Multi-Signal Approaches for Repeated Sampling Schemes in Inertial Sensor Calibration

Gaetan Bakalli, Davide A. Cucci, Ahmed Radi, Naser El-Sheimy, Roberto Molinari, Olivier Scaillet and Stéphane Guerrier

***Abstract*—Inertial sensor calibration plays a progressively important role in many areas of research among which navigation engineering. By performing this task accurately, it is possible to significantly increase general navigation performance by correctly filtering out the deterministic and stochastic measurement errors that characterize such devices. While different techniques are available to model and remove the deterministic errors, there has been considerable research over the past years with respect to modelling the stochastic errors which have complex structures. In order to do the latter, different replicates of these error signals are collected and a model is identified and estimated based on one of these replicates. While this procedure has allowed to improve navigation performance, it has not yet taken advantage of the information coming from all the other replicates collected on the same sensor. However, it has been observed that there is often a change of error behaviour between replicates which can also be explained by different (constant) external conditions under which each replicate was taken. Whatever the reason for the difference between replicates, it appears that the model structure remains the same between replicates but the parameter values vary. Assuming the model structure has been identified, in this work we therefore consider and study the properties of different approaches that allow to combine the information from all replicates considering this phenomenon, confirming their validity both in simulation settings and also when applied to real inertial sensor error signals. By taking into account parameter variation between replicates, this work highlights how these approaches can improve the average navigation precision as well as obtain reliable estimates of the uncertainty of the navigation solution.**

***Index Terms*—Generalized Method of Wavelet Moments, Inertial Sensor Calibration, Stochastic Error, Extended Kalman Filter, Navigation**

I. Introduction

INERTIAL sensors are ubiquitous in modern navigation systems, with applications ranging from space missions, aviation and drones, to personal navigation in smartphones. They provide high-frequency and short-term precise information on the orientation and velocity change of the platform they are placed on. Inertial measurements are typically integrated with other sources to obtain estimates of the platform position and orientation in space. Examples are Global Navigation Satellite Systems within strap-down inertial navigation [1] and cameras for visual-inertial systems [2].

Inertial sensors, like any other sensor, have errors that are both deterministic and stochastic. Deterministic errors such as the stable parts of biases, scale factors and non-orthogonality of the axes can be pre-calibrated and removed from the measurements directly. The additive stochastic part of error can only be taken into account "on-flight" within the estimation process to serve two main purposes: i) estimation of the time-correlated part of those stochastic errors (to remove them from the measurements and improve navigation accuracy [3]) and, ii) estimation of uncertainty associated with the navigation states, such as position and orientation. This requires proper modeling of the stochastic errors of the sensors, often referred to as "stochastic calibration". This task is generally performed in a black-box fashion on a device-per-device basis, acquiring long series of static measurements which are composed by the stochastic error itself, plus constant terms such as gravity and the Earth rotation rate which can be easily removed. Stochastic calibration of inertial sensors has been widely studied in the last decades and has been mainly focused on identifying the model structure (that we denote as F) as well as the parameters (that we denote as $\boldsymbol{\theta}$) that characterize these models. While identifying the model structure F is not necessarily a major challenge since various visual tools and selection criteria are available (see [4]–[7]) as used for example in our case study in Sec. V, the focus has been on identifying (estimating) the parameter vector $\boldsymbol{\theta} \in \boldsymbol{\Theta} \subset \mathbb{R}^p$ which is indeed the focus of this work. More specifically, we assume that the general model structure F can be found within the class of composite models given by

$$X_t = \sum_{l=1}^{L} N_{t,l}, \tag{1}$$

where (X_t) represents the stochastic measurement error indexed over time by t and $(N_{t,l})$ represents the l^{th} component to this error generated by a model such as a white noise, a first-order autoregressive process (Gauss-Markov process), a moving-average process, a random walk or other models such as those discussed in [8] or [9] for example. Having identified the components (and hence F) through the previously mentioned tools, various methods are available to identify/estimate

G. Bakalli is with the QUANT Research Center, Emlyon Business School, Ecully, France (e-mail: bakalli@em-lyon.com).

D. Cucci is with the Geneva School of Economics and Management, University of Geneva, 1205, Switzerland (e-mail: davide.cucci@unige.ch).

A. Radi is with the Technical Researches Center, Cairo, Egypt (e-mail: ahmed.elboraee@ucalgary.ca).

N. El-Sheimy is with the Department of Geomatics Engineering, University of Calgary, Calgary, Alberta T2N 1N4, Canada (e-mail: elsheimy@ucalgary.ca).

R. Molinari is with the Department of Mathematics & Statistics, Auburn University, Auburn, AL 36849, USA (e-mail: robmolinari@auburn.edu).

O. Scaillet is with the Geneva Finance Research Institute, University of Geneva and Swiss Finance Institute, Geneva 1211, Switzerland (e-mail: olivier.scaillet@unige.ch).

S. Guerrier is with the Faculty of Science & Geneva School of Economics and Management, University of Geneva, 1205, Switzerland. (e-mail: stephane.guerrier@unige.ch).

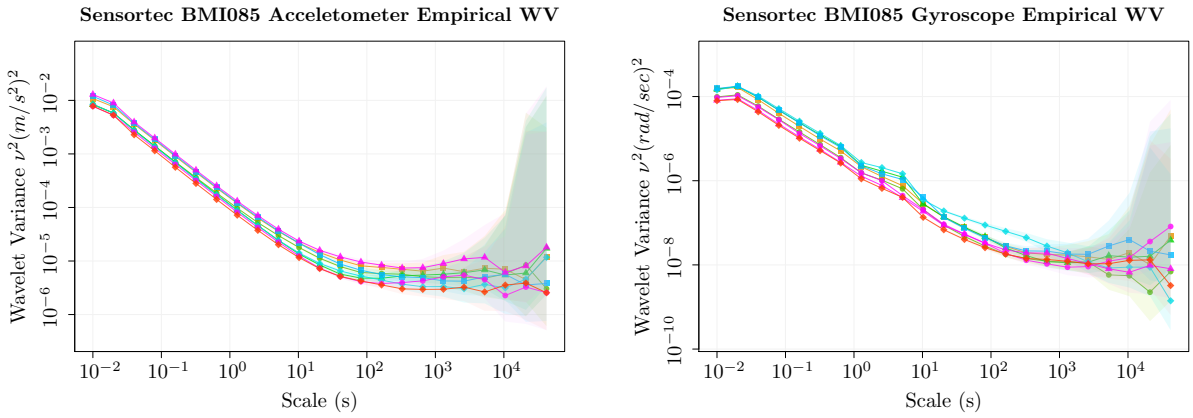

FIG. 1: *Empirical WV (plain doted line) coming from 8 replicates of Bosch Sensortec BMI085 MEMS IMU accelerometer (left plot) and gyroscope (right plot), with their respective* 95% *confidence intervals (shaded areas).*

the parameter $\boldsymbol{\theta}$ that characterize them, going from power spectral density analysis [10], [11] to the correlation of filtered sensor outputs [12]. The majority of these methods aim at decomposing these stochastic signals and/or performing system identification procedures to model them [13], [14]. The most commonly employed techniques are, for example, those based on Maximum-Likelihood Estimation [15], [16] (hereinafter MLE) or methods related to the likelihood setting such as Bayesian-type approaches [5], [17] (where model and parameter identification is performed jointly) as well as the Allan variance [18], [19] (hereinafter AV). Having been initially conceived for the characterization of phase and frequency instability of precision oscillators, the AV approach consists in multiple separate regressions on the linear segments of the AV plots in order to recover the underlying parameters of interest for the stochastic error signal and represents the de-facto standard for inertial sensor stochastic modeling [4], [11]. However, the AV plot is a graphical device that requires a manual inspection and is consequently sensitive to the user's proficiency as well as being burdened with many theoretical limitations including significant (asymptotic) bias in the estimated parameters of the postulated stochastic model [8], [20]. The alternative likelihood-based techniques, on the other hand, suffer from important computational and numerical issues due to the need to estimate the underlying states using a (Extended) Kalman-Filter and then to evaluate the likelihood which can be computationally cumbersome (or intractable) for the size of calibration signals [21].

To overcome the limitations of the AV as well as the important computational limitations of MLE and likelihood-type techniques, the Generalized Method of Wavelet Moments (GMWM) was proposed in [22] and makes use of the quantity called Wavelet Variance (WV) that, in specific settings, is equivalent to the AV up to a constant. As described more formally in the next section, using a matching technique, the WV is of reasonable dimensions even for large signals and allows to easily recover the parameters $\boldsymbol{\theta}$ of the postulated/identified stochastic model F providing a statistically appropriate and computationally feasible technique for stochastic calibration of inertial sensors. However, while the GMWM has improved the task of stochastic calibration for Inertial Measurement Units (IMUs), it still relies on the common approach to calibration which consists in modelling and removing stochastic errors measured on a single experimental run (replicate) of IMU measurements in static conditions. Although this approach still remains valid, it is also common for IMU calibration procedures to perform several independent experimental runs on the same sensor from which the single signal for calibration is chosen. The latter choice however remains somewhat random aside from visually assessing the behavior of the single replicates thereby exposing oneself to the risk of picking a signal that would not generalize well. More importantly, this approach does not consider the information provided by the different replicates which can contribute to improving the modelling process in terms of estimation (and consequent navigation) accuracy. To address this need, [23] and subsequently [24] underlined how the different replicates are indeed important to perform a comprehensive and accurate estimation procedure when assuming a *fixed* parameter $\boldsymbol{\theta}_0$ and put forward some techniques for this purpose.

