. Roumeliotis, "A First-Estimates Jacobian EKF for Improving SLAM Consistency," in *Experimental Robotics*, ser. Springer Tracts in Advanced Robotics, 2009, pp. 373–382.

[13] Li and A. I. Mourikis, "Consistency of EKF-Based Visual-Inertial Odometry," Tech. Rep., 2012.

[14] G. Huang, Anastasios I. Mourikis, and Stergios I. Roumeliotis, "Observability-based Rules for Designing Consistent EKF SLAM Estimators," *IJRR*, vol. 29, no. 5, pp. 502–528, 2010.

[15] J. A. Hesch, D. G. Kottas, S. L. Bowman, and S. I. Roumeliotis, "Towards consistent vision-aided inertial navigation," in *Algorithmic Foundations of Robotics X*. Springer, 2013, pp. 559–574.

[16] J. A. Castellanos, J. Neira, and J. D. Tardós, "Limits to the consistency of EKF-based SLAM," *IFAC*, vol. 37, no. 8, pp. 716–721, 2004.

[17] J. Sola, "Consistency of the monocular ekf-slam algorithm for three different landmark parametrizations," in *ICRA*, 2010.

[18] G. Huang, M. Kaess, and J. J. Leonard, "Towards consistent visual-inertial navigation," in *ICRA*, 2014.

[19] P. Martin and E. Salaun, "Invariant observers for attitude and heading estimation from low-cost inertial and magnetic sensors," in *CDC*, 2007.

[20] S. Bonnabel, P. Martin, and P. Rouchon, "Non-linear observer on Lie Groups for left-invariant dynamics with right-equivariant output," in *IFAC*, 2008.

[21] S. Bonnabel and A. Barrau, "The Invariant Extended Kalman Filter as a Stable Observer," *IEEE Trans. Automat. Contr.*, vol. 62, no. 4, pp. 1797–1812, 2017.

[22] M. Brossard, S. Bonnabel, and A. Barrau, "Unscented Kalman Filtering on Lie Groups for Fusion of IMU and Monocular Vision," " 2018.

[23] R. Hartley, M. G. Jadidi, J. Grizzle, and R. M. Eustice, "Contact-aided invariant extended kalman filtering for legged robot state estimation," in *RSS*, 2018.

[24] A. Barrau, "Non-linear state error based extended Kalman filters with applications to navigation," Ph.D. dissert., 2015.

[25] A. Barrau and S. Bonnabel, "Linear observation systems on groups (I)," " 2017.

[26] C. Hertzberg, R. Wagner, U. Frese, and L. Schröder, "Integrating generic sensor fusion algorithms with sound state representations through encapsulation of manifolds," *Information Fusion*, vol. 14, no. 1, pp. 57–77, 2013.

[27] D. Caruso, "Improving Visual-Inertial Navigation Using Stationary Environmental Magnetic Disturbances," Ph.D. dissert., 2018.

[28] G. S. Chirikjian, *Stochastic Models, Information Theory, and Lie Groups, Volume 2: Analytic Methods and Modern Applications*. Springer Science & Business Media, 2011, vol. 2.