While delivering solutions for the above setting, [23] and [24] nevertheless also underline that these replicates need to be used with caution due to possible changes in behavior between individual signal replicates. Using a WV representation of these signals, Fig. 1 provides an example of this behavior on eight independent recordings in static conditions coming from a Bosch Sensortec BMI085 MEMS IMU accelerometer and gyroscope. It can be seen how the shape of the WV remains roughly the same between replicates (i.e., the underlying model structure F remains constant) but their values differ significantly, mainly over the first scales, since their confidence intervals do not overlap thereby indicating difference beyond simple probabilistic randomness of the WV estimator. The latter conclusion was reached in various settings (as done in Sec. V) also through a statistical test put forward in [23] that checks the null hypothesis that all replicates come

from a model with fixed parameter $\boldsymbol{\theta}_0$ (and hence that the differences between the WV of each replicate are simply due to sampling randomness). Building on this observation, [24] defined this setting as "*near-stationary*" where, instead of considering a stochastic model characterized by a *fixed* parameter vector $\boldsymbol{\theta}_0$, they postulate that the parameters of this model are independent random variables $\boldsymbol{\vartheta}$ that follow a certain stationary probability distribution G. Aside from requiring it to be stationary, [23] and [24] leave the distribution G unspecified by choice since, while explaining possible parameter variation due to internal sensor characteristics, it can also represent the change in parameter values due to observed and/or unobserved external factors during each calibration run, often making the specification of the distribution G extremely complicated.

The solutions proposed in [23] and [24] for the *near-stationary* setting consider estimators that deliver a single estimate (say $\hat{\boldsymbol{\theta}}$) which can be considered somehow "representative" of the random variable $\boldsymbol{\vartheta}$ and that can ideally improve the average navigation accuracy. Indeed, these solutions are appropriate also in the case where all replicates come from a model F with *fixed* parameter $\boldsymbol{\theta}_0$ (i.e. the common setting where G is a point-mass distribution). However these solutions have not been formally studied in this new framework and do not include an additional solution that is put forward in this work. Indeed, assuming the model structure F is identified, this work intends to clearly define the statistical properties of these solutions in the *near-stationary* framework to compare them and identify exactly what theoretical values they target and consequently understand if and how they can improve navigation accuracy compared to a the standard calibration process based on a single replicate.

To present and discuss the proposed approaches and results, this paper is organized as follows. Sec. III formally defines the near-stationary framework and discusses the properties of the different approaches considered for multi-signal calibration. These results are necessary to obtain reliable statistical estimates for navigation purposes and are confirmed in Sec. IV which studies the finite sample performance of the three proposed approaches in a controlled simulation setting. Sec. V presents a case study on real-world inertial sensor calibration error signals which shows how the proposed approaches can generally improve the navigation performance with respect to the current setting where only one replicate is used to calibrate the inertial sensors and feed the navigation filter. Finally, Sec. VI concludes.

II. List of abbreviations

AGMWM	Average Generalized Method of Wavelet Moments
ANESS	Average Normalised Estimation Error Squared
AR1	Auto-Regressive Process of order 1
AV	Allan Variance
AWV	Average Wavelet Variance Estimator
EKF	Extended Kalman Filter
GMWM	Generalized Method of Wavelet Moments
IMU	Inertial Measurement Unit
MLE	Maximum Likelihood Estimator
MS-GMWM	Multi-Signal Generalized Method of Wavelet Moments
MODWT	Maximum Overlap Discrete Wavelet Transform
RW	Random Walk
UAV	Unmanned Aerial Vehicle
WN	White Noise
WV	Wavelet Variance

III. Multi-Signal Calibration

In the following sections, we formally describe and study the solutions, including those put forward in [23] and [24], which are all a direct extension of the GMWM. As mentioned, the latter is currently employed, among others, for sensor calibration on a *single* stochastic error signal issued from an inertial sensor calibration session (see, e.g., [22], [25]). Indeed, in order to estimate the assumed *fixed* parameter vector $\boldsymbol{\theta}_0$ that characterizes the model underlying the stochastic error the GMWM is defined as follows:

$$\tilde{\boldsymbol{\theta}} := \underset{\boldsymbol{\theta}\in\boldsymbol{\Theta}}{\operatorname{argmin}} \|\hat{\boldsymbol{\nu}} - \boldsymbol{\nu}(\boldsymbol{\theta})\|^2_{\boldsymbol{\Omega}}, \tag{2}$$

where, with $\mathbf{Z} \in \mathbb{R}^J$, we have that $\|\mathbf{Z}\|^2_{\boldsymbol{\Omega}} := \mathbf{Z}^\intercal\boldsymbol{\Omega}\mathbf{Z}$. In addition, $\hat{\boldsymbol{\nu}} \in \mathbb{R}^J_+$ represents the unbiased Maximum Overlap Discrete Wavelet Transform (MODWT) WV estimator proposed in [26] estimated on the single error signal issued from the calibration session and defined as

$$\hat{\nu}_j := \frac{1}{M_j}\sum_{t=1}^{M_j} W^2_{j,t} \tag{3}$$

where $(W_{j,t})$ represents the wavelet coefficients for scale j and M_j represents the number of coefficients issued from the MODWT at this scale. Finally, $\boldsymbol{\nu}(\boldsymbol{\theta}) \in \mathbb{R}^J_+$ represents the theoretical WV implied by the parametric model $F_{\boldsymbol{\theta}}$ and $\boldsymbol{\Omega}$ is a positive definite weighting matrix chosen in a suitable way, (see, e.g., [25] and following sections for more details). This approach therefore takes advantage of the relationship between the spectral density function and the WV [26] that allows every model $F_{\boldsymbol{\theta}}$ to have a (known) representation in the form $\boldsymbol{\nu}(\boldsymbol{\theta})$ [27]. For example, the j^{th}-level Haar WV for a white noise process with parameter σ^2 is given by $\nu^2_j(\sigma^2) = \sigma^2/2^j$, highlighting how the WV is a linear function of the parameter for this specific process. In this optic, the GMWM aims to find the value of $\boldsymbol{\theta}$ that makes $\boldsymbol{\nu}(\boldsymbol{\theta})$ as close as possible (under the L_2-metric) to the empirical WV $\hat{\boldsymbol{\nu}}$ and, when the theoretical WV is a linear function of the parameters, it has an explicit solution as shown in [20].

A. Near-Stationary Framework

Compared to the setting where a single error signal is considered, we now consider the case of $K > 1$ replicates from the same IMU in static conditions. Ideally, each signal, indexed by i and with length T_i, is issued from the same stochastic error model with the same fixed parameter values (i.e., $\boldsymbol{\theta}_i = \boldsymbol{\theta}_0, \forall i$)

which are specific to the sensor of interest. However, as discussed in the introduction and following the example given in Fig. 1, it would appear that the model structure F remains constant while the parameters characterizing it change between replicates. We therefore assume that there exists an independent sequence of random variables $\boldsymbol{\vartheta}_i$ (for $i = 1, \dots, K$), with associated probability distribution G defined over a compact set $\boldsymbol{\Theta}$, which we refer to as an *internal sensor model* and we define the processes generated by the sensor as *near-stationary* processes.

Assuming that all deterministic calibration has removed the corresponding errors (e.g. axis non-orthogonalities, etc.), let us define the i^{th} stochastic error signal as

$$(X_t^{(i)}) \sim F_{\boldsymbol{\vartheta}_i}, \quad t = 1, \dots, T_i$$

and

$$\boldsymbol{\vartheta}_i \overset{iid}{\sim} G, \quad \boldsymbol{\vartheta}_i \in \boldsymbol{\Theta} \subset \mathbb{R}^p,$$

As in the GMWM setting, the model $F_{\boldsymbol{\vartheta}_i}$ therefore represents the stochastic process governing the dependence structure over time during the i^{th} calibration session, where the distribution of the innovation sequence is left unspecified. With this setting in mind, we denote the MODWT estimator of WV as $\hat{\boldsymbol{\nu}}_i \in \mathbb{R}_+^J$ where J is a *fixed* integer representing the chosen number of WV scales such that $p \leq J \leq \min_i J_i$ where $J_i \in \mathbb{N}^+$ represents the number of WV scales for the i^{th} signal. It must be noticed that now the estimator $\hat{\boldsymbol{\nu}}_i$ does not target a general fixed WV $\boldsymbol{\nu}(\boldsymbol{\theta}_0)$ but aims to estimate the WV implied by the random parameter vector that generated the i^{th} replicate, i.e., $\boldsymbol{\nu}(\boldsymbol{\vartheta}_i)$.

Considering this new stochastic framework, it would be unreasonable to use the parameter vector estimated on the i^{th} signal to predict the general measurement error of a future signal. As a consequence, it would be more appropriate to define a *fixed* parameter vector that adequately represents and predicts the behaviour of all possible signals issued from the stochastic framework. In order to do so, we adopt the parameter notation from the standard setting and provide a new definition of $\boldsymbol{\theta}_0$ as follows:

$$\boldsymbol{\theta}_0 := \underset{\boldsymbol{\theta} \in \boldsymbol{\Theta}}{\operatorname{argmin}} \, Q\left(\boldsymbol{\theta}\right),$$

where

$$Q\left(\boldsymbol{\theta}\right) := \mathbb{E}\left[\|\boldsymbol{\nu}(\boldsymbol{\vartheta}_i) - \boldsymbol{\nu}(\boldsymbol{\theta})\|_{\boldsymbol{\Omega}}^2\right], \tag{4}$$

with $\mathbb{E}[\cdot]$ denoting the expectation under the distribution G, $\boldsymbol{\nu}(\boldsymbol{\theta})$ representing the theoretical WV implied by the stochastic model evaluated at the fixed parameter vector $\boldsymbol{\theta}$ and $\boldsymbol{\Omega}$ denoting a positive definite weighting matrix. The criterion in (4) is defined as an expectation under the distribution G whose form however, as underlined earlier, does not need to be specified if using the solutions considered in this work. With respect to the weighting matrix, for example, one can choose a *fixed* positive definite matrix for $\boldsymbol{\Omega}$, that we denote as $\boldsymbol{\Omega}_0$, which gives certain weights to how each scale of WV contributes to the estimation process (e.g one possible choice would be the identity matrix which gives equal weights to all scales). If one does not have the value of the elements for the chosen matrix, an estimator of the latter matrix (denoted as $\widehat{\boldsymbol{\Omega}}$) can be used. As long as this matrix is positive definite and assuming identifiability of the function $\boldsymbol{\nu}(\cdot)$, the criterion in (4) is always minimized in $\boldsymbol{\theta}_0$. In an estimation setting, the choice of $\boldsymbol{\Omega}$ is usually limited to minimizing the asymptotic variance which is achieved by choosing $\boldsymbol{\Omega}_0 := \mathbf{V}^{-1}$, where $\mathbf{V}$ is the asymptotic covariance matrix of the estimated WV (see [22], [25]), although a simple diagonal matrix (such as the identity) can often be more than sufficient in practice. With this in mind, the criterion (or loss/objective function) in (4) is an extension of the GMWM objective function which takes into account the internal sensor model G. Indeed, in the case of G being a point-mass distribution, the criterion in (4) simply goes back to representing the standard GMWM criterion in (2). The logic behind choosing this criterion therefore consists in finding a fixed parameter vector $\boldsymbol{\theta}_0$ that minimizes the expected squared-loss between the WV implied by the latter parameter and the WV implied by all possible values of the (parameter) random variable $\boldsymbol{\vartheta}_i$.

B. Multi-Signal Approaches

Given that we cannot directly observe the criterion in (4), we need to consider estimators for this quantity, among which those put forward in [23] and [24] whose finite sample performance was investigated through preliminary simulations and applied studies. These solutions are potential estimators for the quantity of interest $\boldsymbol{\theta}_0$ but, as shown further on, have different properties and actually turn out to be the same under specific or more general circumstances.

However, compared to the solutions put forward in [23] and [24], we define a more general setting where we can assign weights to the information coming from each replicate. More specifically, we define the weights that characterize the studied solutions as follows:

$$w_i := d_i \frac{T_i}{\sum_{k=1}^K T_k},$$

where d_i is a signal-specific constant defined by the user to give more weight to certain signals based on prior knowledge (one would however commonly choose $d_i = 1$ for all i). Based on this definition, conditioned on the choice of d_i, these weights are larger for longer signals therefore giving more weight to those signals that carry more information. We assume that the weights w_i are such that

$$\sum_{i=1}^K w_i = 1, \; w_i \geq 0 \text{ and } K \sum_{i=1}^K w_i^2 \to c, \tag{5}$$

where c is a positive constant. Moreover, we assume that for K sufficiently large there exists a constant M such that $Kw_i \leq M$. These requirements on the weights are mild and are satisfied for most reasonable choices of the constants d_i, and lengths T_i. For example, in the simple situation where $w_i = 1/K$ for all i these requirements are trivially satisfied.

Considering the multiple signal recording setting formalized in the previous paragraphs, the goal of the methods studied in this work is to combine the information from the different signals in an optimal (weighted) manner. The first and most straightforward way to do so would be to take a simple

weighted average of the GMWM estimators issued from the individual signals (we refer to this estimator as the Average GMWM (AGMWM) which was suggested in [23]). More formally, this estimator is defined as follows:

$$\hat{\boldsymbol{\theta}}^{\circ} := \sum_{i=1}^{K} w_i \tilde{\boldsymbol{\vartheta}}_i, \tag{6}$$

where

$$\tilde{\boldsymbol{\vartheta}}_i := \underset{\boldsymbol{\vartheta}\in\boldsymbol{\Theta}}{\operatorname{argmin}} \|\hat{\boldsymbol{\nu}}_i - \boldsymbol{\nu}(\boldsymbol{\vartheta})\|^2_{\boldsymbol{\Omega}},$$

are the individual GMWM parameter estimates for each signal.

The second estimator is new and we call it the Average WV (AWV) estimator which is defined as follows:

$$\hat{\boldsymbol{\theta}}^{\dagger} := \underset{\boldsymbol{\theta}\in\boldsymbol{\Theta}}{\operatorname{argmin}}\, \widehat{Q}^{\dagger}(\boldsymbol{\theta}), \tag{7}$$

where

$$\widehat{Q}^{\dagger}(\boldsymbol{\theta}) := \| \sum_{i=1}^{K} w_i \hat{\boldsymbol{\nu}}_i - \boldsymbol{\nu}(\boldsymbol{\theta})\|^2_{\boldsymbol{\Omega}}.$$

The idea behind this estimator is to replicate the structure of the GMWM estimator and, instead of considering a single estimate of the WV, we take the weighted average of the individual estimated WV. The objective function defining this estimator also resembles the criterion given in (4) and, as we will see further on, indeed targets this criterion.

The final estimator we study is the weighted version of the estimator defined in [23] and [24] and is given by the solution to the objective function resulting from the weighted average of the individual GMWM objective functions. More specifically, this estimator, referred to as the Multi-Signal GMWM (MS-GMWM), is defined as

$$\hat{\boldsymbol{\theta}} := \underset{\boldsymbol{\theta}\in\boldsymbol{\Theta}}{\operatorname{argmin}}\, \widehat{Q}(\boldsymbol{\theta}), \tag{8}$$

where

$$\widehat{Q}(\boldsymbol{\theta}) := \sum_{i=1}^{K} w_i \|\hat{\boldsymbol{\nu}}_i - \boldsymbol{\nu}(\boldsymbol{\theta})\|^2_{\boldsymbol{\Omega}}\,.$$

This estimator is therefore the result of the minimization of a direct estimator of the criterion in (4). Indeed, the empirical WV $\hat{\boldsymbol{\nu}}_i$ is an estimator for the theoretical quantity $\boldsymbol{\nu}(\boldsymbol{\vartheta}_i)$ while the weighted sum over the K signals is aimed at estimating the theoretical expectation $\mathbb{E}[\cdot]$ under the internal sensor model G.

C. Statistical Properties

Having formally defined the methods of interest for the problem at hand, we now lay out a series of assumptions that are necessary to define the asymptotic properties of these estimators. For this reason, we also define $\mathbf{H}(\boldsymbol{\theta}) := \mathbf{A}(\boldsymbol{\theta})^{\intercal}\boldsymbol{\Omega}\mathbf{A}(\boldsymbol{\theta})$ where

$$\mathbf{A}(\boldsymbol{\theta}) := \frac{\partial}{\partial \boldsymbol{\vartheta}^{\top}} \boldsymbol{\nu}(\boldsymbol{\vartheta})\Big|_{\boldsymbol{\vartheta}=\boldsymbol{\theta}}.$$

Moreover, we will use $\xrightarrow{P}$ to denote convergence in probability and $\xrightarrow{\mathcal{D}}$ to denote convergence in distribution.

ASSUMPTION A (Parameter Space): *$\boldsymbol{\theta}_0$ is an interior point of the set $\boldsymbol{\Theta}$ which is compact.*

ASSUMPTION B (Theoretical WV): *The theoretical WV is such that:*

- *$\boldsymbol{\nu}(\boldsymbol{\theta})$ is continuously differentiable $\forall \boldsymbol{\theta} \in \boldsymbol{\Theta}$;*
- *$\boldsymbol{\nu}(\boldsymbol{\theta}_1) = \boldsymbol{\nu}(\boldsymbol{\theta}_2)$ if and only if $\boldsymbol{\theta}_1 = \boldsymbol{\theta}_2$;*
- *$\mathbf{H}(\boldsymbol{\theta}_0)$ exists and is non-singular.*

ASSUMPTION C (Asymptotics):

$$\sqrt{T_i}(\hat{\boldsymbol{\nu}}_i - \boldsymbol{\nu}(\boldsymbol{\vartheta}_i)) \xrightarrow{\mathcal{D}} \mathcal{N}\left(\mathbf{0}, \mathbf{V}(\boldsymbol{\vartheta}_i)\right),$$

conditionally on $\boldsymbol{\vartheta}_i$, where $\mathbf{V}(\boldsymbol{\vartheta}_i) := \lim_{T_i\to\infty} T_i \operatorname{var}(\hat{\boldsymbol{\nu}}_i)$ is continuous in $\boldsymbol{\vartheta}_i$. Moreover, if $\boldsymbol{\Omega}_0$ is estimated by $\widehat{\boldsymbol{\Omega}}$, then we have that

$$||\widehat{\boldsymbol{\Omega}} - \boldsymbol{\Omega}_0||_S \xrightarrow{P} 0,$$

where $||\cdot||_S$ denotes the matrix spectral norm, with the Frobenius norm of $\boldsymbol{\Omega}_0$ being bounded.

Assumption A is a standard regularity condition that ensures that certain quantities are bounded and to allow convergence (however it can be partly relaxed depending on the model of interest). Assumption B ensures that (i) $\boldsymbol{\nu}(\cdot)$ is differentiable (in order to perform expansions); (ii) $\boldsymbol{\nu}(\cdot)$ is injective (in order to have identifiability); and (iii) certain quantities from these expansions exist in order to prove consistency and asymptotic normality of the estimators. Finally, Assumption C requires asymptotic normality of the WV estimator (which was proven under different conditions, see e.g. [25] or [28]), the consistency of $\widehat{\boldsymbol{\Omega}}$ (if an estimator is actually chosen for the weighting matrix $\boldsymbol{\Omega}$) as well as the boundeness of $\boldsymbol{\Omega}_0$.

These assumptions are required to prove results on consistency and asymptotic normality of the multi-signal approaches described earlier which can provide insight to convergence rates of these approaches as well as justify the use of time-dependent bootstrap methods to deliver adequate uncertainty quantification for each of them. Denoting $T := \min_i T_i$, we can now study the first of the considered estimators, namely the AGMWM.

THEOREM 1: *Under Assumptions A to C and letting $K, T \to \infty$, we have that*

$$\hat{\boldsymbol{\theta}}^{\circ} \xrightarrow{P} \mathbb{E}[\boldsymbol{\vartheta}_i] \neq \boldsymbol{\theta}_0$$

Proof of Theorem 1 is given in Appendix A. From this result, it can be noticed how the AGMWM targets the expected value of the internal sensor model G which does not necessarily correspond to the desired value $\boldsymbol{\theta}_0$ defined in (4), except in specific circumstances stated further on.

Considering that the AGMWM does not necessarily target the quantity of interest $\boldsymbol{\theta}_0$, we now study the AWV estimator whose objective function appears closer to the form of the criterion in (4). Indeed, the AWV estimator targets the desired quantity as stated in the following theorem.

THEOREM 2: *Under Assumptions A to C and letting* $K, T \to \infty$, *we have that*

$$\|\hat{\boldsymbol{\theta}}^{\dagger} - \boldsymbol{\theta}_0\| \xrightarrow{P} 0.$$

Proof of Theorem 2 is given in Appendix A. Theorem 2 therefore shows that the AWV targets the desired quantity and is therefore preferable over the AGMWM if one aims at minimizing the criterion in (4).

The third estimator that we would need to study is the MS-GMWM. However, the following proposition underlines how the two estimators (AWV and MS-GMWM) are actually the same estimator.

PROPOSITION 1: *Under Assumptions A and B, we have that*

$$\hat{\boldsymbol{\theta}}^{\dagger} = \hat{\boldsymbol{\theta}}\,.$$

Proof of Proposition 1 is given in Appendix A. Given Proposition 1, we do not need to study the properties of the MS-GMWM since they will be the same as those of the AWV. Considering this, having proved consistency of the AWV, let us now deliver the final property of the AWV which consists in its asymptotic distribution.

PROPOSITION 2: *Under Assumptions A to C and letting* $K, T \to \infty$, *we have that*

$$\sqrt{K}(\hat{\boldsymbol{\theta}}^{\dagger} - \boldsymbol{\theta}_0) \xrightarrow{\mathcal{D}} \mathcal{N}(\mathbf{0}, \boldsymbol{\Lambda}_0)\,,$$

where $\boldsymbol{\Lambda}_0 := \mathbf{H}(\boldsymbol{\theta}_0)^{-1}\mathbf{A}(\boldsymbol{\theta}_0)^{\intercal}\boldsymbol{\Omega}\bar{\mathbf{V}}\boldsymbol{\Omega}\mathbf{A}(\boldsymbol{\theta}_0)\mathbf{H}(\boldsymbol{\theta}_0)^{-1}$ *is the asymptotic covariance matrix, with* $\bar{\mathbf{V}} := \mathbb{E}[\mathbf{V}_i]$.

As a consequence of Proposition 1 and 2, we can also state the following corollary.

COROLLARY 1: *Under Assumptions A to C and letting* $K, T \to \infty$, *we have that*

$$\sqrt{K}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0) \xrightarrow{\mathcal{D}} \mathcal{N}(\mathbf{0}, \boldsymbol{\Lambda})\,.$$

We omit the proof of this corollary since it is a direct consequence of Proposition 1. We conclude this section by delivering one final result which states the case under which the AGMWM actually targets the desired quantity $\boldsymbol{\theta}_0$. This result is provided in the following proposition where $\boldsymbol{W}$ denotes a non-singular matrix.

PROPOSITION 3: *If the theoretical WV is such that* $\boldsymbol{\nu}(\boldsymbol{\theta}) = \boldsymbol{W}\boldsymbol{\theta}$, *we have that*

$$\hat{\boldsymbol{\theta}}^{\circ} = \hat{\boldsymbol{\theta}}^{\dagger} = \hat{\boldsymbol{\theta}}\,.$$

Proof of Proposition 3 is given in Appendix A. This last result therefore states that, whenever the process underlying the signals delivers a theoretical WV which is linear in the parameters of interest, the parameter $\boldsymbol{\theta}_0$ can be estimated with any of the three solutions considered in this work, including the AGMWM. Examples of such processes are the white noise, quantization noise, random walk and drift, or a combination thereof. Algorithm 1 describes the procedure to perform multi-signal calibration using the AWV.

Algorithm 1: Multi-Signal Calibration

Inputs : A model structure F as described in (1) and identified via methods such as those in [4]–[7]; the stochastic error signals $X_t^{(i)}$ $(i = 1, \dots, K)$; weights w_i that must satisfy (5); and $\boldsymbol{\Omega}$ computed as in Sec. III-D.

Output: The estimated parameter $\hat{\boldsymbol{\theta}}^{\dagger} \in \mathbb{R}^p$.

Step 1: Compute the empirical WV with the MODWT estimator as defined in (3) on all observed replicates $i = 1, \dots, K$.

Step 2: Compute the average WV

$$\bar{\boldsymbol{\nu}} := \sum_{i=1}^{K} w_i \hat{\boldsymbol{\nu}}_i$$

over all replicates where $\hat{\boldsymbol{\nu}}_i$ is defined in Sec. III-A.

Step 3: Solve the following optimization problem:

$$\hat{\boldsymbol{\theta}}^{\dagger} = \underset{\boldsymbol{\theta} \in \boldsymbol{\Theta}}{\operatorname{argmin}} \|\bar{\boldsymbol{\nu}} - \boldsymbol{\nu}(\boldsymbol{\theta})\|^2_{\boldsymbol{\Omega}},$$

where $\boldsymbol{\nu}(\boldsymbol{\theta})$ denotes the WV implied by the model structure F.

D. Discussion

The choice of the matrix $\boldsymbol{\Omega}$ may not be completely obvious in the stochastic framework considered in this work. If one chooses an estimator $\widehat{\boldsymbol{\Omega}}$ then, in the standard single replicate setting, one can choose the inverse of the estimated covariance matrix of the empirical WV, or a diagonal matrix proportional to the latter. Since the matrix $\boldsymbol{\Omega}$ only affects the asymptotic efficiency of the resulting estimator and does not affect the consistency as long as it is positive definite, then one could choose the following matrix:

$$\widehat{\boldsymbol{\Omega}}_K := \sum_{i=1}^{K} w_i^2 \widehat{\boldsymbol{\Omega}}_i,$$

where $\widehat{\boldsymbol{\Omega}}_i$ represents the estimator for $\boldsymbol{\Omega}_0$ for the i^{th} replicate. The weighted average of the matrices that would be used on the individual replicates is indeed a valid choice and, for this reason, is what is going to be used in the next applied sections.

IV. SIMULATION STUDIES

In this section we provide further support to the results presented in Sec. III by studying the finite sample performance of the suggested approaches. To do so we mimic the standard scenario where the inertial sensors have been calibrated in a static setting and all deterministic components of their error signals have been removed leaving only the stochastic

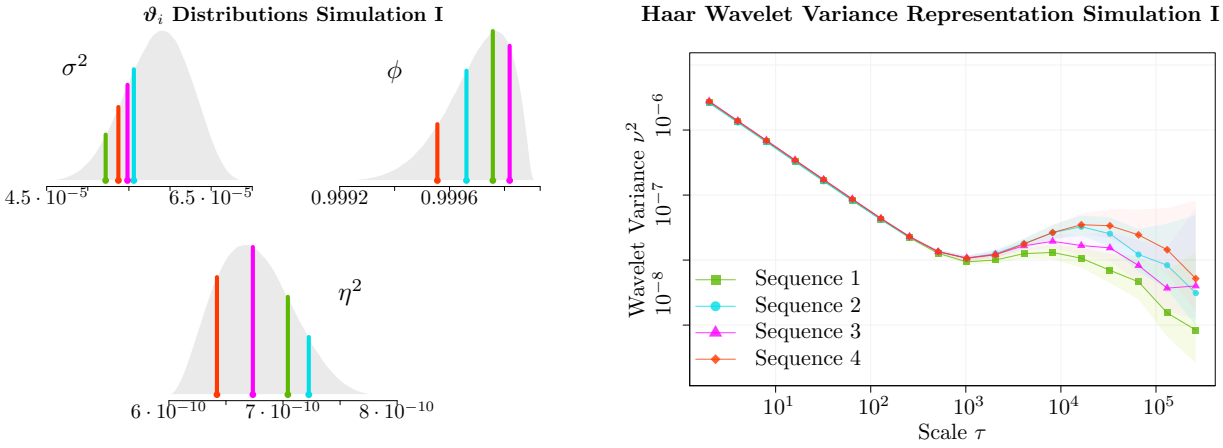

FIG. 2: *Left: Marginal densities of the internal sensor model G for the parameters σ_i^2, ϕ_i and η_i^2 considered in the Simulation I setting (WN + AR1) with horizontal colored lines representing four randomly selected values from each density. Right: WV plots and 95% confidence intervals for the empirical WV of the signals generated by the parameter values selected from the respective densities in the top part (each color in the top part corresponds to the color of the WV in the bottom part). Units on axes are not provided given that the WV plot is based on simulated data.*

component (X_t). Therefore in this section we only study the performance of the considered methods (including the AWV represented in Algorithm 1) for parameter estimation of the stochastic model F, whereas their impact on navigation will be studied in Sec. V. More specifically, based on the previous results we only compare two of the considered solutions, namely the AGMWM and AWV (since the MS-GMWM is equivalent to the latter). To do so we perform simulation studies based on composite stochastic processes that often characterize the stochastic signals from inertial sensor measurements. The first is a relatively common example consisting in the sum of a White Noise (WN) process with a first-order AutoRegressive (AR1) process (the latter consisting in a re-parametrization of a Gauss-Markov process), while the second consists in a sum of these two processes with the addition of a Random Walk (RW). In this second simulation setting, we therefore also consider the presence of non-stationary processes in the error signals also commonly found in stochastic signal calibration.

In order to generate settings that closely resemble the WV plots that are observed in stochastic calibration sessions, we choose to represent the internal sensor model G through independent and rescaled Beta distributions (i.e., each element of the parameter vector $\boldsymbol{\vartheta}_i$ comes from a separate rescaled Beta distribution). In addition, we choose to study the estimators in a setting where we observe $K = 6$ replicates which reasonably mimics a realistic calibration process with different runs. All replicates have the same length, i.e., $T_i = T = 10^6$ (for all i), which allows us to choose $J = 13$ scales of WV issued from the MODWT. Moreover, we choose $\boldsymbol{\Omega}$ by taking the average of the individual matrices for each replicate as discussed at the end of Sec. III. We repeat this setting $B = 500$ times to investigate the empirical distribution of the estimators studied. Finally, to be able to understand if the estimators are targeting the correct values, we compute the value $\boldsymbol{\theta}_0$ via numerical simulations by minimizing $Q(\boldsymbol{\theta})$ given in (4) based on 10^3 values of $\boldsymbol{\vartheta}_i$ randomly generated

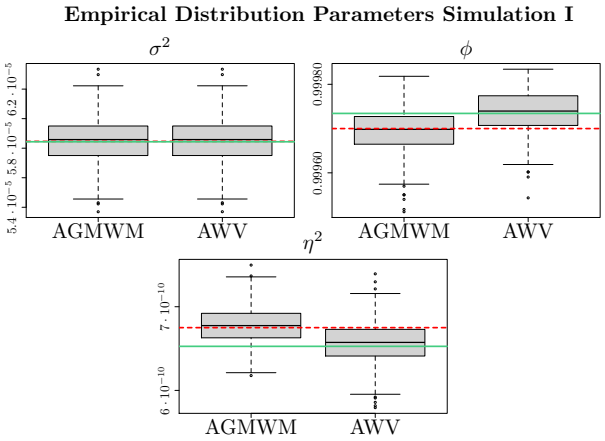

FIG. 3: *Empirical distributions of the AGMWM (left boxplot) and AWV (right boxplot) for the parameters of the stochastic error model (WN + AR1) of Simulation I ($K = 6$ and $T = 10^6$). The red dashed line represents the parameter value $\boldsymbol{\theta}^\circ$, while the full green line represents $\boldsymbol{\theta}_0$.*

from the chosen internal sensor model G (this allows us to well approximate the true value of $\boldsymbol{\theta}_0$), while $\boldsymbol{\theta}^\circ$ is computed for each element of $\boldsymbol{\vartheta}_i$ based on its corresponding distribution. In our simulation settings and in the case study in the subsequent section, we consider combinations of the following processes (components) for the model F:

- White Noise (WN): $N_t^{(i)} \sim \mathcal{F}(0, \sigma_i^2)$,
- First Order Auto-Regressive process (AR1): $N_t^{(i)} = \phi_i N_{t-1}^{(i)} + \varepsilon_t^{(i)}$ where $\varepsilon_t^{(i)} \sim \mathcal{F}(0, \eta_i^2)$,
- Random Walk (RW): $N_t^{(i)} = N_{t-1}^{(i)} + \iota_t^{(i)}$ where $\iota_t^{(i)} \sim \mathcal{F}(0, \gamma_i^2)$.

In the above cases, we let $\mathcal{F}$ represent any stationary probability distribution (e.g. Gaussian) with null expectation and an

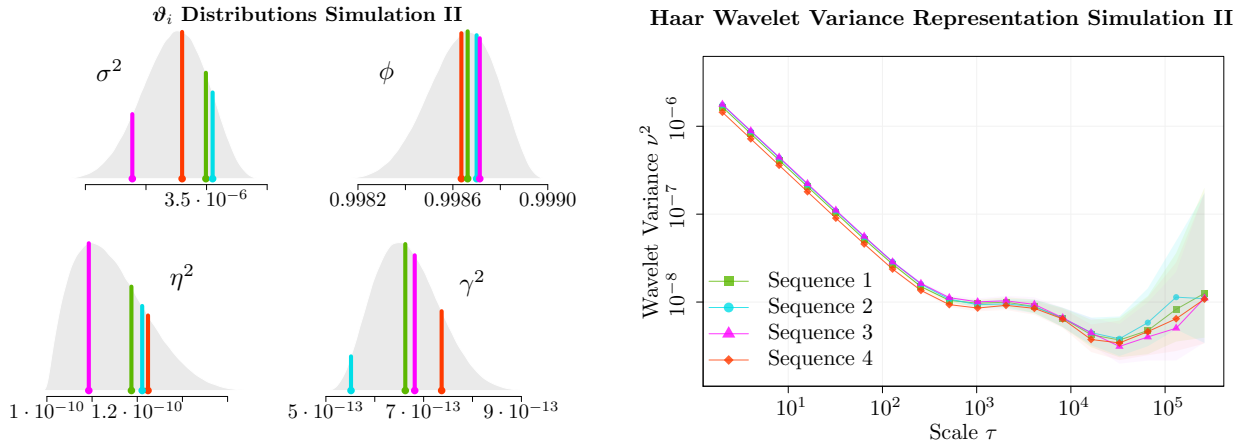

FIG. 4: *Left: Marginal densities of the internal sensor model G for the parameters σ_i^2, ϕ_i, η_i^2 and γ_i^2 considered in the Simulation II setting (WN + AR1 + RW) with horizontal colored lines representing four randomly selected values from each density. Right: WV plots and 95% confidence intervals for the empirical WV of the signals generated by the parameter values selected from the respective densities in the top part (each color in the top part corresponds to the color of the WV in the bottom part). Units on axes are not provided given that the WV plot is based on simulated data.*

unspecified variance parameter.

A. Simulation I

For the first simulation, the parameter vector for the i^{th} replicate is defined as follows $\boldsymbol{\vartheta}_i := [\sigma_i^2, \phi_i, \eta_i^2]$, where σ_i^2 represents the WN parameter, ϕ_i is the autoregressive parameter of the AR1, and η_i^2 is the innovation variance parameter of the AR1. In the near-stationary setting, we therefore have that $\boldsymbol{\vartheta}_i \sim G$, and, as discussed earlier, for these simulations we choose G to be a Beta distribution with different parameter values since this distribution lies within finite bounds (i.e. $[0, 1]$ support) and can take on different asymmetric forms depending on its parameters (see e.g. left plot in Fig. 4). Moreover, we scale each distribution for it to be consistent with the realistic range of values that the parameters of each process can take. Hence $\boldsymbol{\vartheta}_i$ is composed by

- $\sigma_i^2 = 4 \cdot 10^{-5} + Y_i^{(1)}(7 \cdot 10^{w-5} - 4 \cdot 10^{-5})$, where $Y_i^{(1)} \sim$ Beta $(8, 5)$,
- $\phi_i = 9.99 \cdot 10^{-1} + Y_i^{(2)}(9.999 \cdot 10^{-1} - 9.99 \cdot 10^{-1})$, where $Y_i^{(2)} \sim$ Beta$(7, 2)$,
- $\eta_i^2 = 6 \cdot 10^{-10} + Y_i^{(3)}(8 \cdot 10^{-10} - 6 \cdot 10^{-10})$, where $Y_i^{(3)} \sim$ Beta$(3, 5)$.

with $Y_i^{(j)}, j = 1, 2, 3$ therefore representing a variable from a Beta distribution used to mimic the random nature of the parameter values as postulated by the near-stationary framework, which is then multiplied by a scaling constant that projects its range onto realistic values for the process parameters. An insight into the described simulation setting is given in Fig. 2 where in the left part we can observe the rescaled Beta density functions (grey surfaces) from which we generate the respective parameter values that compose $\boldsymbol{\vartheta}_i$. Hence, the internal sensor model G is the multivariate distribution composed of independent variables σ_i^2, ϕ_i and η_i^2. The vertical colored lines represent randomly sampled values for the parameters following their respective distributions where common colors indicate those values that were generated jointly to deliver four different values of $\boldsymbol{\vartheta}_i$. These colors are then used to represent the empirical WV computed on signals generated from each value of $\boldsymbol{\vartheta}_i$ which can be seen in the right part of Fig. 2. We can notice how the different WVs are extremely close at the first scales and then differ at the larger scales. This plot is very similar to those seen in many applied settings as shown in Sec. V.

Having described how the data are generated, we now apply the procedure of Algorithm 1 to compute the AWV estimator and use (6) to compute the AGMWM estimator. When applying the estimators to the setting described above, we observe the results shown in Fig. 3. The red dashed line represents the true value of $\boldsymbol{\theta}^\circ$ and the full green line represents the (approximated) value of interest $\boldsymbol{\theta}_0$. The boxplots represent the empirical distribution of the estimated parameter values for the AGMWM (left boxplot) and AWV (right boxplot) respectively. While all boxplots appear to support the results on asymptotic normality of the estimators derived in Sec. III, it can be observed that the corresponding elements of $\boldsymbol{\theta}^\circ$ and $\boldsymbol{\theta}_0$ appear to differ (especially for the AR1 process which is non-linear in the WV). As a result of these differences, it is also obvious to detect how the two estimators target these different quantities since the AGMWM is centered around the red dashed line ($\boldsymbol{\theta}^\circ$) and the AWV around the full green line ($\boldsymbol{\theta}_0$). This therefore supports the consistency results in Sec. III which indeed state that these estimators target these respective quantities.

B. Simulation II

As mentioned at the start of this section, we perform a second simulation study in a similar way to the first one but, in this case, we add a RW process to the other two. This implies that the generated signals are non-stationary which is in fact

the case for many stochastic error signals issued from inertial calibration sessions. For this simulation, we have that $\boldsymbol{\vartheta}_i := [\sigma_i^2, \phi_i, \eta_i^2, \gamma_i^2]$ where, in addition to the parameters specified in the previous simulation, γ_i^2 represents the parameter of the RW process. The internal sensor model is composed of the following random parameter distributions:

- $\sigma_i^2 = 2 \cdot 10^{-6} + Y_i^{(1)}(4 \cdot 10^{-6} - 2 \cdot 10^{-6})$, where $Y_i^{(1)} \sim \text{Beta}(8, 5)$,
- $\phi_i = 9.98 \cdot 10^{-1} + Y_i^{(2)}(9.99 \cdot 10^{-1} - 9.98 \cdot 10^{-1})$, where $Y_i^{(2)} \sim \text{Beta}(7, 4)$,
- $\eta_i^2 = 1 \cdot 10^{-10} + Y_i^{(3)}(1.5 \cdot 10^{-10} - 1 \cdot 10^{-10})$, where $Y_i^{(3)} \sim \text{Beta}(3, 5)$;
- $\gamma^2 = 0.5 \cdot 10^{-12} + Y_i^{(4)}(1 \cdot 10^{-12} - 0.5 \cdot 10^{-12})$, where $Y_i^{(4)} \sim \text{Beta}(4, 8)$;

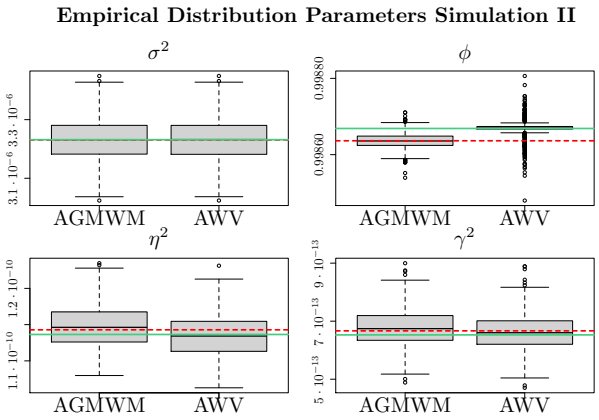

FIG. 5: *Empirical distributions of the AGMWM (left boxplot) and AWV (right boxplot) for the parameters of the stochastic error model (WN + AR1 + RW) of Simulation II ($K = 6$ and $T = 10^6$). The red dashed line represents the parameter value $\boldsymbol{\theta}^\circ$, while the full green line represents $\boldsymbol{\theta}_0$.*

Again, to give a visual support to the setting of this simulation, we provide an example of the parameter distributions (grey areas) along with four randomly sampled values for $\boldsymbol{\vartheta}_i$ represented by the four different colors in Fig. 4 left panels. Also in this case, it is possible to notice how the empirical WV generated from these different parameter values differ across the scales and, it can also be seen how some can be significantly different from the others at the first scales as highlighted by the non-overlapping confidence intervals of the respective WVs (shaded areas in the WV plot). In a similar manner to the first simulation we represent the results when applying the two considered estimators to this near-stationary setting. These results, represented in Fig. 5, confirm the conclusions made in the first simulation where both estimators appear normally distributed and both target their respective values of reference, i.e., $\boldsymbol{\theta}^\circ$ for the AGMWM and $\boldsymbol{\theta}_0$ for the AWV. Having given empirical support to the conclusions made in Sec. III, we now study how these conclusions deliver advantages in applied cases. In the next section, we therefore study the results in terms of navigation

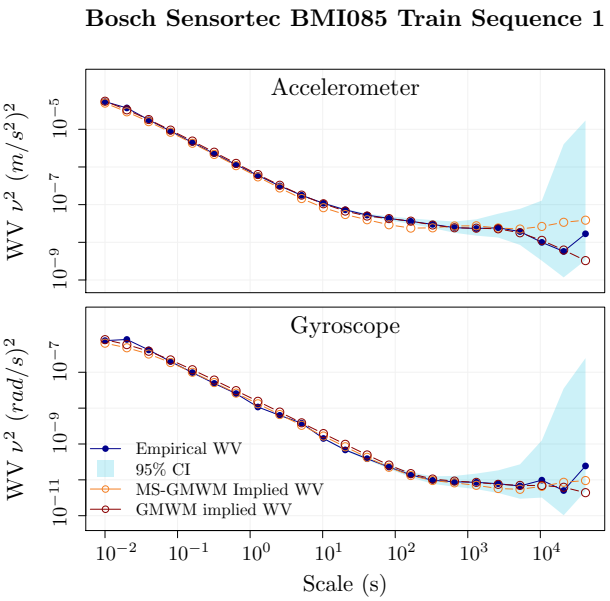

FIG. 6: *Empirical WV (blue doted line) for the first training sequence of accelerometer and gyroscope of a Bosch Sensortec BMI085 6-Axis IMU and their respective 95% confidence intervals (blue shaded area). Red dotted lines represent the implied WV from the individual solution of the GMWM on this first sequence, while orange dotted lines represent the WV implied by the MS-GMWM computed trough the AWV.*

performance when using the AWV estimator which targets the value of interest $\boldsymbol{\theta}_0$.

V. CASE STUDY - IMPACT ON NAVIGATION

The purpose of this section is to compare how navigation performances change when estimating stochastic models for the inertial sensors using a single replicate of the calibration data (as it is currently done), based on the GMWM, or using all replicates jointly based on the AWV estimator put forward in this work.

We collect static measurements from a Bosch Sensortec BMI085 6-Axis IMU [1], a low-cost MEMs IMU (< 5 USD per unit, when purchased in volumes) for navigation applications, e.g., in UAVs. Such an inertial module combines a 3-axis gyroscope and a 3-axis accelerometer. We collect $K = 16$ replicates $X_t^{(i)}$ $(i = 1, \ldots, K)$ of sensor data in static conditions at 20 °C in a temperature controlled chamber, each one lasting 12 hours. Since the sensor is static, the acquired data consists of samples of the noise processes only. The sensor runs at a frequency of 200 Hz, thus each error signal contains approximately 8.5 million sample points. Such long signals are of increasing interest even for low-cost MEMs IMUs since these devices are now employed in navigation and mapping, e.g. in UAVs, for longer and longer time frames [29] and the proper characterization of their stochastic properties is critical to determine an accurate orientation solution even when GPS

[1]The Bosch Sensortec BMI085: https://www.bosch-sensortec.com/products/motion-sensors/imus/bmi085/

observations are available [30]. In the following paragraphs, we focus on the error signals from the X-axis gyroscope and accelerometer.

To identify the error process we visually analyse the empirical WV of eight sequences that we consider for training purposes (i.e., used to estimate the model parameters), while leaving the remaining eight for validation, as discussed later on. The empirical WV of the training sequences are shown in Fig. 1. We observe that the considered devices are characterized by a non-negligible bias-instability, as it can be seen from the relatively flat part of the WV at the larger scales. This behaviour is common in low-cost inertial sensors and it is typically modeled with a sum of first order autoregressive processes (AR1), or equivalently, first order Gauss-Markov processes, as suggested for example in [11], [16]. We find that three AR1 processes are well suited to model each training sequence for the gyroscopes, and four for the accelerometers, respectively. We note that in both cases one of such AR1 processes always has a very short correlation time, far smaller than 1 s. This process models the intrinsic bandwidth limitation of the sensor (visible in the elbow at the first two scales of the WV) and is typically replaced with a white noise (an Angular/Velocity Random Walk) in practice. We estimate one model separately on each sequence in the training set, obtaining models $\mathcal{M}_i$, with $i \in [1, ..., 8]$. Next, we apply the AWV method proposed in this work employing all eight training sequences together, obtaining the model denoted as $\mathcal{M}_{\text{MS}}$. The estimated training models $\mathcal{M}_i$ appear to adequately fit the empirical WV of their respective training sequence, thus supporting the choice of the general model (an example consisting in the first training sequence is provided in Fig. 6). The fits for each sequence are given in Appendix B Fig. B.1 and B.2.

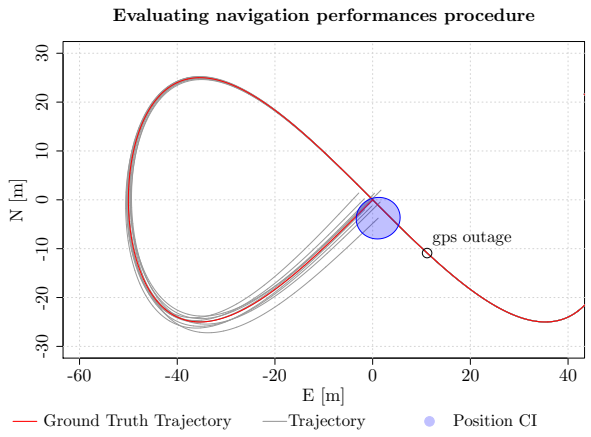

FIG. 7: *Procedure for evaluating navigation performances. An EKF estimates the trajectory of an UAV in* 250 *Monte-Carlo runs (a few of them are represented through the thin dark lines). The GNSS position and velocity information are no longer available after the marked point (GPS outage). The position and orientation error and the coverage of the uncertainty of the navigation states (blue circle), as estimated with by the EKF, are assessed based on the ground truth trajectory (red line) during the last* 15*s of the GNSS outage period.*

Each fitted model lies within the confidence intervals of the empirical WV. Considering these representations, it is straightforward to detect differences in the models fit to the signals via the individual and joint approaches. Given this, in order to confirm whether to use a single replicate or a multi-signal approach we perform the near-stationarity test put forward in [23], with its respective null and alternative hypotheses being $\mathcal{H}_0 : \boldsymbol{\vartheta}_i = \boldsymbol{\theta}_0, \forall i$ and $\mathcal{H}_a : \mathcal{H}_0$ is false. To perform this test we simulate 100 bootstrap replicates under the estimated $F_{\hat{\boldsymbol{\theta}}^\dagger}$ which, keeping in mind the discrete nature of the bootstrapped test statistic, gives us a zero p-value thereby allowing us to reject the null hypothesis that all replicates are issued from the same data-generating process with $\boldsymbol{\vartheta}_i = \boldsymbol{\theta}_0$ for all i (i.e. G is a Dirac point mass distribution). The estimated parameters of the models $\mathcal{M}_i$ are included in Appendix C, Fig. C.1 and C.2. We note that a substantial variability can be observed within the latter fits and that, as expected, the parameters obtained with the AWV method do not correspond to their mean.

We investigate the navigation performance on the $8 + 1$ different models. The estimated stochastic models are used to configure an Extended Kalman Filter (EKF) for INS/GNSS navigation [1]. This filter fuses inertial and GNSS readings, leveraging on the provided stochastic models, to estimate the vehicle navigation states (position, velocity and orientation). It allows us to compare the performance of the different models available for the inertial sensor in terms of position and orientation errors as well as consistency of the confidence intervals for the navigation states within a realistic navigation scenario. We consider a ground-truth trajectory typical of a small fixed-wing Unmanned Aerial Vehicle (UAV) performing an aerial mapping mission. A 30s GNSS outage period is considered after 9.5 minutes. All the true kinematic properties of the sensors are known (position, velocity, etc.) from the reference trajectory and they are used to generate synthetic, noise-free sensor readings for both the inertial and the GNSS sensors. Realistic noisy readings are then generated for the inertial sensors by adding samples from the noise replicates collected during static acquisitions to the synthetic noise-free readings. Here, we employ the remaining eight static data sequences we collected and that were never used in the previously described stochastic calibration step. As for the GNSS readings, the added noise is WN with standard deviation 2.5 cm, which corresponds to the assumed uncertainty carrier-phase differential of GNSS typically employed in mapping missions. While following the common emulation approach adopted in the inertial sensor calibration literature this far, this simulation set-up may not entirely reflect the real-world setting where sensor measurement errors should ideally be observed under varying external conditions (e.g. dynamics). Nevertheless, given the many current theoretical and practical challenges of measuring and modelling these errors in the latter context, the methods put forward in this work aim to better approximate the more realistic (dynamic) context and therefore to reduce the potential navigation errors due to the stochastic measurement errors.

With the above in mind, a forward navigation solution is computed using an EKF from the noisy sensor readings. We consider $9 \times 8 = 72$ different cases in which the EKF

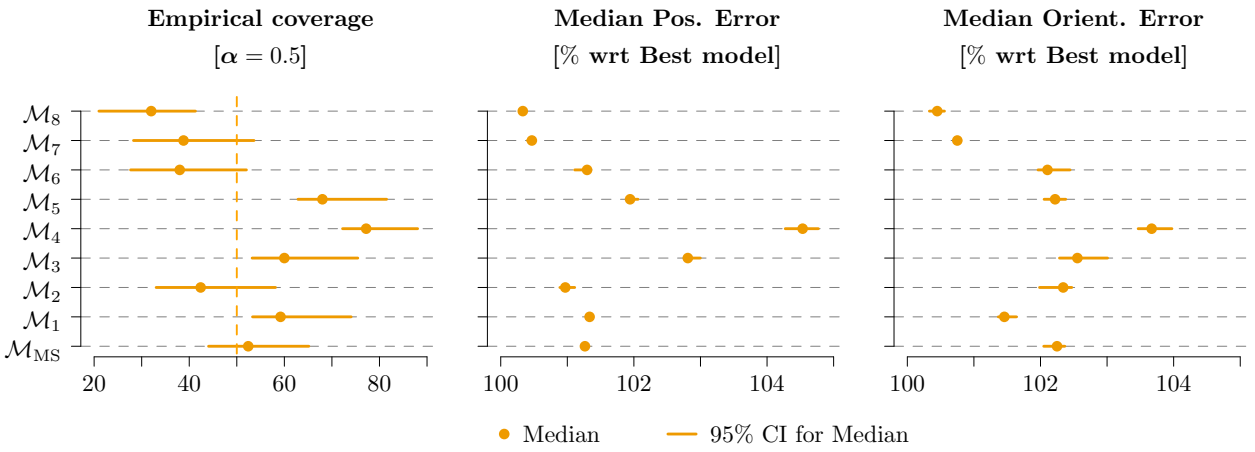

FIG. 8: *Empirical coverage of the 50% confidence intervals derived from the EKF covariance matrices, median position and orientation errors achieved by using each estimated model to predict the error on all replicates, $\mathcal{M}_i$ represents the model estimated on replicate i and $\mathcal{M}_{MS}$ represents the model estimated via the multi-signal AWV. The results are expressed in percentage with respect the best performing model on one specific static acquisition.*

is configured to use one of the $8 + 1 = 9$ model sets fitted on the static acquisition replicates, while the noise data corrupting inertial readings comes from one of the eight different static acquisition sequences kept for validation, each time considering a different, continuous chunk of data.

The 250 solutions for each case are aggregated and compared in terms of relative position and orientation error and consistency of the confidence intervals computed by the EKF: we compute 50% confidence intervals (approximately corresponding to the common choice of $\pm\sigma$ intervals) from the navigation state covariance matrix estimated by the EKF and we count how many times the true navigation states (from the reference trajectory) fall within such confidence intervals. Note that it is equivalent to check whether the Average Normalised Estimation Error Squared (ANEES), as defined for example in [31, Chapter 3.7.4], falls within its expected bounds, and it allows to quantify whether the employed stochastic models for the inertial errors lead to a over- or under-confident estimation of the navigation state uncertainty. The position and orientation error and the coverage metrics are evaluated each 0.5s in the last 15s of the GNSS outage period to better highlight their evolution when the navigation filter works in standalone mode, e.g., relying only on inertial data (represented in Fig. 7). The results are presented in Fig. 8.

It is possible to see that the differences in position and orientation error, computed in percentage with respect to the best performing model, vary up to 5 % depending on which stochastic model is selected for the inertial sensor. If using a rough reference level for improvement achieved through deterministic calibration, in [32] they report gains up to roughly 37 % although these quantities are not absolutely comparable since additional procedures were performed in the latter study to improve navigation performance (in addition to other comparative problems). As a consequence these differences in improvement may seem small but, among many practitioners, attitude quality improvement is deemed proportional to the square (or even the cube) of the IMU size and weight (as well as cost). In addition to navigation accuracy, the importance in stochastic calibration consists mainly in uncertainty quantification which can be seen through the differences in coverage which are much more significant: when computing a confidence interval for position and orientation with level $\alpha = 0.5$ (50%), we find that the empirical coverage of certain models fit on a single sequence, e.g., $\mathcal{M}_4$ and $\mathcal{M}_8$, is as low as 10 % or as high as 90 % in some cases. This implies that, when configured with such models, the EKF is largely over- or under-confident in the estimation of the uncertainty of the navigation states. Even though the actual errors in such states remain relatively small, the quantification of their uncertainty is substantially unreliable which prevents, for example, proper decision making in safety-critical navigation applications, or consistent information fusion in more complex scenarios such as simultaneous localisation and mapping, where further sensor information (e.g., from cameras) need to be taken into account. On the other hand, the model estimated with one of the methods put forward and studied in this work, $\mathcal{M}_{\text{MS}}$, achieves extremely similar position and orientation performances with respect to the best individual model, while at the same time providing a reliable and correct uncertainty quantification of the position and orientation estimates. We remark that *by chance* one single sequence may lead to the estimation of a stochastic model which performs well in practice, but at the same time the opposite may hold, for example if training sequences 3, 4, or 8 were to be selected. These results indicate that the AWV (or a multi-signal method) can deliver a more robust (stable) estimation of the stochastic models that underlie inertial sensor measurement errors, compensating for the intrinsic variability of the single realizations of calibration data.

VI. CONCLUSIONS

In this work, we studied methods and delivered further evidence for the need of a multi-signal approach when dealing

with inertial sensor calibration. Indeed, in many practical settings, one can observe a near-stationary behavior of replicate IMU stochastic error signals which needs to be taken into account when performing estimation for model selection and construction of accurate navigation filters. Having compared different existing and new approaches to address this problem, we determined their asymptotic properties and their common features which were empirically supported in controlled simulation settings as well as in applied case study scenarios. In the latter case, this work also highlighted how the use of a single replicate to perform stochastic calibration may not be a suitable choice and confirmed that a multi-signal solution is the most appropriate in such settings. As a result of this work, it is now possible to select the most appropriate multi-signal calibration approach according to the goal of interest and consequently achieve improved navigation performance both in terms of accuracy as well as in terms of uncertainty quantification during navigation. Finally, this study can extend to all approaches based on moment-matching (e.g. Generalized Methods of Moments) beyond the WV and IMU calibration.

ACKNOWLEDGMENT

This work was supported in part by the SNSF Grant $\#100018-182582$, in part by the SNSF Professorships Grant $\#176843$ and by the Innosuisse-Boomerang Grant $\#37308.1$ IP-ENG. Roberto Molinari was also partially supported by NSF grant SES-2150615.

REFERENCES

[1] D. Titterton, J. L. Weston, and J. Weston, *Strapdown inertial navigation technology*. IET, 2004, vol. 17.

[2] G. Huang, "Visual-inertial navigation: A concise review," in *2019 International Conference on Robotics and Automation (ICRA)*. IEEE, 2019, pp. 9572–9582.

[3] J. H. Wall, D. M. Bevly *et al.*, "Characterization of various IMU error sources and the effect on navigation performance," in *Proceedings of the 18th international technical meeting of the satellite division of the institute of navigation (ION GNSS 2005)*, 2005, pp. 967–978.

[4] N. El-Sheimy, H. Hou, and X. Niu, "Analysis and modeling of inertial sensors using allan variance," *IEEE Transactions on instrumentation and measurement*, vol. 57, no. 1, pp. 140–149, 2007.

[5] O. Stepanov and A. Motorin, "Problem-oriented approach to identification of sensor error models and its application to navigation data processing," *IFAC-PapersOnLine*, vol. 50, no. 1, pp. 2830–2835, 2017.

[6] J. J. Balamuta, R. Molinari, S. Guerrier, and W. Yang, "A computationally efficient framework for automatic inertial sensor calibration," *IEEE Sensors Journal*, vol. 18, no. 4, pp. 1636–1646, 2017.

[7] A. Radi, G. Bakalli, N. El-Sheimy, S. Guerrier, and R. Molinari, "An automatic calibration approach for the stochastic parameters of inertial sensors," in *Proceedings of the 30th International Technical Meeting of The Satellite Division of the Institute of Navigation (ION GNSS+ 2017)*, 2017, pp. 3053–3060.

[8] S. Guerrier, R. Molinari, and Y. Stebler, "Theoretical limitations of allan variance-based regression for time series model estimation," *IEEE Signal Processing Letters*, vol. 23, no. 5, pp. 597–601, 2016.

[9] ——, "Wavelet-based improvements for inertial sensor error modeling," *IEEE Transactions on Instrumentation and Measurement*, vol. 65, no. 12, pp. 2693–2700, 2016.

[10] R. Allen and D. Chang, "Performance testing of the systron donner quartz gyro," *Jpl Engineering Memorandum, EM*, pp. 343–1297, 1993.

[11] I. Board, "IEEE standard specification format guide and test procedure for single-axis interferometric fiber optic gyros," *IEEE Std*, pp. 952–1997, 1998.

[12] Y. Yuksel, N. El-Sheimy, and A. Noureldin, "Error modeling and characterization of environmental effects for low cost inertial MEMS units," in *Proceedings of IEEE/ION PLANS 2010*, 2010, pp. 598–612.

[13] B. Claus, "Multiscale statistical signal processing: identification of a multiscale AR process from a sample of an ordinary signal," *IEEE transactions on signal processing*, vol. 41, no. 12, pp. 3266–3274, 1993.

[14] R. Johansson, M. Verhaegen, and C. T. Chou, "Stochastic theory of continuous-time state-space identification," *IEEE Transactions on Signal Processing*, vol. 47, no. 1, pp. 41–51, 1999.

[15] J. Nikolic, P. Furgale, A. Melzer, and R. Siegwart, "Maximum likelihood identification of inertial sensor noise model parameters," *IEEE Sensors Journal*, vol. 16, no. 1, pp. 163–176, 2015.

[16] Y. Yuksel and H. B. Kaygisiz, "Notes on stochastic errors of low cost mems inertial units," *línea*, 2011.

[17] O. Stepanov and A. Motorin, "Performance criteria for the identification of inertial sensor error models," *Sensors*, vol. 19, no. 9, p. 1997, 2019.

[18] D. W. Allan, "Statistics of atomic frequency standards," *Proceedings of the IEEE*, vol. 54, no. 2, pp. 221–230, 1966.

[19] ——, "Historicity, strengths, and weaknesses of Allan variances and their general applications," *Gyroscopy and Navigation*, vol. 7, no. 1, pp. 1–17, 2016.

[20] S. Guerrier, J. Jurado, M. Khaghani, G. Bakalli, M. Karemera, R. Molinari, S. Orso, J. Raquet, C. Schubert, J. Skaloud *et al.*, "Wavelet-based moment-matching techniques for inertial sensor calibration," *IEEE Transactions on Instrumentation and Measurement*, 2020.

[21] Y. Stebler, S. Guerrier, J. Skaloud, and M.-P. Victoria-Feser, "Constrained expectation-maximization algorithm for stochastic inertial error modeling: Study of feasibility," *Measurement Science and Technology*, vol. 22, no. 8, p. 085204, 2011.

[22] S. Guerrier, J. Skaloud, Y. Stebler, and M. Victoria-Feser, "Wavelet-variance-based estimation for composite stochastic processes," *Journal of the American Statistical Association*, vol. 108, no. 503, 2013.

[23] G. Bakalli, A. Radi, N. El-Sheimy, R. Molinari, and S. Guerrier, "A computational multivariate-based technique for inertial sensor calibration," in *Proceedings of the 30th International Technical Meeting of The Satellite Division of the Institute of Navigation (ION GNSS+ 2017)*, 2017, pp. 3028–3038.

[24] A. Radi, G. Bakalli, S. Guerrier, N. El-Sheimy, A. B. Sesay, and R. Molinari, "A multisignal wavelet variance-based framework for inertial sensor stochastic error modeling," *IEEE Transactions on Instrumentation and Measurement*, vol. 68, no. 12, pp. 4924–4936, 2019.

[25] S. Guerrier, R. Molinari, M.-P. Victoria-Feser, and H. Xu, "Robust two-step wavelet-based inference for time series models," *Journal of the American Statistical Association*, pp. 1–50, 2021.

[26] D. B. Percival and A. T. Walden, *Wavelet methods for time series analysis*. Cambridge university press, 2000, vol. 4.

[27] N. F. Zhang, "Allan variance of time series models for measurement data," *Metrologia*, vol. 45, no. 5, p. 549, 2008.

[28] A. Serroukh, A. T. Walden, and D. B. Percival, "Statistical properties and uses of the wavelet variance estimator for the scale analysis of time series," *Journal of the American Statistical Association*, vol. 95, no. 449, pp. 184–196, 2000.

[29] G. Jouvet, T. Stastny, P. Oettershagen, E. van Dongen, M. Hugentobler, T. Mantel, A. Melzer, Y. Weidmann, M. Funk, and R. Siegwart, "Sun2ice: Monitoring calving glaciers from solar-powered uavs," in *EGU General Assembly Conference Abstracts*, 2018, p. 11682.

[30] D. A. Cucci, M. Rehak, and J. Skaloud, "Bundle adjustment with raw inertial observations in UAV applications," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 130, pp. 1–12, 2017.

[31] Y. Bar-Shalom, X. R. Li, and T. Kirubarajan, *Estimation with applications to tracking and navigation: theory algorithms and software*. John Wiley & Sons, 2004.

[32] P. Clausen, "Calibration aspects of INS navigation," p. 237, 2019. [Online]. Available: http://infoscience.epfl.ch/record/264793